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 Davidian2

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. Denote by \( \hat{X}_m \) the ordinary least squares estimator for \( X \), i.e., \( \hat{X}_m = (D^TD)^{-1}D^TW_{m\times 1} \). Assuming \( \sigma^2 \) known and viewing \( X \) as an unknown parameter, \( \hat{X}_m \) is a complete sufficient statistic for \( X \), and \( \hat{X}_m|X \sim N_p\{X, \sigma^2(D^TD)^{-1}\} \). Therefore, by the Factorization Theorem,

\[
f_{W|X}(w|x; \sigma^2) = f_{W|\hat{X}_m}(w|\hat{x}_m; \sigma^2)f_{X|\hat{X}_m}(\hat{x}_m|x; \sigma^2),
\]

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

\[
f_{Y,W|H}(y, w|h; \Omega) = \int f_{Y,X,H}(y|x, h; \theta, \zeta)f_{W|X}(w|x; \sigma^2)f_{X|H}(x|h; \tau^{(a)})dx
\]

\[
= f_{W|\hat{X}_m}(w|\hat{x}_m; \sigma^2) \times \int f_{Y,X,H}(y|x, h; \theta, \zeta)f_{X|\hat{X}_m}(\hat{x}_m|x; \sigma^2)f_{X|H}(x|h; \tau^{(a)})dx. \quad (A.1)
\]

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

\[
\bar{\Delta} = \int f_{Y,X,H}(y|x, h; \theta, \zeta)f_{X|\hat{X}_m}(\hat{x}_m|x; \sigma^2)f_{X|H}(x|h; \tau^{(a)})dx
\]

\[-f_{Y,X,H}(y|\hat{x}_m, h; \theta, \zeta)f_{X|H}(\hat{x}_m|h; \tau^{(a)}).
\]
After the change of variable $z = G^{-1}_m(x - \hat{x}_m)$ and rearrangement of terms, we find that

$$
\tilde{\Delta} = \int \left\{ f_{Y|X,H}(y|x_m + G_m z, h; \theta, \zeta) f_{x_m|X}(\hat{x}_m + G_m z|h; \tau^{(a)}) - 
f_{Y|X_m,H}(y|x_m, h; \theta, \zeta) f_{x_m|H}(\hat{x}_m|h; \tau^{(a)}) \right\} \phi(z) \, dz.
$$

(A.2)

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

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

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

$$
f_{w|x_m}(w|\hat{x}_m; \sigma^2) \int f_{Y|X,H}(y|x, h; \theta, \zeta) f_{x_m|x}(\hat{x}_m|x; \sigma^2) f_{x|m}(\hat{x}_m|h; \tau^{(a)}) \, dx,
$$

(A.3)

and

$$
f_{w|x_m}(w|\hat{x}_m; \sigma^2) f_{Y|X_m,H}(y|\hat{x}_m, h; \theta, \zeta) f_{x|m}(\hat{x}_m|h; \tau^{(a)}),
$$

(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 $\hat{\nu}_1$ and $\hat{\nu}_2$

Notations:

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

- $\Omega$ denotes the $d \times 1$ vector of all unknown parameters in the joint model.

- $\Omega_{-\sigma^2}$ denotes $\Omega$ excluding $\sigma^2$. 


• $E[\psi \{Q_i; \Omega_{-\sigma^2}(0), \sigma^2(0)\}] = 0$ uniquely defines $\Omega(0) = \{\Omega_{-\sigma^2}(0)^T, \sigma^2(0)\}^T$;
• $E[\psi(b) \{Q_i^{(b)}; \Omega_{-\sigma^2}(\lambda), \sigma^2(\lambda)\}] = 0$ uniquely defines $\Omega(\lambda) = \{\Omega_{-\sigma^2}(\lambda)^T, \sigma^2(\lambda)\}^T$ for $\lambda > 0$, where the expectations are with respect to the true densities of $Q_i$ and $Q_i^{(b)}$, respectively, $\sigma^2(\lambda) = (1 + \lambda)\sigma^2(0)$, and $\psi(b) \{Q_i^{(b)}; \Omega\} = B^{-1} \sum_{b=1}^B \psi(Q_{b,i}(\lambda); \Omega)$.

