

# Web-based Supplementary Materials for “Latent-model Robustness in Joint Models for a Primary Endpoint and a Longitudinal Process”

Xianzheng Huang,<sup>1,\*</sup> Leonard A. Stefanski,<sup>2</sup> and Marie Davidian<sup>2</sup>

<sup>1</sup> Department of Statistics, University of South Carolina,

Columbia, South Carolina 29208, U.S.A.

<sup>2</sup> Department of Statistics, North Carolina State University,

Raleigh, North Carolina 27695, U.S.A.

*\*email:* huang@stat.sc.edu

## Web Appendix A: Proof of Theorem 1

For brevity, the subject index  $i$  is dropped in the following argument. Denote by  $\widehat{\mathbf{X}}_m$  the ordinary least squares estimator for  $\mathbf{X}$ , i.e.,  $\widehat{\mathbf{X}}_m = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{W}_{m \times 1}$ . Assuming  $\sigma^2$  known and viewing  $\mathbf{X}$  as an unknown parameter,  $\widehat{\mathbf{X}}_m$  is a complete sufficient statistic for  $\mathbf{X}$ , and  $\widehat{\mathbf{X}}_m | \mathbf{X} \sim N_p\{\mathbf{X}, \sigma^2(\mathbf{D}^T \mathbf{D})^{-1}\}$ . Therefore, by the Factorization Theorem,

$$f_{\mathbf{W}|\mathbf{X}}(\mathbf{w}|\mathbf{x}; \sigma^2) = f_{\mathbf{W}|\widehat{\mathbf{x}}_m}(\mathbf{w}|\widehat{\mathbf{x}}_m; \sigma^2) f_{\widehat{\mathbf{x}}_m|\mathbf{X}}(\widehat{\mathbf{x}}_m|\mathbf{x}; \sigma^2),$$

where  $f_{\mathbf{W}|\widehat{\mathbf{x}}_m}(\mathbf{w}|\widehat{\mathbf{x}}_m; \sigma^2)$  is free of  $\mathbf{x}$ ; and  $f_{\widehat{\mathbf{x}}_m|\mathbf{X}}(\widehat{\mathbf{x}}_m|\mathbf{x}; \sigma^2) = |\mathbf{G}_m|^{-1} \phi\{\mathbf{G}_m^{-1}(\widehat{\mathbf{x}}_m - \mathbf{x})\}$ ,  $\phi(\cdot)$  is the density of the  $p$ -dimensional standard normal density, with  $\mathbf{G}_m \mathbf{G}_m^T = \sigma^2(\mathbf{D}^T \mathbf{D})^{-1}$ . It follows that the observed data density in equation (2) in the article can be rewritten as

$$\begin{aligned} f_{\mathbf{Y}, \mathbf{W}|\mathbf{H}}(\mathbf{y}, \mathbf{w}|\mathbf{h}; \Omega) &= \int f_{\mathbf{Y}|\mathbf{X}, \mathbf{H}}(\mathbf{y}|\mathbf{x}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{W}|\mathbf{X}}(\mathbf{w}|\mathbf{x}; \sigma^2) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\mathbf{x}|\mathbf{h}; \boldsymbol{\tau}^{(a)}) d\mathbf{x} \\ &= f_{\mathbf{W}|\widehat{\mathbf{x}}_m}(\mathbf{w}|\widehat{\mathbf{x}}_m; \sigma^2) \times \\ &\quad \int f_{\mathbf{Y}|\mathbf{X}, \mathbf{H}}(\mathbf{y}|\mathbf{x}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\widehat{\mathbf{x}}_m|\mathbf{X}}(\widehat{\mathbf{x}}_m|\mathbf{x}; \sigma^2) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\mathbf{x}|\mathbf{h}; \boldsymbol{\tau}^{(a)}) d\mathbf{x}. \end{aligned} \quad (\text{A.1})$$

