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

Xianzheng Huang,[1,*] Leonard A. Stefanski,[2] and Marie Davidian[2]

[1] Department of Statistics, University of South Carolina,

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

[2] 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. Recall in Theorem 1, $\mathbf{S} = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{W}_{m\times 1}$, which is the ordinary least squares estimator for $\mathbf{X}$ if one treats $\mathbf{X}$ as a parameter. Assuming σ^2 known and viewing $\mathbf{X}$ as an unknown parameter, $\mathbf{S}$ is a complete sufficient statistic for $\mathbf{X}$, and $\mathbf{S}|\mathbf{X} \sim N_p\{\mathbf{X}, \sigma^2(\mathbf{D}^T\mathbf{D})^{-1}\}$. Therefore, by the Factorization Theorem,

$$f_{\mathbf{W}|\mathbf{X}}(\boldsymbol{w}|\boldsymbol{x};\sigma^2) = f_{\mathbf{W}|\mathbf{S}}(\boldsymbol{w}|\boldsymbol{s};\sigma^2) f_{\mathbf{S}|\mathbf{X}}(\boldsymbol{s}|\boldsymbol{x};\sigma^2),$$

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

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

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

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

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

$$\begin{aligned} \widetilde{\Delta} &= \int \Big\{ f_{\mathbf{Y}|\mathbf{X},\mathbf{H}}(\boldsymbol{y}|\boldsymbol{s} + \mathbf{G}\boldsymbol{z}, \boldsymbol{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f^{(a)}_{\mathbf{X}|\mathbf{H}}(\boldsymbol{s} + \mathbf{G}\boldsymbol{z}|\boldsymbol{h}; \boldsymbol{\tau}^{(a)}) - \\ &\qquad f_{\mathbf{Y}|\mathbf{S},\mathbf{H}}(\boldsymbol{y}|\boldsymbol{s}, \boldsymbol{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f^{(a)}_{\mathbf{X}|\mathbf{H}}(\boldsymbol{s}|\boldsymbol{h}; \boldsymbol{\tau}^{(a)}) \Big\} \phi(\boldsymbol{z}) d\boldsymbol{z}. \end{aligned} \tag{A.2}$$

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, $\mathbf{G} \to \mathbf{0}_{p\times p}$, and the integrand in (A.2) converges to zero. Thus, as the longitudinal process information increases without bound, $\widetilde{\Delta} \to 0$ whenever integration and limit can be interchanged in (A.2). A sufficient condition for the interchange of limit and integration is that

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

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

$$f_{\mathbf{W}|\mathbf{S}}(\boldsymbol{w}|\boldsymbol{s}; \sigma^2) \int f_{\mathbf{Y}|\mathbf{X},\mathbf{H}}(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f_{\mathbf{S}|\mathbf{X}}(\boldsymbol{s}|\boldsymbol{x}; \sigma^2) f^{(a)}_{\mathbf{X}|\mathbf{H}}(\boldsymbol{x}|\boldsymbol{h}; \boldsymbol{\tau}^{(a)}) d\boldsymbol{x}, \tag{A.3}$$

$$f_{\mathbf{W}|\mathbf{S}}(\boldsymbol{w}|\boldsymbol{s}; \sigma^2) f_{\mathbf{Y}|\mathbf{S},\mathbf{H}}(\boldsymbol{y}|\boldsymbol{s}, \boldsymbol{h}; \boldsymbol{\theta}, \boldsymbol{\zeta}) f^{(a)}_{\mathbf{X}|\mathbf{H}}(\boldsymbol{s}|\boldsymbol{h}; \boldsymbol{\tau}^{(a)}), \tag{A.4}$$

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 $\widehat{\boldsymbol{\nu}}_1$ and $\widehat{\boldsymbol{\nu}}_2$

Notations:

- For $i = 1, \ldots, n$, $\mathbf{Q}_i$ is the observed data for subject i; $\mathbf{Q}_i^{(B)}$ is the B sets of λ-remeasured data for subject i. $\mathbf{Q} = \{\mathbf{Q}_i, i = 1, \ldots, n\}$; $\mathbf{Q}^{(B)} = \{\mathbf{Q}_i^{(B)}, i = 1, \ldots, n\}$.

- $\boldsymbol{\Omega}$ denotes the $d \times 1$ vector of all unknown parameters in the joint model. $\boldsymbol{\Omega}_{-\sigma^2}$ denotes $\boldsymbol{\Omega}$ excluding σ^2.

