Supplement to "Multidimensional credibility: A new approach based on joint distribution function"

Limin Wen[a], Wei Liu[a], Yiying Zhang[b,*]

a. School of Mathematics and Statistics, Jiangxi Normal University, Nanchang, 330022, China.
b. Department of Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China.

February 1, 2024

Abstract

In this supplementary material designed for "*Multidimensional credibility: A new approach based on joint distribution function*", we furnish meticulous proofs for each theorem, proposition, and requisite lemma presented in the main paper. For context, notation, and definitions, we refer to the paper.

*Corresponding author. E-mail: zhangyy3@sustech.edu.cn

1 Proof of Theorem 2.1

Based on Assumptions 2.1 and 2.2, we have:

$$\mathbb{E}[\mathbf{Y}] = \mathbf{1}_n \otimes \boldsymbol{\mu}_0,$$

where $\mathbf{1}_n$ denotes a column vector with dimension n. According to the double expectation formula, we have

$$\mathbb{Cov}(\mathbf{Y}_i, \mathbf{Y}_j) = \mathbb{E}\left[\mathbb{Cov}(\mathbf{Y}_i, \mathbf{Y}_j|\boldsymbol{\Theta})\right] + \mathbb{Cov}\left[\mathbb{E}(\mathbf{Y}_i|\boldsymbol{\Theta}), \mathbb{E}(\mathbf{Y}_j|\boldsymbol{\Theta})\right] = \begin{cases} T + \Sigma_0, & i = j \\ T, & i \neq j \end{cases}.$$

Thus, we have

$$\mathbb{Cov}(\boldsymbol{\mu}(\boldsymbol{\Theta}), \mathbf{Y}) = \mathbb{E}\left[\mathbb{Cov}(\boldsymbol{\mu}(\boldsymbol{\Theta}), \mathbf{Y}|\boldsymbol{\Theta})\right] + \mathbb{Cov}\left[\boldsymbol{\mu}(\boldsymbol{\Theta}), \mathbb{E}(\mathbf{Y}|\boldsymbol{\Theta})\right] = \mathbf{1}_n' \otimes T$$

and

$$\mathbb{Cov}(\mathbf{Y}, \mathbf{Y}) = I_n \otimes \Sigma_0 + (\mathbf{1}_n\mathbf{1}_n') \otimes T,$$

where I_n is the n-dimensional identity matrix. By applying the matrix inversion formula:

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1} DA^{-1},$$

we obtain

$$\mathbb{Cov}^{-1}(\mathbf{Y}, \mathbf{Y}) = I_n \otimes \Sigma_0^{-1} - \left(\mathbf{1}_n \otimes \Sigma_0^{-1}\right)\left(T^{-1} + n\Sigma_0^{-1}\right)^{-1}\left(\mathbf{1}_n' \otimes \Sigma_0^{-1}\right).$$

Consequently, the optimal linear estimate of the conditional mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$ is:

$$\begin{aligned}
&\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} \\
&= \boldsymbol{\mu}_0 + (\mathbf{1}_n' \otimes T)\left[I_n \otimes \Sigma_0^{-1} - \left(\mathbf{1}_n \otimes \Sigma_0^{-1}\right)\left(T^{-1} + n\Sigma_0^{-1}\right)^{-1}\left(\mathbf{1}_n' \otimes \Sigma_0^{-1}\right)\right](\mathbf{Y} - \mathbf{1}_n \otimes \boldsymbol{\mu}_0) \\
&= \boldsymbol{\mu}_0 + nT\Sigma_0^{-1}\left(I_p - \left(T^{-1} + n\Sigma_0^{-1}\right)^{-1} n\Sigma_0^{-1}\right)\left(\overline{\mathbf{Y}} - \boldsymbol{\mu}_0\right) \\
&= \boldsymbol{\mu}_0 + Z_{C,n}\left(\overline{\mathbf{Y}} - \boldsymbol{\mu}_0\right) \\
&= Z_{C,n}\overline{\mathbf{Y}} + (I_p - Z_{C,n})\boldsymbol{\mu}_0,
\end{aligned}$$

which completes the proof of Theorem 2.1.