• $\tilde{\Omega}(0)$ solves $\sum_{i=1}^n \psi(Q_i; \Omega) = 0$; $\tilde{\Omega}_b(\lambda)$ solves $\sum_{i=1}^n \psi^{(b)} \{Q_i^{(b)}(\lambda); \Omega\} = 0$.

• $t^*_1(\lambda) = [T_1(\lambda)](k)[\tilde{\Omega}_1(k)]^{-1/2}$, where $T_1(\lambda) = n^{1/2} \left\{ \tilde{\Omega}_{-\sigma^2}(\lambda) - \tilde{\Omega}_{-\sigma^2}(0) \right\}$.

• $t^*_2(\lambda) = [T_2(\lambda)](k)[\tilde{\Omega}_2(k)]^{-1/2}$, where $T_2(\lambda) = n^{-1/2} \sum_{i=1}^n \psi^{(b)} \{Q_i^{(b)}(\lambda); \tilde{\Omega}_{-\sigma^2}(0), (1 + \lambda)\bar{\sigma}^2(0)\}$.

• For a positive definite matrix $\Pi$, define $\Pi^{1/2}$ as the positive definite square root such that $\Pi^{1/2}(\Pi^{1/2})^T = \Pi$, and $\Pi^{-1/2}$ as the inverse of $\Pi^{1/2}$. Define $T_1^* = \tilde{\nu}_1^{-1/2} T_1(\lambda)$ and $T_2^* = \tilde{\nu}_2^{-1/2} T_2(\lambda)$.

• Consider the null hypothesis, $H_0 : \Omega_{-\sigma^2}(\lambda) - \Omega_{-\sigma^2}(0) = 0$, and the contiguous alternative hypothesis, $H_\alpha : \Omega_{-\sigma^2}(\lambda) - \Omega_{-\sigma^2}(0) = n^{-1/2} \Delta^*(\lambda)$.

Define the following Hessian matrices and the associated empirical estimators.

\[
A_1 \{\Omega(0)\} = E \left\{ -\partial \psi(Q_i; \Omega)/\partial \Omega^T \right\} \bigg| \Omega = \Omega(0);
\]

\[
\widehat{A}_1 \{Q; \tilde{\Omega}(0)\} = -n^{-1} \sum_{i=1}^n \partial \psi(Q_i; \Omega)/\partial \Omega^T \bigg| \Omega = \tilde{\Omega}(0);
\]

\[
A_2 \{\Omega(\lambda)\} = E \left[ -\partial \psi^{(b)} \{Q_i^{(b)}; \Omega\}/\partial \Omega^T \right] \bigg| \Omega = \Omega(\lambda);
\]

\[
\widehat{A}_2 \{Q^{(b)}(\lambda); \tilde{\Omega}(\lambda)\} = -n^{-1} \sum_{i=1}^n \partial \psi^{(b)} \{Q_i^{(b)}; \Omega\}/\partial \Omega^T \bigg| \Omega = \tilde{\Omega}(\lambda).
\]

Similarly denote by $A_2 \{\Omega_{-\sigma^2}(0), (1 + \lambda)\sigma^2(0)\}$ the expectation $E \left[ -\partial \psi^{(b)} \{Q_i^{(b)}; \Omega\}/\partial \Omega^T \right]$ evaluated at $\Omega = \{\Omega_{-\sigma^2}(0)^T, (1 + \lambda)\sigma^2(0)\}^T$. And define $\widehat{A}_2 \{Q^{(b)}(\lambda); \tilde{\Omega}_{-\sigma^2}(0), (1 + \lambda)\bar{\sigma}^2(0)\}$ as the average $-n^{-1} \sum_{i=1}^n \partial \psi^{(b)} \{Q_i^{(b)}; \Omega\}/\partial \Omega^T$ evaluated at $\Omega = \{\tilde{\Omega}_{-\sigma^2}(0)^T, (1 + \lambda)\bar{\sigma}^2(0)\}^T$. 