Next consider the integral in expression (A.1), and specifically the difference

$$\begin{aligned} \widetilde{\Delta} &= \int f_{\mathbf{Y}|\mathbf{X}, \mathbf{H}}(\mathbf{y}|\mathbf{x}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\widehat{\mathbf{x}}_m|\mathbf{X}}(\widehat{\mathbf{x}}_m|\mathbf{x}; \sigma^2) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\mathbf{x}|\mathbf{h}; \boldsymbol{\tau}^{(a)}) d\mathbf{x} \\ &\quad - f_{\mathbf{Y}|\widehat{\mathbf{x}}_m, \mathbf{H}}(\mathbf{y}|\widehat{\mathbf{x}}_m, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\widehat{\mathbf{x}}_m|\mathbf{h}; \boldsymbol{\tau}^{(a)}). \end{aligned}$$

After the change of variable  $\mathbf{z} = \mathbf{G}_m^{-1}(\mathbf{x} - \widehat{\mathbf{x}}_m)$  and rearrangement of terms, we find that

$$\begin{aligned} \widetilde{\Delta} = \int \left\{ f_{\mathbf{Y}|\mathbf{X},\mathbf{H}}(\mathbf{y}|\widehat{\mathbf{x}}_m + \mathbf{G}_m\mathbf{z}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\widehat{\mathbf{x}}_m + \mathbf{G}_m\mathbf{z}|\mathbf{h}; \boldsymbol{\tau}^{(a)}) - \right. \\ \left. f_{\mathbf{Y}|\widehat{\mathbf{x}}_m, \mathbf{H}}(\mathbf{y}|\widehat{\mathbf{x}}_m, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\widehat{\mathbf{x}}_m|\mathbf{h}; \boldsymbol{\tau}^{(a)}) \right\} \phi(\mathbf{z}) d\mathbf{z}. \end{aligned} \quad (\text{A.2})$$

We formalize the condition that the longitudinal process information increases by assuming that the minimum eigenvalue of  $\mathbf{D}^T\mathbf{D}$  diverges to  $+\infty$ . In this case,  $\mathbf{G}_m \rightarrow \mathbf{0}_{p \times p}$ , and the integrand in (A.2) converges to zero. Thus, as the longitudinal process information increases without bound,  $\widetilde{\Delta} \rightarrow 0$  whenever integration and limit can be interchanged in (A.2). A sufficient condition for the interchange of limit and integration is that

$$|f_{\mathbf{Y}|\mathbf{X},\mathbf{H}}(\mathbf{y}|\mathbf{x}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\mathbf{x}|\mathbf{h}; \boldsymbol{\tau}^{(a)})| \leq M$$

for some positive constant  $M$  at each fixed  $\mathbf{y}$ , as then the integrand in (A.2) is bounded by an integrable function,  $2M\phi(\mathbf{z})$ , and the result follows via the Lebesgue Dominated Convergence Theorem. Finally, note that as  $\widetilde{\Delta} \rightarrow 0$ , the ratio of the expressions

$$f_{\mathbf{W}|\widehat{\mathbf{x}}_m}(\mathbf{w}|\widehat{\mathbf{x}}_m; \sigma^2) \int f_{\mathbf{Y}|\mathbf{X},\mathbf{H}}(\mathbf{y}|\mathbf{x}, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\widehat{\mathbf{x}}_m|\mathbf{X}}(\widehat{\mathbf{x}}_m|\mathbf{x}; \sigma^2) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\mathbf{x}|\mathbf{h}; \boldsymbol{\tau}^{(a)}) d\mathbf{x}, \quad (\text{A.3})$$

$$f_{\mathbf{W}|\widehat{\mathbf{x}}_m}(\mathbf{w}|\widehat{\mathbf{x}}_m; \sigma^2) f_{\mathbf{Y}|\widehat{\mathbf{x}}_m, \mathbf{H}}(\mathbf{y}|\widehat{\mathbf{x}}_m, \mathbf{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{X}|\mathbf{H}}^{(a)}(\widehat{\mathbf{x}}_m|\mathbf{h}; \boldsymbol{\tau}^{(a)}), \quad (\text{A.4})$$

approaches one. Therefore, the ratio of the density in (2), re-expressed in (A.3), over (A.4) also approaches one as the longitudinal process information increases.