2 A useful lemma for constrained problem

Lemma 1. *Consider a constrained optimization problem*

$$\begin{cases} \min\limits_{x_1,\ldots,x_p} f\left(x_1, x_2, \cdots, x_p\right) \\ s.t.\ g\left(x_1, x_2, \cdots, x_p\right) \geq 0 \end{cases}, \tag{1}$$

as well as the optimization problem without constraints:

$$\min_{x_1,\ldots,x_p} f\left(x_1, x_2, \cdots, x_p\right). \tag{2}$$

and let $\left(x_1^*, x_2^*, \cdots, x_p^*\right)$ *be the optimal solution obtained from (2). If* $\left(x_1^*, x_2^*, \cdots, x_p^*\right)$ *precisely satisfies the condition* $g\left(x_1^*, x_2^*, \cdots, x_p^*\right) \geq 0$*, then problems (1) and (2) are equivalent.*

Proof. Let $\left(y_1^*, y_2^*, \cdots, y_p^*\right)$ be the solution of problem (1). Evidently, we have

$$f\left(x_1^*, x_2^*, \cdots, x_p^*\right) \leq f\left(y_1^*, y_2^*, \cdots, y_p^*\right).$$

On the other hand, due to $g\left(x_1^*, x_2^*, \cdots, x_p^*\right) \geq 0$, $\left(x_1^*, x_2^*, \cdots, x_p^*\right)$ is within the feasible domain of equation (1). Therefore, with the presence of

$$f\left(x_1^*, x_2^*, \cdots, x_p^*\right) \geq f\left(y_1^*, y_2^*, \cdots, y_p^*\right),$$

it follows that

$$f\left(x_1^*, x_2^*, \cdots, x_p^*\right) = f\left(y_1^*, y_2^*, \cdots, y_p^*\right),$$

indicating that problems (1) and (2) are equivalent. □

3 Proof of Theorem 3.1

Due to

$$\mathbb{E}\left[H_i\left(\mathbf{y}\right)\right] = \mathbb{E}\left[\mathbb{E}\left(H_i(\mathbf{y})|\,\boldsymbol{\Theta}\right)\right] = F_0\left(\mathbf{y}\right),$$

we can apply the variance formula to derive

$$\begin{aligned} & \mathbb{E}\left[\left(F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right) - \alpha_0\left(\mathbf{y}\right) - \sum_{s=1}^{n} \alpha_s H_s\left(\mathbf{y}\right)\right)^2\right] \\ = \ & \mathbb{V}\left(F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right) - \sum_{s=1}^{n} \alpha_s H_s\left(\mathbf{y}\right)\right) + \left\{\left(1 - \sum_{s=1}^{n} \alpha_s\right) F_0\left(\mathbf{y}\right) - \alpha_0\left(\mathbf{y}\right)\right\}^2 \\ \geq \ & \mathbb{V}\left(F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right) - \sum_{s=1}^{n} \alpha_s H_s\left(\mathbf{y}\right)\right). \end{aligned} \tag{3}$$

It is straightforward to verify that the equality above holds if and only if

$$\alpha_0\left(\mathbf{y}\right) = (1 - \sum_{s=1}^{n} \alpha_s) F_0\left(\mathbf{y}\right),$$

Therefore, the solution for $\alpha_0\left(\mathbf{y}\right)$ is given by

$$\widehat{\alpha_0}\left(\mathbf{y}\right) = (1 - \sum_{s=1}^{n} \alpha_s) F_0\left(\mathbf{y}\right). \tag{4}$$

Substituting equation (4) into (3) and denoting

$$\varphi\left(\mathbf{y}\right) = \mathbb{E}\left(F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right) - \widehat{\alpha_0}\left(\mathbf{y}\right) - \sum_{s=1}^{n} \alpha_s H_s\left(\mathbf{y}\right)\right)^2,$$

the problem (14) in the main paper is equivalent to the problem

$$\min_{\alpha_s \in \mathbb{R}} \int_{\mathbb{R}^p} \varphi\left(\mathbf{y}\right) d\mathbf{y}. \tag{5}$$

To find the optimal solution to the problem (5), we let

$$\Phi = \int_{\mathbb{R}^p} \varphi\left(\mathbf{y}\right) d\mathbf{y}$$

and take the partial derivative of α_s with respect to Φ. Thus, we obtain

$$\frac{\partial \Phi}{\partial \alpha_s} = \int_{\mathbb{R}^p} \mathbb{Cov}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right), H_s\left(\mathbf{y}\right)\right] d\mathbf{y} - \sum_{s=1}^{n} \alpha_s \int_{\mathbb{R}^p} \mathbb{Cov}\left[H_t\left(\mathbf{y}\right), H_s\left(\mathbf{y}\right)\right] d\mathbf{y}. \tag{6}$$