- $E\big[\boldsymbol{\psi}\left\{\mathbf{Q}_i; \boldsymbol{\Omega}_{-\sigma^2}(0), \sigma^2(0)\right\}\big] = \mathbf{0}$ uniquely defines $\boldsymbol{\Omega}(0) = \big\{\boldsymbol{\Omega}_{-\sigma^2}(0)^T,\, \sigma^2(0)\big\}^T$; $E\big[\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}; \boldsymbol{\Omega}_{-\sigma^2}(\lambda), \sigma^2(\lambda)\right\}\big] = \mathbf{0}$ 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 $\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}\right\} = B^{-1}\sum_{b=1}^{B}\boldsymbol{\psi}\{\mathbf{Q}_{b,i}(\lambda); \boldsymbol{\Omega}\}$.

- $\widetilde{\boldsymbol{\Omega}}(0)$ solves $\sum_{i=1}^{n}\boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega}) = \mathbf{0}$; $\widetilde{\boldsymbol{\Omega}}_B(\lambda)$ solves $\sum_{i=1}^{n}\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda); \boldsymbol{\Omega}\right\} = \mathbf{0}$.

- $t_1^*(\lambda) = [\mathbf{T}_1(\lambda)]_{(k)}[\widehat{\boldsymbol{\nu}}_1]_{(k)}^{-1/2}$, where $\mathbf{T}_1(\lambda) = n^{1/2}\left\{\widetilde{\boldsymbol{\Omega}}_{-\sigma^2}(\lambda) - \widetilde{\boldsymbol{\Omega}}_{-\sigma^2}(0)\right\}$.

- $t_2^*(\lambda) = [\mathbf{T}_2(\lambda)]_{(k)}[\widehat{\boldsymbol{\nu}}_2]_{(k)}^{-1/2}$, where $\mathbf{T}_2(\lambda) = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda); \widetilde{\boldsymbol{\Omega}}_{-\sigma^2}(0), (1+\lambda)\tilde{\sigma}^2(0)\right\}$.

- 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 $\mathbf{T}_1^* = \widehat{\boldsymbol{\nu}}_1^{-1/2}\mathbf{T}_1(\lambda)$ and $\mathbf{T}_2^* = \widehat{\boldsymbol{\nu}}_2^{-1/2}\mathbf{T}_2(\lambda)$.

- 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.

$$
\begin{aligned}
\mathbf{A}_1\left\{\boldsymbol{\Omega}(0)\right\} &= E\left\{-\partial\boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega})/\partial\boldsymbol{\Omega}^T\right\}\Big|_{\boldsymbol{\Omega} = \boldsymbol{\Omega}(0)};\\
\widehat{\mathbf{A}}_1\left\{\mathbf{Q}; \widetilde{\boldsymbol{\Omega}}(0)\right\} &= -n^{-1}\sum_{i=1}^{n}\partial\boldsymbol{\psi}(\mathbf{Q}_i; \boldsymbol{\Omega})/\partial\boldsymbol{\Omega}^T\Big|_{\boldsymbol{\Omega} = \widetilde{\boldsymbol{\Omega}}(0)};
\end{aligned}
$$

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

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

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

Using the influence function approximation, we have

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

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

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

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

where

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

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

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

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

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

(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,

$$\begin{aligned}
\mathbf{T}_2(\lambda) &\approx n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda);\boldsymbol{\Omega}_{-\sigma^2}(0),(1+\lambda)\sigma^2(0)\right\}+\\
&\quad n^{-1/2}\sum_{i=1}^{n}\frac{\partial\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda);\boldsymbol{\Omega}\right\}}{\partial\boldsymbol{\Omega}^T}\Big|_{\boldsymbol{\Omega}=\left\{\boldsymbol{\Omega}_{-\sigma^2}(0)^T,(1+\lambda)\sigma^2(0)\right\}^T}\left\{\widetilde{\boldsymbol{\Omega}}(0)-\boldsymbol{\Omega}(0)\right\}\\
&\approx n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda);\boldsymbol{\Omega}_{-\sigma^2}(0),(1+\lambda)\sigma^2(0)\right\}-\\
&\quad \mathbf{A}_2\left\{\boldsymbol{\Omega}_{-\sigma^2}(0),(1+\lambda)\sigma^2(0)\right\}n^{-1/2}\sum_{i=1}^{n}\mathbf{A}_1^{-1}\left\{\boldsymbol{\Omega}(0)\right\}\boldsymbol{\psi}\{\mathbf{Q}_i;\boldsymbol{\Omega}(0)\} \qquad \text{(B.5)}\\
&= n^{-1/2}\sum_{i=1}^{n}\Big[\boldsymbol{\psi}^{(B)}\left\{\mathbf{Q}_i^{(B)}(\lambda);\boldsymbol{\Omega}_{-\sigma^2}(0),(1+\lambda)\sigma^2(0)\right\}\\
&\quad -\mathbf{A}_2\left\{\boldsymbol{\Omega}_{-\sigma^2}(0),(1+\lambda)\sigma^2(0)\right\}\mathbf{A}_1^{-1}\left\{\boldsymbol{\Omega}(0)\right\}\boldsymbol{\psi}\left\{\mathbf{Q}_i;\boldsymbol{\Omega}(0)\right\}\Big]\\
&= n^{-1/2}\sum_{i=1}^{n}\mathbf{R}_{2i}^*, \qquad \text{(B.6)}
\end{aligned}$$

where

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

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

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

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 $\mathbf{\Omega}(\lambda)$ under H_a,

