LPRE criterion based estimating equation approaches for the error-in-covariables multiplicative regression models

Qihua Wang [1,2] and Dahai Hu[3]

[1] *School of Statistics and Mathematics, Zhejiang Gongshang University Hangzhou, Zhejiang 310018, China.*

[2]Academy of Mathematics and Systems, Chinese Academy of Sciences *Beijing 100190, China.*

[3] *Department of Statistics and Finance, University of Science and Technology of China Hefei Anhui, 230026, China.*

This supplementary material contains the further simulation where the assumptions are violated and the proofs of the theorems in detail.

S1 Further simulation

Theoretically, the proposed approaches require the assumption of $E(\varepsilon - 1/\varepsilon|Z) = 0$. We also carried out some simulation studies to assess the performances of the proposed estimators when the assumption is violated. In the uniform measurement error scenario with $n = 200$, we considered two distributions of ε: $\log\varepsilon \sim Beta(2,4)$ and $\log\varepsilon \sim 0.5\times t(5)$. For the first distribution, $E(\varepsilon - 1/\varepsilon|Z) \neq 0$. For the second distribution, $E(\varepsilon|Z)$ does not exist. The simulation results was displayed in Table S1.

Insert Table S1 here

When $E(\varepsilon - 1/\varepsilon|Z) \neq 0$, both the naive estimators for c_0, α_0 and γ_0 are biased seriously. However, two proposed methods and the classical CLS method are of small bias for the parameters α_0 and γ_0. This implies that the three methods can still correct the bias for the estimators of α_0 and γ_0 even if the the assumption for ε is violated. It is also noted that all the estimators for the intercept term c_0 are biased seriously. The reason may be that both $E[\log\varepsilon|Z]$ and $E(\varepsilon - 1/\varepsilon|Z)$ are not equal to zero. According to the least square theory, $E[\log\varepsilon|Z] \neq 0$ yields biased intercept. Similarly, for the multiplicative regression model, when $E(\varepsilon - 1/\varepsilon|Z) \neq 0$, there exists a positive value a such that $E(a\varepsilon - 1/(a\varepsilon)|Z) = 0$, then $a\varepsilon$ can be treated as the new model error term. This implies that the proposed estimators cause bias for the intercept term c_0. When $E(\varepsilon|Z)$ does not exist, our proposed CMS and CEE methods can also correct the bias but with larger standard error than the classical CLS method.

Also, our proposed CMS and CEE methods require the symmetry and the low tail of the measurement error U(Condition C4 and C5 in the Appendix) for the asymptotic properties. Simulation studies were carried out to examine our proposed estimators when the assumptions are violated. In the scenario of $n = 200$ and $\log\varepsilon \sim N(0, 0.25)$, we generated U in three cases. In case 1, U was generated from the standardized $Beta(2, 4)$ distribution and scaled to have standard deviation of 0.5, and hence the skewness of U is 0.467 and the symmetry condition is moderately violated. In case 2, U was generated from the standardized χ_1^2 truncated at 5 and divided by 2, and hence the measurement error U is extremely skewed with skewness 1.68. In case 3, U was generated from $t(5)$ and scaled to have standard deviation 0.5 and hence the Condition C5 in the Appendix is violated. The simulation results was displayed in Table S2.

Insert Table S2 here

In all the cases considered, both the proposed estimators and the CLS estimator of α_0 and γ_0 are of far smaller bias than both the naive estimators, and hence the three method can still correct the bias for the estimators of α_0 and γ_0. Except for CMS, the bias of the estimators of c_0 for all the other methods are also small. In conclusion, the proposed CMS method is sensitive to the assumption of U whereas the proposed CEE method and the CLS method are robust.

S2 The proof of Theorem

Proof of theorem 1

Proof: For simplification, denote $U_{cms}(\beta) = \sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} \hat{T}^*(\mathcal{O}_{i,r}, \beta)$.

Step 1: Proof of asymptotic normality of $n^{-1/2}U_{cms}(\beta_0)$.