## Web Appendix B: Derivations of Variance Estimators $\widehat{\nu}_1$ and $\widehat{\nu}_2$

Notations:

- For  $i = 1, \dots, n$ ,  $\mathbf{Q}_i$  is the observed data for subject  $i$ ;  $\mathbf{Q}_i^{(B)}$  is the  $B$  sets of  $\lambda$ -remeasured data for subject  $i$ .  $\mathbf{Q} = \{\mathbf{Q}_i, i = 1, \dots, n\}$ ;  $\mathbf{Q}^{(B)} = \{\mathbf{Q}_i^{(B)}, i = 1, \dots, n\}$ .
- $\boldsymbol{\Omega}$  denotes the  $d \times 1$  vector of all unknown parameters in the joint model.  
 $\boldsymbol{\Omega}_{-\sigma^2}$  denotes  $\boldsymbol{\Omega}$  excluding  $\sigma^2$ .

- $E[\boldsymbol{\psi}\{\mathbf{Q}_i; \boldsymbol{\Omega}_{-\sigma^2}(0), \sigma^2(0)\}] = \mathbf{0}$  uniquely defines  $\boldsymbol{\Omega}(0) = \{\boldsymbol{\Omega}_{-\sigma^2}(0)^T, \sigma^2(0)\}^T$ ;  
 $E[\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}_{-\sigma^2}(\lambda), \sigma^2(\lambda)\}] = \mathbf{0}$  uniquely defines  $\boldsymbol{\Omega}(\lambda) = \{\boldsymbol{\Omega}_{-\sigma^2}(\lambda)^T, \sigma^2(\lambda)\}^T$  for  $\lambda > 0$ , where the expectations are with respect to the true densities of  $\mathbf{Q}_i$  and  $\mathbf{Q}_i^{(B)}$ , respectively,  $\sigma^2(\lambda) = (1 + \lambda)\sigma^2(0)$ , and  $\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}\} = B^{-1} \sum_{b=1}^B \boldsymbol{\psi}\{\mathbf{Q}_{b,i}(\lambda); \boldsymbol{\Omega}\}$ .
- $\tilde{\boldsymbol{\Omega}}(0)$  solves  $\sum_{i=1}^n \boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega}) = \mathbf{0}$ ;  $\tilde{\boldsymbol{\Omega}}_B(\lambda)$  solves  $\sum_{i=1}^n \boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}\} = \mathbf{0}$ .
- $t_1^*(\lambda) = [\mathbf{T}_1(\lambda)]_{(k)}[\hat{\boldsymbol{\nu}}_1]_{(k)}^{-1/2}$ , where  $\mathbf{T}_1(\lambda) = n^{1/2} \left\{ \tilde{\boldsymbol{\Omega}}_{-\sigma^2}(\lambda) - \tilde{\boldsymbol{\Omega}}_{-\sigma^2}(0) \right\}$ .
- $t_2^*(\lambda) = [\mathbf{T}_2(\lambda)]_{(k)}[\hat{\boldsymbol{\nu}}_2]_{(k)}^{-1/2}$ , where  $\mathbf{T}_2(\lambda) = n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi}^{(B)}\left\{ \mathbf{Q}_i^{(B)}(\lambda); \tilde{\boldsymbol{\Omega}}_{-\sigma^2}(0), (1 + \lambda)\tilde{\sigma}^2(0) \right\}$ .
- For a positive definite matrix  $\boldsymbol{\Pi}$ , define  $\boldsymbol{\Pi}^{1/2}$  as the positive definite square root such that  $\boldsymbol{\Pi}^{1/2}(\boldsymbol{\Pi}^{1/2})^T = \boldsymbol{\Pi}$ , and  $\boldsymbol{\Pi}^{-1/2}$  as the inverse of  $\boldsymbol{\Pi}^{1/2}$ . Define  $\mathbf{T}_1^* = \hat{\boldsymbol{\nu}}_1^{-1/2} \mathbf{T}_1(\lambda)$  and  $\mathbf{T}_2^* = \hat{\boldsymbol{\nu}}_2^{-1/2} \mathbf{T}_2(\lambda)$ .
- Consider the null hypothesis,  $H_0 : \boldsymbol{\Omega}_{-\sigma^2}(\lambda) - \boldsymbol{\Omega}_{-\sigma^2}(0) = \mathbf{0}$ , and the contiguous alternative hypothesis,  $H_a : \boldsymbol{\Omega}_{-\sigma^2}(\lambda) - \boldsymbol{\Omega}_{-\sigma^2}(0) = n^{-1/2} \boldsymbol{\Delta}^*(\lambda)$ .