3
An estimator for the variance-covariance matrix of $T_1(\lambda)$, $\hat{\nu}_1$:

Using the influence function approximation, we have

$$n^{1/2} \left\{ \tilde{\Omega}(\lambda) - \Omega(\lambda) \right\} = n^{-1/2} \sum_{i=1}^{n} A_2^{-1} \{ \Omega(\lambda) \} \psi^{(b)} \{ Q_i^{(b)}(\lambda); \Omega(\lambda) \} + o_p(1), \quad (B.1)$$

$$n^{1/2} \left\{ \tilde{\Omega}(0) - \Omega(0) \right\} = n^{-1/2} \sum_{i=1}^{n} A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \} + o_p(1). \quad (B.2)$$

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

$$T_1(\lambda) = n^{1/2} \left\{ \Omega(\lambda) - \Omega(0) \right\} + n^{-1/2} \sum_{i=1}^{n} \left[ A_2^{-1} \{ \Omega(\lambda) \} \psi^{(b)} \{ Q_i^{(b)}(\lambda); \Omega(\lambda) - A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \} \} \right] + o_p(1)$$

$$= n^{1/2} \left\{ \Omega(\lambda) - \Omega(0) \right\} + n^{-1/2} \sum_{i=1}^{n} R^*_i + o_p(1), \quad (B.3)$$

where

$$R^*_i = A_2^{-1} \{ \Omega(\lambda) \} \psi^{(b)} \{ Q_i^{(b)}(\lambda); \Omega(\lambda) \} - A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \}. \quad (B.4)$$

Based on the approximation in (B.3), an estimator for the variance-covariance matrix of $T_1$ is given by

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

where $R_i = n^{-1} \sum_{i=1}^{n} R_{1i}$, and

$$R_{1i} = \tilde{A}_2^{-1} \{ Q_i^{(b)}; \tilde{\Omega}(\lambda) \} \psi^{(b)} \{ Q_i^{(b)}(\lambda); \tilde{\Omega}(\lambda) \} - \tilde{A}_1^{-1} \{ Q_i; \tilde{\Omega}(0) \} \psi \{ Q_i; \tilde{\Omega}(0) \}.$$
(II) An estimator for the variance-covariance matrix of $T_2(\lambda)$, $\hat{\nu}_2$:

The first-order Taylor expansion of $T_2(\lambda)$ around $\Omega(0)$ gives the following approximation,

$$T_2(\lambda) \approx n^{-1/2} \sum_{i=1}^{n} \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \} +$$

$$n^{-1/2} \sum_{i=1}^{n} \frac{\partial \psi^{(B)}}{\partial \Omega} \{ Q_i^{(B)}(\lambda); \Omega \} \bigg| \Omega = \{ \Omega_{\sigma^2}(0)^T, (1 + \lambda)\sigma^2(0) \}^T \left\{ \tilde{\Omega}(0) - \Omega(0) \right\}$$

$$\approx n^{-1/2} \sum_{i=1}^{n} \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \} -$$

$$A_2 \{ \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \} n^{-1/2} \sum_{i=1}^{n} A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \}$$

$$\approx n^{-1/2} \sum_{i=1}^{n} R_{2i}^*,$$  \hspace{1cm} (B.5)

where

$$R_{2i}^* = \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \} - A_2 \{ \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \} A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \}. $$  \hspace{1cm} (B.6)

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

$$n^{-1} \sum_{i=1}^{n} \frac{\partial \psi^{(B)}}{\partial \Omega} \{ Q_i^{(B)}(\lambda); \Omega \} \bigg| \Omega = \{ \Omega_{\sigma^2}(0)^T, (1 + \lambda)\sigma^2(0) \}^T \overset{p}{\to} A_2 \{ \Omega_{\sigma^2}(0), (1 + \lambda)\sigma^2(0) \}.$$  \hspace{1cm} (B.7)

Based on the approximation in (B.6), an estimator for the asymptotic variance-covariance matrix of $T_2(\lambda)$ is given by

$$\hat{\nu}_2 = (n - 1)^{-1} \sum_{i=1}^{n} (R_{2i} - \bar{R}_2)(R_{2i} - \bar{R}_2)^T,$$

where $\bar{R}_2 = n^{-1} \sum_{i=1}^{n} R_{2i}$ and

$$R_{2i} = \psi^{(B)} \{ Q_i^{(B)}(\lambda); \tilde{\Omega}_{\sigma^2}(0), (1 + \lambda)\tilde{\sigma}^2(0) \} - \tilde{A}_2 \{ \tilde{\Omega}_{\sigma^2}(0), (1 + \lambda)\tilde{\sigma}^2(0) \} \tilde{A}_1^{-1} \{ \tilde{\Omega}(0) \} \psi \{ Q_i; \tilde{\Omega}(0) \}.$$
Web Appendix C: Asymptotic Equivalence Between $T_1^*(\lambda)$ and $T_2^*(\lambda)$