Note that

$$\begin{aligned} U_{cms}(\beta_0) &= \sum_{i=1}^{n} [Y_i^{-1} \hat{R}_i^{(1)}(\beta_0) - Y_i \hat{R}_i^{(1)}(-\beta_0)] \\ &= \sum_{i=1}^{n} Y_i^{-1} [\hat{R}_i^{(1)}(\beta_0) - R_i^{(1)}(\beta_0)] - \sum_{i=1}^{n} Y_i [\hat{R}_i^{(1)}(-\beta_0) - R_i^{(1)}(-\beta_0)] \\ &\quad + \sum_{i=1}^{n} [Y_i^{-1} R_i^{(1)}(\beta_0) - Y_i R_i^{(1)}(-\beta_0)] \\ &:= A_1 - A_2 + A_3, \end{aligned}$$

and

$$\begin{aligned} &\hat{R}_i^{(1)}(\beta_0) - R_i^{(1)}(\beta_0) \\ =&\{\hat{\varphi}_0^{-1}(\gamma_0) - \varphi_0^{-1}(\gamma_0)\} n_i^{-1} \sum_{r=1}^{n_i} \exp(\hat{Z}_{i,r}^T \beta_0) \hat{Z}_{i,r} \\ &- J\{\hat{\varphi}_0^{-2}(\gamma_0)\hat{\varphi}_1(\gamma_0) - \varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)\} n_i^{-1} \sum_{r=1}^{n_i} \exp(\hat{Z}_{i,r}^T \beta_0). \end{aligned}$$

Thus, the term A_1 can be decomposed as $A_1 = A_{1,1} - A_{1,2}$, where

$$\begin{aligned} A_{1,1} =& \{\hat{\varphi}_0^{-1}(\gamma_0) - \varphi_0^{-1}(\gamma_0)\} \sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} Y_i^{-1} \exp(\hat{Z}_{i,r}^T \beta_0) \hat{Z}_{i,r}, \\ A_{1,2} =& J\{\hat{\varphi}_0^{-2}(\gamma_0)\hat{\varphi}_1(\gamma_0) - \varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)\} \sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} Y_i^{-1} \exp(\hat{Z}_{i,r}^T \beta_0). \end{aligned}$$

For $A_{1,1}$ and $A_{1,2}$, we have

$$\begin{aligned} A_{1,1} =& - \left\{\varphi_0^{-1}(\gamma_0)E(\varepsilon^{-1}Z) + J\varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)E(\varepsilon^{-1})\right\}n\{\hat{\varphi}_0(\gamma_0) - \varphi_0(\gamma_0)\} \\ &+ o_p(n^{1/2}), \\ A_{1,2} =& - 2J\varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)E(\varepsilon^{-1})n\{\hat{\varphi}_0(\gamma_0) - \varphi_0(\gamma_0)\} \\ &+ J\varphi_0^{-1}(\gamma_0)E(\varepsilon^{-1})n\{\hat{\varphi}_1(\gamma_0) - \varphi_1(\gamma_0)\} + o_p(n^{1/2}). \end{aligned}$$

Thus, we can obtain

$$\begin{aligned} A_1 =&\{-\varphi_0^{-1}(\gamma_0)E(\varepsilon^{-1}Z) + J\varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)E(\varepsilon^{-1})\}n\{\hat{\varphi}_0(\gamma_0) - \varphi_0(\gamma_0)\} \\ &- J\varphi_0^{-1}(\gamma_0)E(\varepsilon^{-1})n\{\hat{\varphi}_1(\gamma_0) - \varphi_1(\gamma_0)\} + o_p(n^{1/2}). \end{aligned}$$

Similar to A_1, the term A_2 can be represented as

$$\begin{aligned} A_2 =& - \{\varphi_0^{-1}(\gamma_0)E(\varepsilon Z) + J\varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)E(\varepsilon)\}n\{\hat{\varphi}_0(\gamma_0) - \varphi_0(\gamma_0)\} \\ &+ J\varphi_0^{-1}(\gamma_0)E(\varepsilon)n\{\hat{\varphi}_1(\gamma_0) - \varphi_1(\gamma_0)\} + o_p(n^{1/2}). \end{aligned}$$