Define the following Hessian matrices and the associated empirical estimators.

$$\begin{aligned} \mathbf{A}_1\{\boldsymbol{\Omega}(0)\} &= E\left\{-\partial\boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega})/\partial\boldsymbol{\Omega}^T\right\} \Big|_{\boldsymbol{\Omega} = \boldsymbol{\Omega}(0)}; \\ \hat{\mathbf{A}}_1\{\mathbf{Q}; \tilde{\boldsymbol{\Omega}}(0)\} &= -n^{-1} \sum_{i=1}^n \partial\boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega})/\partial\boldsymbol{\Omega}^T \Big|_{\boldsymbol{\Omega} = \tilde{\boldsymbol{\Omega}}(0)}; \\ \mathbf{A}_2\{\boldsymbol{\Omega}(\lambda)\} &= E\left[-\partial\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}\}/\partial\boldsymbol{\Omega}^T\right] \Big|_{\boldsymbol{\Omega} = \boldsymbol{\Omega}(\lambda)}; \\ \hat{\mathbf{A}}_2\{\mathbf{Q}^{(B)}(\lambda); \tilde{\boldsymbol{\Omega}}(\lambda)\} &= -n^{-1} \sum_{i=1}^n \partial\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}\}/\partial\boldsymbol{\Omega}^T \Big|_{\boldsymbol{\Omega} = \tilde{\boldsymbol{\Omega}}(\lambda)}. \end{aligned}$$

Similarly denote by  $\mathbf{A}_2\{\boldsymbol{\Omega}_{-\sigma^2}(0), (1 + \lambda)\sigma^2(0)\}$  the expectation  $E[-\partial\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}\}/\partial\boldsymbol{\Omega}^T]$  evaluated at  $\boldsymbol{\Omega} = \{\boldsymbol{\Omega}_{-\sigma^2}(0)^T, (1 + \lambda)\sigma^2(0)\}^T$ . And define  $\hat{\mathbf{A}}_2\{\mathbf{Q}^{(B)}(\lambda); \tilde{\boldsymbol{\Omega}}_{-\sigma^2}(0), (1 + \lambda)\tilde{\sigma}^2(0)\}$  as the average  $-n^{-1} \sum_{i=1}^n \partial\boldsymbol{\psi}^{(B)}\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}\}/\partial\boldsymbol{\Omega}^T$  evaluated at  $\boldsymbol{\Omega} = \{\tilde{\boldsymbol{\Omega}}_{-\sigma^2}(0)^T, (1 + \lambda)\tilde{\sigma}^2(0)\}^T$ .

(I) An estimator for the variance-covariance matrix of  $\mathbf{T}_1(\lambda)$ ,  $\hat{\boldsymbol{\nu}}_1$ :

Using the influence function approximation, we have

$$n^{1/2} \left\{ \tilde{\boldsymbol{\Omega}}(\lambda) - \boldsymbol{\Omega}(\lambda) \right\} = n^{-1/2} \sum_{i=1}^n \mathbf{A}_2^{-1} \left\{ \boldsymbol{\Omega}(\lambda) \right\} \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) \right\} + \boldsymbol{o}_p(1), \quad (\text{B.1})$$

$$n^{1/2} \left\{ \tilde{\boldsymbol{\Omega}}(0) - \boldsymbol{\Omega}(0) \right\} = n^{-1/2} \sum_{i=1}^n \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\} + \boldsymbol{o}_p(1). \quad (\text{B.2})$$