According to the double expectation formula, we have

$$\begin{aligned} & \mathbb{Cov}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right), H_s\left(\mathbf{y}\right)\right] \\ = \; & \mathbb{E}\left[\mathbb{Cov}\left(F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right), H_s\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] + \mathbb{Cov}\left\{F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right), \mathbb{E}\left[H_s\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right]\right\} \\ = \; & \mathbb{V}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] \end{aligned} \tag{7}$$

and

$$\begin{aligned} \mathbb{Cov}\left[H_t\left(\mathbf{y}\right), H_s\left(\mathbf{y}\right)\right] &= \mathbb{E}\left[\mathbb{Cov}\left(H_t\left(\mathbf{y}\right), H_s\left(\mathbf{y}\right)|\,\boldsymbol{\Theta}\right)\right] + \mathbb{V}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] \\ &= \begin{cases} \mathbb{V}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] & t \neq s \\ \mathbb{E}\left[\mathbb{V}\left(H_t\left(\mathbf{y}\right)|\,\boldsymbol{\Theta}\right)\right] + \mathbb{V}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] & t = s \end{cases}. \end{aligned} \tag{8}$$

Substituting equations (7) and (8) into (6) and setting it equal to 0, we obtain

$$\begin{aligned} \frac{\partial \Phi}{\partial \alpha_s} &= (1 - \sum_{s=1}^{n} \alpha_s) \int_{\mathbb{R}^p} \mathbb{V}\left[F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)\right] d\mathbf{y} - \alpha_t \int_{\mathbb{R}^p} \mathbb{E}\left[\mathbb{V}\left(H_t\left(\mathbf{y}\right)|\,\boldsymbol{\Theta}\right)\right] d\mathbf{y} \\ &= -\sum_{s=1}^{n} \alpha_s \tau_0^2 - \alpha_t \sigma_0^2 + \tau_0^2 = 0, \quad t = 1, 2, \cdots, n. \end{aligned}$$

Summing up the above equation for t from 1 to n, it gives

$$\sum_{s=1}^{n} \alpha_s = \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2}.$$

Furthermore, the solution for α_t is given by

$$\alpha_t = \frac{\tau_0^2}{n\tau_0^2 + \sigma_0^2}, \quad t = 1, 2, \cdots, n. \tag{9}$$

Substituting equations (4) and (9) into (13) of the main paper, we can obtain the optimal linear estimate of $F\left(\mathbf{y}|\,\boldsymbol{\Theta}\right)$ as

$$\begin{aligned} \widehat{F}\left(\mathbf{y}|\,\boldsymbol{\Theta}\right) &= \left(1 - \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2}\right) F_0\left(\mathbf{y}\right) + \frac{\tau_0^2}{n\tau_0^2 + \sigma_0^2} \sum_{i=1}^{n} H_i\left(\mathbf{y}\right) \\ &= \left(1 - \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2}\right) F_0\left(\mathbf{y}\right) + \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2} F_n\left(\mathbf{y}\right) \\ &= Z_{N,n} F_n\left(\mathbf{y}\right) + (1 - Z_{N,n}) F_0\left(\mathbf{y}\right). \end{aligned}$$

4 Proof of Proposition 3.1

By applying the plug-in method, we have

$$\begin{aligned}\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} &= \int_{\mathbb{R}^p} \mathbf{y} d\left[Z_{N,n}F_n(\mathbf{y}) + (1-Z_{N,n})F_0(\mathbf{y})\right] \\ &= Z_{N,n}\int_{\mathbb{R}^p} \mathbf{y} dF_n(\mathbf{y}) + (1-Z_{N,n})\int_{\mathbb{R}^p} \mathbf{y} dF_0(\mathbf{y}) \\ &= Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_{\mathbf{0}},\end{aligned}$$

which yields the desired result.