Therefore, according to Condition C3, we have

$$\begin{aligned} A_1 - A_2 =&JE(\varepsilon + \varepsilon^{-1})\varphi_0^{-2}(\gamma_0)\varphi_1(\gamma_0)n\{\hat{\varphi}_0(\gamma_0) - \varphi_0(\gamma_0)\} \\ &- JE(\varepsilon + \varepsilon^{-1})\varphi_0^{-1}(\gamma_0)n\{\hat{\varphi}_1(\gamma_0) - \varphi_1(\gamma_0)\} + o_p(n^{1/2}). \end{aligned} \tag{S2.1}$$

Given any γ, by definition

$$\{\hat{\varphi}_0(\gamma) + \varphi_0(\gamma)\}\{\hat{\varphi}_0(\gamma) - \varphi_0(\gamma)\} = \tilde{n}^{-1}\sum_{i=1}^{n}\xi_i\{h_i^{(0)}(\gamma) - \varphi_0^2(\gamma)\}. \tag{S2.2}$$

Owning to (S2.2), the consistency of $\hat{\varphi}_0(\gamma)$ and $\lim \tilde{n}/n = 1 - \rho_1$, we obtain

$$
\begin{aligned}
&n\{\hat{\varphi}_0(\gamma)-\varphi_0(\gamma)\}\\
=&\{2(1-\rho_1)\varphi_0(\gamma)\}^{-1}\sum_{i=1}^{n}\xi_i\{h_i^{(0)}(\gamma)-\varphi_0^2(\gamma)\}+o_p(n^{1/2}).
\end{aligned}
\tag{S2.3}
$$

Clearly,

$$
\hat{\varphi}_0(\gamma)\hat{\varphi}_1(\gamma)-\varphi_0(\gamma)\varphi_1(\gamma)=(2\tilde{n})^{-1}\sum_{i=1}^{n}\xi_i\{h_i^{(1)}(\gamma)-2\varphi_0(\gamma)\varphi_1(\gamma)\}. \tag{S2.4}
$$

According to (S2.3) and the consistency of $\hat{\varphi}_1(\gamma)$, we obtain

$$
\begin{aligned}
&n\{\hat{\varphi}_0(\gamma)\hat{\varphi}_1(\gamma)-\varphi_0(\gamma)\varphi_1(\gamma)\}\\
=&n\{\hat{\varphi}_0(\gamma)\hat{\varphi}_1(\gamma)-\varphi_0(\gamma)\hat{\varphi}_1(\gamma)\}+n\{\varphi_0(\gamma)\hat{\varphi}_1(\gamma)-\varphi_0(\gamma)\varphi_1(\gamma)\}\\
=&n\{\hat{\varphi}_0(\gamma)-\varphi_0(\gamma)\}\hat{\varphi}_1(\gamma)+n\{\hat{\varphi}_1(\gamma)-\varphi_1(\gamma)\}\varphi_0(\gamma)\\
=&n\{\hat{\varphi}_0(\gamma)-\varphi_0(\gamma)\}\varphi_1(\gamma)+n\{\hat{\varphi}_1(\gamma)-\varphi_1(\gamma)\}\varphi_0(\gamma)+o_p(n^{1/2}).
\end{aligned}
\tag{S2.5}
$$

Combining (S2.3), (S2.4) and (S2.5), we have

$$
\begin{aligned}
&n\{\hat{\varphi}_1(\gamma)-\varphi_1(\gamma)\}\\
=&\{2(1-\rho_1)\varphi_0(\gamma)\}^{-1}\sum_{i=1}^{n}\xi_i\{h_i^{(1)}(\gamma)\\
&-\varphi_0^{-1}(\gamma)\varphi_1(\gamma)h_i^{(0)}(\gamma)-\varphi_0(\gamma)\varphi_1(\gamma)\}+o_p(n^{1/2}).
\end{aligned}
\tag{S2.6}
$$