Subtracting (B.2) from (B.1) yields

$$\begin{aligned} \mathbf{T}_1(\lambda) &= n^{1/2} \left\{ \boldsymbol{\Omega}(\lambda) - \boldsymbol{\Omega}(0) \right\} + \\ &\quad n^{-1/2} \sum_{i=1}^n \left[ \mathbf{A}_2^{-1} \left\{ \boldsymbol{\Omega}(\lambda) \right\} \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) - \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\} \right] + \boldsymbol{o}_p(1) \\ &= n^{1/2} \left\{ \boldsymbol{\Omega}(\lambda) - \boldsymbol{\Omega}(0) \right\} + n^{-1/2} \sum_{i=1}^n \mathbf{R}_{1i}^* + \boldsymbol{o}_p(1), \end{aligned} \quad (\text{B.3})$$

where

$$\mathbf{R}_{1i}^* = \mathbf{A}_2^{-1} \left\{ \boldsymbol{\Omega}(\lambda) \right\} \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) \right\} - \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\}. \quad (\text{B.4})$$

Based on the approximation in (B.3), an estimator for the variance-covariance matrix of  $\mathbf{T}_1$  is given by

$$\hat{\boldsymbol{\nu}}_1 = (n-1)^{-1} \sum_{i=1}^n (\mathbf{R}_{1i} - \bar{\mathbf{R}}_1)(\mathbf{R}_{1i} - \bar{\mathbf{R}}_1)^T,$$

where  $\bar{\mathbf{R}}_1 = n^{-1} \sum_{i=1}^n \mathbf{R}_{1i}$ , and

$$\mathbf{R}_{1i} = \hat{\mathbf{A}}_2^{-1} \left\{ \mathbf{Q}^{(B)}; \tilde{\boldsymbol{\Omega}}(\lambda) \right\} \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \tilde{\boldsymbol{\Omega}}(\lambda) \right\} - \hat{\mathbf{A}}_1^{-1} \left\{ \mathbf{Q}; \tilde{\boldsymbol{\Omega}}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \tilde{\boldsymbol{\Omega}}(0) \right\}.$$

(II) An estimator for the variance-covariance matrix of  $\mathbf{T}_2(\lambda)$ ,  $\widehat{\boldsymbol{\nu}}_2$ :

The first-order Taylor expansion of  $\mathbf{T}_2(\lambda)$  around  $\boldsymbol{\Omega}(0)$  gives the following approximation,

$$\begin{aligned} \mathbf{T}_2(\lambda) &\approx n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} + \\ &\quad n^{-1/2} \sum_{i=1}^n \frac{\partial \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega} \right\}}{\partial \boldsymbol{\Omega}^T} \Big|_{\boldsymbol{\Omega} = \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0)^T, (1+\lambda)\sigma^2(0) \right\}^T} \left\{ \widetilde{\boldsymbol{\Omega}}(0) - \boldsymbol{\Omega}(0) \right\} \\ &\approx n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} - \\ &\quad \mathbf{A}_2 \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} n^{-1/2} \sum_{i=1}^n \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\} \end{aligned} \quad (\text{B.5})$$

$$\begin{aligned} &= n^{-1/2} \sum_{i=1}^n \left[ \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} \right. \\ &\quad \left. - \mathbf{A}_2 \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\} \right] \\ &= n^{-1/2} \sum_{i=1}^n \mathbf{R}_{2i}^*, \end{aligned} \quad (\text{B.6})$$

where

$$\mathbf{R}_{2i}^* = \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} - \mathbf{A}_2 \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\} \mathbf{A}_1^{-1} \left\{ \boldsymbol{\Omega}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \right\}. \quad (\text{B.7})$$

The approximation in (B.5) follows from (B.2) and that, as  $n \rightarrow \infty$ ,

$$n^{-1} \sum_{i=1}^n \frac{\partial \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega} \right\}}{\partial \boldsymbol{\Omega}^T} \Big|_{\boldsymbol{\Omega} = \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0)^T, (1+\lambda)\sigma^2(0) \right\}^T} \xrightarrow{p} \mathbf{A}_2 \left\{ \boldsymbol{\Omega}_{-\sigma^2}(0), (1+\lambda)\sigma^2(0) \right\}.$$