5 Proof of Proposition 3.3

Due to

$$\mathbb{E}\left[\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] = \mathbb{E}\left\{\mathbb{E}\left[\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\middle|\,\boldsymbol{\Theta}\right]\right\} = 0,$$

it follows that

$$\begin{aligned}&\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\Theta)\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] \\ = \;& \mathbb{E}\left[\left(Z_{N,n}\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)+(1-Z_{N,n})\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\right)\right. \\ &\left.\left(\left(Z_{N,n}\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)+(1-Z_{N,n})\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\right)\right)'\right] \\ = \;& Z_{N,n}^2\mathbb{E}\left[\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right]+(1-Z_{N,n})^2\mathbb{E}\left[\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\boldsymbol{\mu}_0-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] \\ = \;& Z_{N,n}^2\mathbb{E}\left\{\mathbb{E}\left[\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\overline{\mathbf{Y}}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\middle|\,\boldsymbol{\Theta}\right]\right\}+(1-Z_{N,n})^2\mathbb{V}\left[\boldsymbol{\mu}(\boldsymbol{\Theta})\right] \\ = \;& \frac{Z_{N,n}^2}{n}\Sigma_0+(1-Z_{N,n})^2T.\end{aligned}$$

Additionally, as the sample size n tends to infinity ($n\to\infty$), we have:

$$\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right]\to 0_{p\times p},$$

where $0_{p\times p}$ represents the zero matrix. Furthermore, for any p-dimensional real vector $\boldsymbol{\xi}$, we have

$$\boldsymbol{\xi}'\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right]\boldsymbol{\xi} = \mathbb{E}\left[\boldsymbol{\xi}'\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\right]^2\to 0.$$

Based on the arbitrariness of $\boldsymbol{\xi}$, we conclude that:

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(v)}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right]\to 0.$$

6 Proof of Proposition 3.4

Since

$$\Sigma_n = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{Y}_i - \overline{\mathbf{Y}}\right)\left(\mathbf{Y}_i - \overline{\mathbf{Y}}\right)' = U_n - \overline{\mathbf{Y}}\cdot\overline{\mathbf{Y}}'$$

and

$$\Sigma_0 = \mathbb{E}\left(\Sigma(\mathbf{\Theta})\right) = U_0 - \boldsymbol{\mu}_0\boldsymbol{\mu}_0',$$

it follows that

$$\begin{aligned}
\widehat{\Sigma_{N,n}(\mathbf{\Theta})} &= \int_{\mathbb{R}^p}\left[\mathbf{y} - \widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})}\right]\left[\mathbf{y} - \widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})}\right]' d\widehat{F}\left(\mathbf{y}|\,\mathbf{\Theta}\right)\\
&= \int_{\mathbb{R}^p}\mathbf{y}\mathbf{y}'d\widehat{F}\left(\mathbf{y}|\,\mathbf{\Theta}\right) - \widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})}\cdot\widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})}'\\
&= Z_{N,n}U_n + (1-Z_{N,n})U_0 - \left(Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0\right)\left(Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0\right)'\\
&= Z_{N,n}^2\left(U_n - \overline{\mathbf{Y}}\cdot\overline{\mathbf{Y}}'\right) + (1-Z_{N,n})^2\left(U_0 - \boldsymbol{\mu}_0\boldsymbol{\mu}_0'\right)\\
&\quad + Z_{N,n}(1-Z_{N,n})\left(U_n - \overline{\mathbf{Y}}\boldsymbol{\mu}_0' - \boldsymbol{\mu}_0\overline{\mathbf{Y}}' + U_0\right)\\
&= \omega_{1,n}\Sigma_n + \omega_{2,n}\Sigma_0 + (1-\omega_{1,n}-\omega_{2,n})M_0.
\end{aligned}$$

7 Proof of Theorem 3.2

By the strong law of large numbers, we have

$$\overline{\mathbf{Y}} \to \boldsymbol{\mu}(\mathbf{\Theta}),\ a.s.,\quad \text{and}\quad \frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i\mathbf{Y}_i' \to \mathbb{E}\left(\mathbf{Y}_1\mathbf{Y}_1'|\mathbf{\Theta}\right),\ a.s.,$$

when $n\to\infty$. Using the continuity theorem of almost sure convergence, it gives

$$\Sigma_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i\mathbf{Y}_i' - \overline{\mathbf{Y}}\cdot\overline{\mathbf{Y}}' \to \Sigma(\mathbf{\Theta}),\ a.s.$$