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

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

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

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

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

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

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

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

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

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

Web Appendix D: SIMEX Plots for the SWAN Data

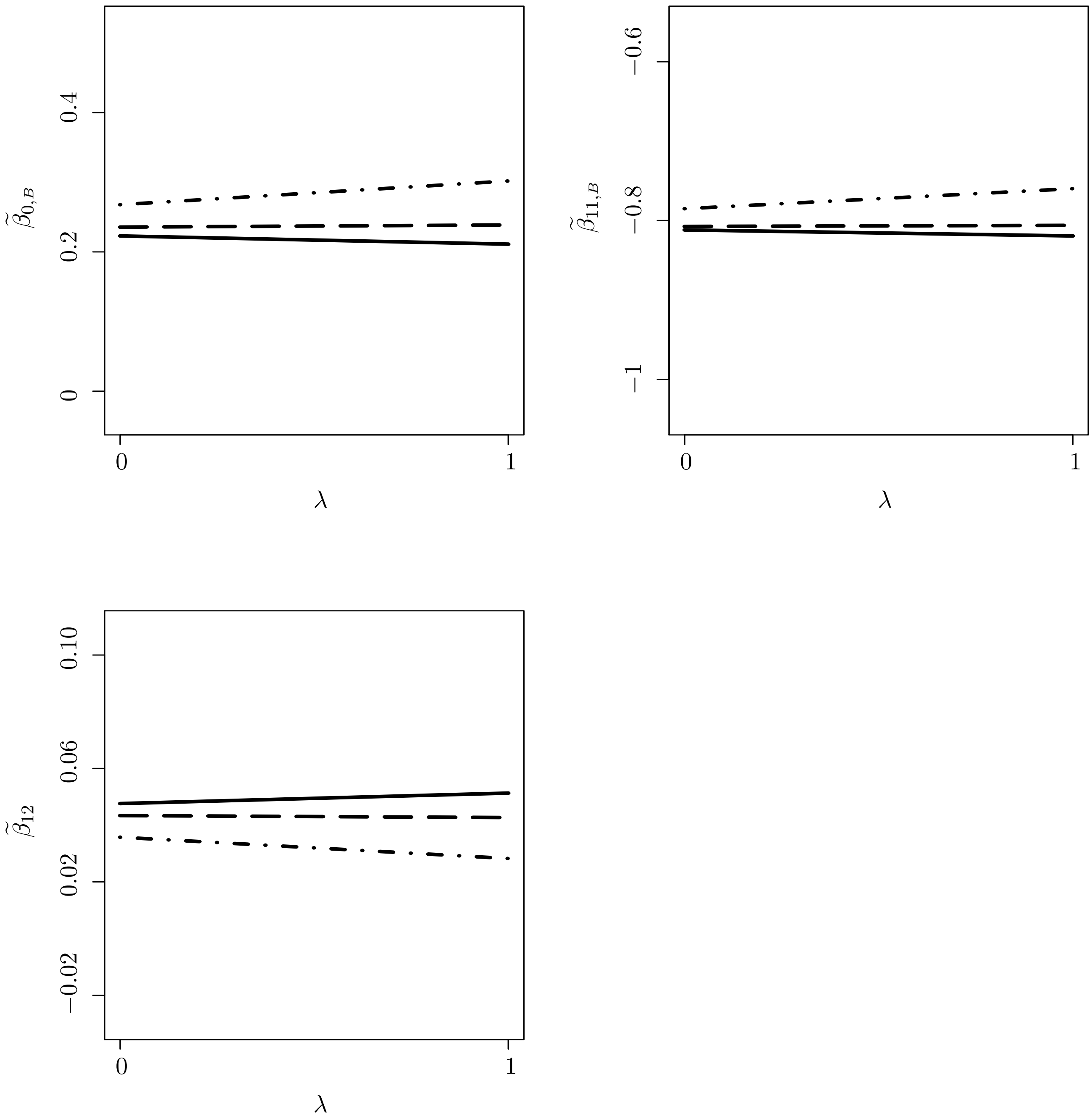

Figure 1. SIMEX plots of the MLEs $\widetilde{\boldsymbol{\theta}}_B^{(c)}(\lambda)$, $\widetilde{\boldsymbol{\theta}}_B^{(m)}(\lambda)$, and $\widetilde{\boldsymbol{\theta}}_B^{(n)}(\lambda)$ computed from the SWAN data. The line types are, $\widetilde{\boldsymbol{\theta}}_B^{(c)}(\lambda)$: long dashed;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