Based on the approximation in (B.6), an estimator for the asymptotic variance-covariance matrix of  $\mathbf{T}_2(\lambda)$  is given by

$$\widehat{\boldsymbol{\nu}}_2 = (n-1)^{-1} \sum_{i=1}^n (\mathbf{R}_{2i} - \overline{\mathbf{R}}_2)(\mathbf{R}_{2i} - \overline{\mathbf{R}}_2)^T,$$

where  $\overline{\mathbf{R}}_2 = n^{-1} \sum_{i=1}^n \mathbf{R}_{2i}$  and

$$\mathbf{R}_{2i} = \boldsymbol{\psi}^{(B)} \left\{ \mathbf{Q}_i^{(B)}(\lambda); \widetilde{\boldsymbol{\Omega}}_{-\sigma^2}(0), (1+\lambda)\tilde{\sigma}^2(0) \right\} - \widehat{\mathbf{A}}_2 \left\{ \widetilde{\boldsymbol{\Omega}}_{-\sigma^2}(0), (1+\lambda)\tilde{\sigma}^2(0) \right\} \widehat{\mathbf{A}}_1^{-1} \left\{ \widetilde{\boldsymbol{\Omega}}(0) \right\} \boldsymbol{\psi} \left\{ \mathbf{Q}_i; \widetilde{\boldsymbol{\Omega}}(0) \right\}.$$

## Web Appendix C: Asymptotic Equivalence Between $\mathbf{T}_1^*(\lambda)$ and $\mathbf{T}_2^*(\lambda)$

First consider the first-order Taylor expansion of  $\mathbf{T}_2(\lambda)$  around  $\boldsymbol{\Omega}(\lambda)$  under  $H_a$ ,

$$\begin{aligned}
\mathbf{T}_2(\lambda) &= n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) \} + \\
&\quad n^{-1/2} \sum_{i=1}^n \frac{\partial \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega} \}}{\partial \boldsymbol{\Omega}^T} \Big|_{\boldsymbol{\Omega} = \boldsymbol{\Omega}(\lambda)} \left\{ \tilde{\boldsymbol{\Omega}}(0) - \boldsymbol{\Omega}(0) - n^{-1/2} \boldsymbol{\Delta}^*(\lambda) \right\} + \mathbf{o}_p(1) \\
&= n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) \} - \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \mathbf{A}_1^{-1} \{ \boldsymbol{\Omega}(0) \} n^{-1/2} \sum_{i=1}^n \boldsymbol{\psi} \{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \} \\
&\quad + \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \boldsymbol{\Delta}^*(\lambda) + \mathbf{o}_p(1) \\
&= \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \boldsymbol{\Delta}^*(\lambda) + \\
&\quad n^{-1/2} \sum_{i=1}^n \left[ \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}(\lambda) \} - \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \mathbf{A}_1^{-1} \{ \boldsymbol{\Omega}(0) \} \boldsymbol{\psi} \{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \} \right] + \mathbf{o}_p(1) \\
&= \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \boldsymbol{\Delta}^*(\lambda) + \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} n^{-1/2} \sum_{i=1}^n \mathbf{R}_{1i}^* + \mathbf{o}_p(1). \tag{C.1}
\end{aligned}$$

From (B.3), under  $H_a$ ,  $\mathbf{T}_1(\lambda) = \boldsymbol{\Delta}^*(\lambda) + n^{-1/2} \sum_{i=1}^n \mathbf{R}_{1i}^* + \mathbf{o}_p(1)$ . Therefore, (C.1) implies

$$\mathbf{T}_2(\lambda) = \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \mathbf{T}_1(\lambda) + \mathbf{o}_p(1). \tag{C.2}$$