Furthermore, from (23) of the main paper, we have $Z_{N,n}\to 1$ when $n\to\infty$. Hence, we obtain

$$\widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})} = Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0 \to \boldsymbol{\mu}(\mathbf{\Theta}),\ a.s.$$

and

$$\widehat{\Sigma_{N,n}(\mathbf{\Theta})} = \omega_{1,n}\Sigma_n + \omega_{2,n}\Sigma_0 + (1-\omega_{1,n}-\omega_{2,n})M_0 \to \Sigma(\mathbf{\Theta}),\ a.s.$$

8 Proof of Theorem 3.3

Note that

$$\sqrt{n}\left(\widehat{\boldsymbol{\mu}_{N,n}(\mathbf{\Theta})} - \boldsymbol{\mu}(\mathbf{\Theta})\right) = \sqrt{n}\left(\overline{\mathbf{Y}} - \boldsymbol{\mu}(\mathbf{\Theta})\right) + \sqrt{n}(1-Z_{N,n})\left(\boldsymbol{\mu}_0 - \overline{\mathbf{Y}}\right).$$

Obviously, the sample mean $\overline{\mathbf{Y}}$ and the aggregate mean $\boldsymbol{\mu}_0$ are both bounded in probability, when $n \to \infty$, we have

$$(1 - Z_{N,n})\sqrt{n} = O\left(\frac{1}{\sqrt{n}}\right).$$

Thus, it gives

$$\sqrt{n}(1 - Z_{N,n})\left(\boldsymbol{\mu}_0 - \overline{\mathbf{Y}}\right) \xrightarrow{P} 0.$$

Furthermore, by the multidimensional central limit theorem of independent and identically distributed random variables (cf. Serfling, 2009), we have

$$\sqrt{n}\left(\overline{\mathbf{Y}} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right) \xrightarrow{L} N(0, \Sigma(\boldsymbol{\Theta})).$$

Therefore, by the Slutsky's theorem, we obtain

$$\sqrt{n}\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right) = \sqrt{n}\left(\overline{\mathbf{Y}} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right) \xrightarrow{L} N(0, \Sigma(\boldsymbol{\Theta})).$$

9 Proof of Proposition 3.5

Firstly, for the traditional multivariate credibility estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$, the credibility factor matrix is give by

$$Z_{C,n} = \begin{pmatrix} Z_{11}^C & Z_{12}^C \\ Z_{21}^C & Z_{22}^C \end{pmatrix} = \begin{pmatrix} \frac{n\tau_1^2\delta_2 - n\nu_1\delta_3}{\delta_1\delta_2 - \delta_3^2} & \frac{n\nu_1\sigma_2^2 - n\nu_2\tau_2^2}{\delta_1\delta_2 - \delta_3^2} \\ \frac{n\nu_1\sigma_1^2 - n\nu_2\tau_1^2}{\delta_1\delta_2 - \delta_3^2} & \frac{n\tau_2^2\delta_1 - n\nu_1\delta_3}{\delta_1\delta_2 - \delta_3^2} \end{pmatrix}.$$

Furthermore, according to Theorem 2.1, we have

$$\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} = \begin{pmatrix} \widehat{\mu_{C,n}^{(1)}(\Theta)} \\ \widehat{\mu_{C,n}^{(2)}(\Theta)} \end{pmatrix} = \begin{pmatrix} Z_{11}^C \overline{Y}^{(1)} + \left(1 - Z_{11}^C\right)\mu_1 + Z_{12}^C\left(\overline{Y}^{(2)} - \mu_2\right) \\ Z_{22}^C \overline{Y}^{(2)} + \left(1 - Z_{22}^C\right)\mu_2 + Z_{21}^C\left(\overline{Y}^{(1)} - \mu_1\right) \end{pmatrix}.$$