First consider the first-order Taylor expansion of $T_2(\lambda)$ around $\Omega(\lambda)$ under $H_a$,

$$T_2(\lambda) = n^{-1/2} \sum_{i=1}^{n} \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega(\lambda) \} +$$

$$n^{-1/2} \sum_{i=1}^{n} \frac{\partial \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega \}}{\partial \Omega} \bigg|_{\Omega = \Omega(\lambda)} \left\{ \Omega(0) - \Omega(0) - n^{-1/2} \Delta^*(\lambda) \right\} + o_p(1)$$

$$= n^{-1/2} \sum_{i=1}^{n} \psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega(\lambda) \} - A_2 \{ \Omega(\lambda) \} \Omega^{-1} \{ \Omega(0) \} n^{-1/2} \sum_{i=1}^{n} \psi \{ Q_i; \Omega(0) \} + A_2 \{ \Omega(\lambda) \} \Delta^*(\lambda) + o_p(1)$$

$$= A_2 \{ \Omega(\lambda) \} \Delta^*(\lambda) + n^{-1/2} \sum_{i=1}^{n} [\psi^{(B)} \{ Q_i^{(B)}(\lambda); \Omega(\lambda) \} - A_2 \{ \Omega(\lambda) \} \Omega^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \}] + o_p(1)$$

$$= A_2 \{ \Omega(\lambda) \} \Delta^*(\lambda) + A_2 \{ \Omega(\lambda) \} n^{-1/2} \sum_{i=1}^{n} R_{1i}^* + o_p(1). \quad \text{(C.1)}$$

From (B.3), under $H_a$, $T_1(\lambda) = \Delta^*(\lambda) + n^{-1/2} \sum_{i=1}^{n} R_{1i}^* + o_p(1)$. Therefore, (C.1) implies

$$T_2(\lambda) = A_2 \{ \Omega(\lambda) \} T_1(\lambda) + o_p(1). \quad \text{(C.2)}$$

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

$$R_{2i}^* = A_2 \{ \Omega_{-\sigma^2}(0), \sigma^2(\lambda) \} \left[ A_2^{-1} \{ \Omega_{-\sigma^2}(0), \sigma^2(\lambda) \} \psi^{(B)} \{ Q_i^{(B)}; \Omega_{-\sigma^2}(0), \sigma^2(\lambda) \} - A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \} \right] -$$

$$\rightarrow A_2 \{ \Omega(\lambda) \} \left[ A_2^{-1} \{ \Omega(\lambda) \} \psi^{(B)} \{ Q_i^{(B)}; \Omega(\lambda) \} - A_1^{-1} \{ \Omega(0) \} \psi \{ Q_i; \Omega(0) \} \right] -$$

$$= A_2 \{ \Omega(\lambda) \} R_{1i}^*.$$

Therefore, under $H_a$, as $n \to \infty$, $\text{var}(R_{2i}^*) \to A_2 \{ \Omega(\lambda) \} \text{var}(R_{1i}^*)A_2^T \{ \Omega(\lambda) \}$, thus

$$\text{var}(R_{2i}^*)^{-1} \to A_2^{-1} \{ \Omega(\lambda) \}^T \text{var}(R_{1i}^*)^{-1} A_2^{-1} \{ \Omega(\lambda) \}. \quad \text{(C.3)}$$
Lastly, by the definition of $T_2^*(\lambda)$ and (C.2),

$$T_2^T(\lambda)T_2^*(\lambda) = T_2^T(\lambda)\text{var}(R_{2i})^{-1}T_2(\lambda)$$

$$= T_1^T(\lambda)A_2^T\Omega(\lambda)A_2T_1(\lambda) + o_p(1)$$

[by (C.3)]

$$\rightarrow T_1^T(\lambda)\text{var}(R_{1i})^{-1}T_1(\lambda)$$

$$= T_1^*T(\lambda)T_1^*(\lambda),$$

which establishes the asymptotic equivalence between $T_1^*(\lambda)$ and $T_2^*(\lambda)$.

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 $\lambda = 0$. 