Next relate  $\text{var}(\mathbf{R}_{2i}^*)$  and  $\text{var}(\mathbf{R}_{1i}^*)$  under  $H_a$  as  $n \rightarrow \infty$ . By (B.7),

$$\begin{aligned}
\mathbf{R}_{2i}^* &= \mathbf{A}_2 \{ \boldsymbol{\Omega}_{-\sigma^2}(0), \sigma^2(\lambda) \} \left[ \mathbf{A}_2^{-1} \{ \boldsymbol{\Omega}_{-\sigma^2}(0), \sigma^2(\lambda) \} \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}_{-\sigma^2}(0), \sigma^2(\lambda) \} - \right. \\
&\quad \left. \mathbf{A}_1^{-1} \{ \boldsymbol{\Omega}(0) \} \boldsymbol{\psi} \{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \} \right]. \\
&\longrightarrow \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \left[ \mathbf{A}_2^{-1} \{ \boldsymbol{\Omega}(\lambda) \} \boldsymbol{\psi}^{(B)} \{ \mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}(\lambda) \} - \mathbf{A}_1^{-1} \{ \boldsymbol{\Omega}(0) \} \boldsymbol{\psi} \{ \mathbf{Q}_i; \boldsymbol{\Omega}(0) \} \right] \\
&= \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \mathbf{R}_{1i}^*.
\end{aligned}$$

Therefore, under  $H_a$ , as  $n \rightarrow \infty$ ,  $\text{var}(\mathbf{R}_{2i}^*) \rightarrow \mathbf{A}_2 \{ \boldsymbol{\Omega}(\lambda) \} \text{var}(\mathbf{R}_{1i}^*) \mathbf{A}_2^T \{ \boldsymbol{\Omega}(\lambda) \}$ , thus

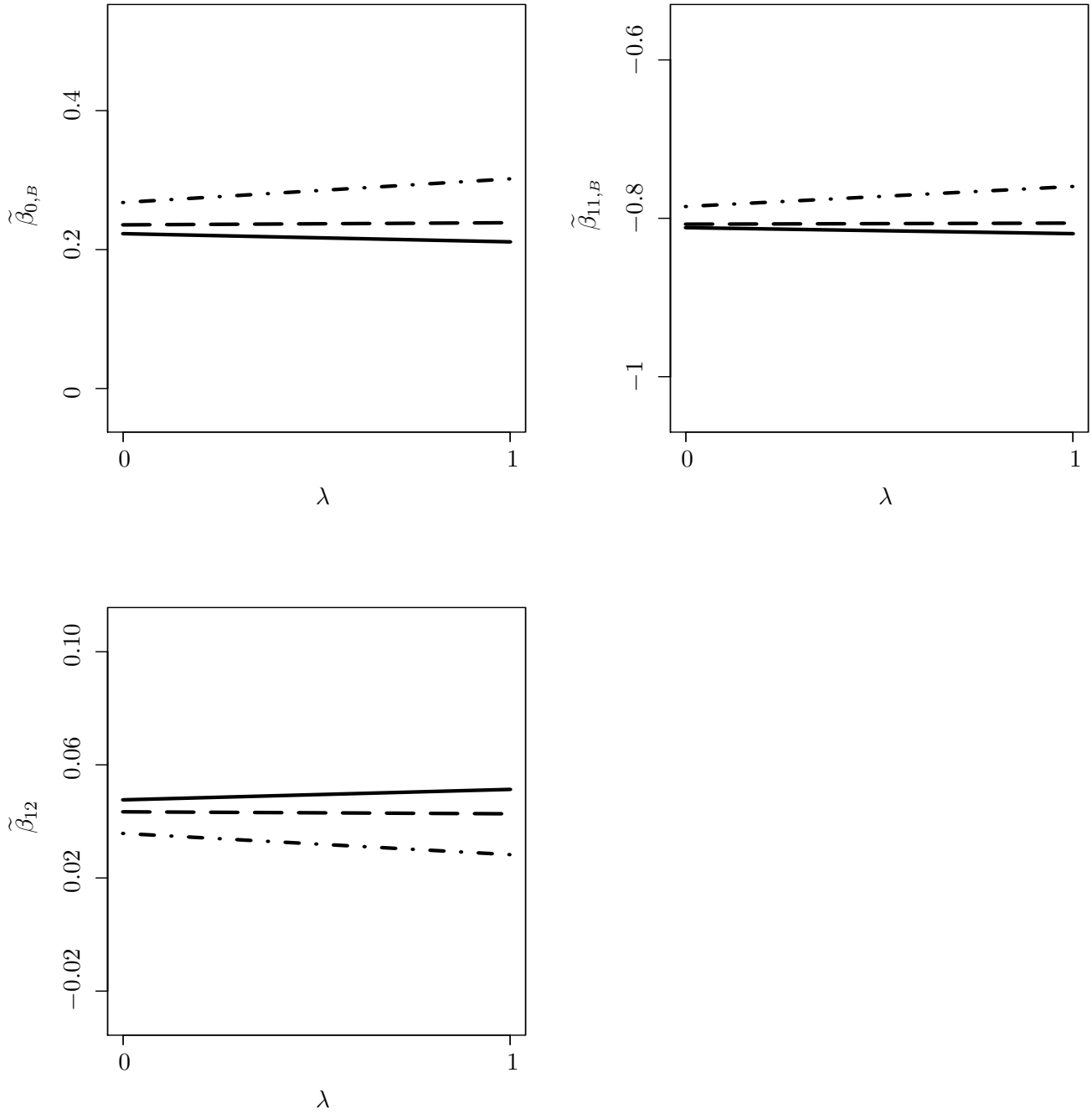
$$\text{var}(\mathbf{R}_{2i}^*)^{-1} \rightarrow \mathbf{A}_2^{-1} \{ \boldsymbol{\Omega}(\lambda) \}^T \text{var}(\mathbf{R}_{1i}^*)^{-1} \mathbf{A}_2^{-1} \{ \boldsymbol{\Omega}(\lambda) \}. \tag{C.3}$$