Using the double expectation formula, we have

$$\begin{aligned}
& \mathbb{E}\left[\left(\widehat{\mu_{C,n}^{(1)}(\Theta)} - \mu_1(\Theta)\right)^2\right] \\
= \; & \mathbb{E}\left[\left(Z_{11}^C(\overline{Y}^{(1)} - \mu_1(\Theta)) + (1 - Z_{11}^C)(\mu_1 - \mu_1(\Theta)) + Z_{12}^C(\overline{Y}^{(2)} - \mu_2)\right)^2\right] \\
= \; & \left(Z_{11}^C\right)^2 \mathbb{E}\left[\left(\overline{Y}^{(1)} - \mu_1(\Theta)\right)^2\right] + \left(1 - Z_{11}^C\right)^2 \mathbb{E}\left[(\mu_1 - \mu_1(\Theta))^2\right] \\
& + \left(Z_{12}^C\right)^2 \mathbb{E}\left[\left(\overline{Y}^{(2)} - \mu_2\right)^2\right] + 2Z_{12}^C(1 - Z_{11}^C)\mathbb{E}\left[(\mu_1 - \mu_1(\Theta))(\overline{Y}^{(2)} - \mu_2)\right] \\
= \; & \left(Z_{11}^C\right)^2 \frac{\sigma_1^2}{n} + \left(1 - Z_{11}^C\right)^2 \tau_1^2 + \left(Z_{12}^C\right)^2\left(\frac{\sigma_2^2}{n} + \tau_2^2\right) - 2\nu_1 Z_{12}^C(1 - Z_{11}^C).
\end{aligned}$$

Similarly, we obtain

$$\mathbb{E}\left[\left(\widehat{\mu^{(2)}_{C,n}(\Theta)}-\mu_2(\Theta)\right)^2\right] = \left(Z^C_{22}\right)^2\frac{\sigma_2^2}{n}+\left(1-Z^C_{22}\right)^2\tau_2^2 + \left(Z^C_{21}\right)^2\left(\frac{\sigma_1^2}{n}+\tau_1^2\right)-2\nu_1 Z^C_{21}(1-Z^C_{22}).$$

Therefore, the mean of the weighted F-norm error of $\widehat{\boldsymbol{\mu}_{C,n}(\Theta)}$ is

$$\begin{aligned}
& \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|^2_{\boldsymbol{\xi}}\right]\\
= \; & \mathbb{E}\left[\xi_1\left(\widehat{\mu^{(1)}_{C,n}(\Theta)}-\mu_1(\Theta)\right)^2\right]+\mathbb{E}\left[\xi_2\left(\widehat{\mu^{(2)}_{C,n}(\Theta)}-\mu_2(\Theta)\right)^2\right]\\
= \; & \left(\xi_1\left(Z^C_{11}\right)^2+\xi_2\left(Z^C_{21}\right)^2\right)\frac{\delta_1}{n}+\left(\xi_1\left(Z^C_{22}\right)^2+\xi_2\left(Z^C_{12}\right)^2\right)\frac{\delta_2}{n}\\
& -2\nu_1\left(\xi_2 Z^C_{21}\left(1-Z^C_{22}\right)+\xi_1 Z^C_{12}\left(1-Z^C_{11}\right)\right)\\
& +\xi_2\left(1-2Z^C_{22}\right)\tau_2^2+\xi_1\left(1-2Z^C_{11}\right)\tau_1^2\\
= \; & \Delta_1+\Delta_2+\Delta_3+\Delta_4.
\end{aligned}$$

On the other hand, according to Proposition 3.1, we have

$$\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}=\begin{pmatrix}\widehat{\mu^{(1)}_{N,n}(\Theta)}\\ \widehat{\mu^{(2)}_{N,n}(\Theta)}\end{pmatrix}=\begin{pmatrix}Z_{N,n}\overline{Y}^{(1)}+(1-Z_{N,n})\mu_1\\ Z_{N,n}\overline{Y}^{(2)}+(1-Z_{N,n})\mu_2\end{pmatrix}.$$

Similarly, it gives

$$\begin{aligned}
\mathbb{E}\left[\left(\widehat{\mu^{(1)}_{N,n}(\Theta)}-\mu_1(\Theta)\right)^2\right] =& \mathbb{E}\left[\left(Z_{N,n}\left(\overline{Y}^{(1)}-\mu_1(\Theta)\right)+(1-Z_{N,n})(\mu_1-\mu_1(\Theta))\right)^2\right]\\
=& Z^2_{N,n}\mathbb{E}\left[\left(\overline{Y}^{(1)}-\mu_1(\Theta)\right)^2\right]+(1-Z_{N,n})^2\mathbb{E}\left[(\mu_1-\mu_1(\Theta))^2\right]\\
& +2Z_{N,n}(1-Z_{N,n})\mathbb{E}\left[\left(\overline{Y}^{(1)}-\mu_1(\Theta)\right)(\mu_1-\mu_1(\Theta))\right]\\
=& Z^2_{N,n}\mathbb{E}\left[\sigma_1^2(\Theta)/n\right]+(1-Z_{N,n})^2\mathbb{V}(\mu_1(\Theta))\\
=& Z^2_{N,n}\frac{\sigma_1^2}{n}+(1-Z_{N,n})^2\tau_1^2
\end{aligned}$$

and

$$\mathbb{E}\left[\left(\widehat{\mu^{(2)}_{N,n}(\Theta)}-\mu_2(\Theta)\right)^2\right]=Z^2_{N,n}\frac{\sigma_2^2}{n}+(1-Z_{N,n})^2\tau_2^2.$$

Thus, the mean of the weighted F-norm error of $\widehat{\boldsymbol{\mu}_{N,n}(\Theta)}$ is given by

$$\begin{aligned}
\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|^2_{\boldsymbol{\xi}}\right] &= \mathbb{E}\left[\xi_1\left(\widehat{\mu^{(1)}_{N,n}(\Theta)}-\mu_1(\Theta)\right)^2\right]+\mathbb{E}\left[\xi_2\left(\widehat{\mu^{(2)}_{N,n}(\Theta)}-\mu_2(\Theta)\right)^2\right]\\
&= \frac{n\tau_0^4\left(\xi_1\sigma_1^2+\xi_2\sigma_2^2\right)}{\left(n\tau_0^2+\sigma_0^2\right)^2}+\frac{\sigma_0^4\left(\xi_1\tau_1^2+\xi_2\tau_2^2\right)}{\left(n\tau_0^2+\sigma_0^2\right)^2}.
\end{aligned}$$

10 Proof of Proposition 3.6

Firstly, according to equation (26) of the main paper, when $\rho_1 = \rho_2 = 0$, we have

$$Z_{C,n} = \begin{pmatrix} \frac{n\tau_1^2}{n\tau_1^2+\sigma_1^2} & 0 \\ 0 & \frac{n\tau_2^2}{n\tau_2^2+\sigma_2^2} \end{pmatrix}.$$

Therefore, we obtain the mean of the weighted F-norm for the error of $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ as follows:

$$\begin{aligned} \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] &= \xi_1 \mathbb{E}\left[\left(\widehat{\mu_{C,n}^{(1)}(\Theta)} - \mu_1(\Theta)\right)^2\right] + \xi_2 \mathbb{E}\left[\left(\widehat{\mu_{C,n}^{(2)}(\Theta)} - \mu_2(\Theta)\right)^2\right] \\ &= \xi_1 \frac{\tau_1^2\sigma_1^2}{n\tau_1^2+\sigma_1^2} + \xi_2 \frac{\tau_2^2\sigma_2^2}{n\tau_2^2+\sigma_2^2}. \end{aligned}$$

Furthermore, according to equation (25) of the main paper, we obtain

$$\begin{aligned} & \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] - \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] \\ = \;& \xi_1\left(\frac{\tau_1^2\sigma_1^2}{n\tau_1^2+\sigma_1^2} - \frac{n\tau_0^4\sigma_1^2+\sigma_0^4\tau_1^2}{(n\tau_0^2+\sigma_0^2)^2}\right) + \xi_2\left(\frac{\tau_2^2\sigma_2^2}{n\tau_2^2+\sigma_2^2} - \frac{n\tau_0^4\sigma_2^2+\sigma_0^4\tau_2^2}{(n\tau_0^2+\sigma_0^2)^2}\right) \\ = \;& \xi_1 \frac{-n(\tau_0^2\sigma_1^2-\sigma_0^2\tau_1^2)^2}{(n\tau_1^2+\sigma_1^2)(n\tau_0^2+\sigma_0^2)^2} + \xi_2 \frac{-n(\tau_0^2\sigma_2^2-\sigma_0^2\tau_2^2)^2}{(n\tau_2^2+\sigma_2^2)(n\tau_0^2+\sigma_0^2)^2} \\ \leq \;& 0. \end{aligned}$$

References

Serfling, R. J. (2009). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.