Lastly, by the definition of  $\mathbf{T}_2^*(\lambda)$  and (C.2),

$$\begin{aligned}
\mathbf{T}_2^{*T}(\lambda)\mathbf{T}_2^*(\lambda) &= \mathbf{T}_2^T(\lambda)\text{var}(\mathbf{R}_{2i}^*)^{-1}\mathbf{T}_2(\lambda) \\
&= \mathbf{T}_1^T(\lambda)\mathbf{A}_2^T\{\boldsymbol{\Omega}(\lambda)\}\text{var}(\mathbf{R}_{2i}^*)^{-1}\mathbf{A}_2^T\{\boldsymbol{\Omega}(\lambda)\}\mathbf{T}_1(\lambda) + \boldsymbol{o}_p(1) \\
\left[ \text{by (C.3)} \right] &\longrightarrow \mathbf{T}_1^T(\lambda)\text{var}(\mathbf{R}_{1i}^*)^{-1}\mathbf{T}_1(\lambda) \\
&= \mathbf{T}_1^{*T}(\lambda)\mathbf{T}_1^*(\lambda),
\end{aligned}$$

which establishes the asymptotic equivalence between  $\mathbf{T}_1^*(\lambda)$  and  $\mathbf{T}_2^*(\lambda)$ .

#### Web Appendix D: SIMEX Plots for the SWAN Data



**Figure 1.** SIMEX plots of the MLEs  $\tilde{\theta}_B^{(c)}(\lambda)$ ,  $\tilde{\theta}_B^{(m)}(\lambda)$ , and  $\tilde{\theta}_B^{(n)}(\lambda)$  computed from the SWAN data. The line types are,  $\tilde{\theta}_B^{(c)}(\lambda)$ : long dashed;  $\tilde{\theta}_B^{(m)}(\lambda)$ : dash-dotted; and  $\tilde{\theta}_B^{(n)}(\lambda)$ : solid. The ranges of the vertical axes are set to be one estimated standard deviation of  $\tilde{\theta}^{(n)}(0)$  below and above the average of the three types of estimates at  $\lambda = 0$ .