Plugging both (S2.3) and (S2.6) into (S2.1) yields $A_1-A_2=-\sum_{i=1}^{n}\xi_iJr_i+o_p(n^{1/2})$. Owing to the fact that $E(h_i^{(0)}(\gamma))=\varphi_0^2(\gamma)$ and $E(h_i^{(1)}(\gamma))=2\varphi_0(\gamma)\varphi_1(\gamma)$ for any given γ, we have that $E(r_i)=0$. Furthermore, $A_3=\sum_{i=1}^{n}v_i$, and $E(v_i)=0$, under Condition C3. Summarizing the preceding results, we have

$$U_{cms}(\beta_0) = \sum_{i=1}^{n}\{v_i - \xi_i J r_i\} + o_p(n^{1/2}).$$

From the above discussion, it follows that $v_i - \xi_i J r_i, i \in \mathcal{A}_k$ is i.i.d with mean zero and the terms between $\mathcal{A}_k$ and $\mathcal{A}_l$ are independent for $k \neq l$. It then follows from the multivariate central limit theorem that

$$n^{-1/2}U_{cms}(\beta_0) \xrightarrow{D} N(0, \Sigma_{cms}).$$

Step 2: Proof of asymptotic normality of $n^{1/2}(\hat{\beta}_{cms} - \beta_0)$.

Denote $V_n(\beta) = n^{-1}\partial U_{cms}(\beta)/\partial\beta^T$ and recall that $\hat{R}_i^{(2)}(\beta) = \partial\hat{R}_i^{(1)}(\beta)/\partial\beta^T$, then

$$V_n(\beta) = n^{-1}\sum_{i=1}^{n}\{Y_i^{-1}\hat{R}_i^{(2)}(\beta) + Y_i\hat{R}_i^{(2)}(-\beta)\}.$$

Take $R_i^{(2)}(\beta) = \partial R_i^{(1)}(\beta)/\partial\beta^T$. A simple calculation yields

$$\begin{aligned} R_i^{(2)}(\beta) =& \{n_i\varphi_0^2(\gamma)\}^{-1}\sum_{r=1}^{n_i}\exp(\hat{Z}_{i,r}^T\beta)\big[\varphi_0(\gamma)\hat{Z}_{i,r}^{\otimes 2} - J\varphi_1(\gamma)\hat{Z}_{i,r}^T \\ &- \hat{Z}_{i,r}\varphi_1^T(\gamma)J^T - J\varphi_2(\gamma)J^T + 2\varphi_0^{-1}(\gamma)\{J\varphi_1(\gamma)\}^{\otimes 2}\big], \end{aligned}$$

where $\varphi_2(\gamma) = \partial\varphi_1(\gamma)/\partial\gamma^T$. Denote $\hat{\varphi}_2(\gamma) = \partial\hat{\varphi}_1(\gamma)/\partial\gamma^T$. Owning to $\hat{\varphi}_1(\gamma) = \partial\hat{\varphi}_0(\gamma)/\partial\gamma$ and $\hat{\varphi}_2(\gamma) = \partial\hat{\varphi}_1(\gamma)/\partial\gamma^T$, $\hat{R}_i^{(2)}(\beta)$ can be obtained by replacing $\varphi_k(\gamma)$ in $R_i^{(2)}(\beta)$ with $\hat{\varphi}_k(\gamma)$, where $k = 0, 1, 2$. It follows from some simple calculation that $\varphi_2(\gamma) = E[U^{\otimes 2}\exp(U^T\gamma)]$ and $\hat{\varphi}_2(\gamma) =$

$\hat{\varphi}_0^{-1}(\gamma)\{(2\tilde{n})^{-1}\sum_{i=1}^n \xi_i h_i^{(2)}(\gamma) - \hat{\varphi}_1^{\otimes 2}(\gamma)\}$, where

$$h_i^{(2)}(\gamma) = \{n_i(n_i-1)\}^{-1}\sum_{r\neq s}(W_{i,r}-W_{i,s})^{\otimes 2}\exp\{(W_{i,r}-W_{i,s})^T\gamma\}.$$

Recall the expressions of $E[\exp(\hat{Z}_{i,r}^T\beta)|Z_i]$ and $E[\exp(\hat{Z}_{i,r}^T\beta)\hat{Z}_{i,r}|Z_i]$. We also note that

$$\begin{aligned} E[\exp(\hat{Z}_{i,r}^T\beta)\hat{Z}_{i,r}^{\otimes 2}|Z_i] = & \exp(Z_i^T\beta)\{\varphi_0(\gamma)Z_i^{\otimes 2} + J\varphi_1(\gamma)Z_i^T \\ & + Z_i\varphi_1^T(\gamma)J^T + J\varphi_2(\gamma)J^T\}. \end{aligned}$$

Therefore, we can obtain that $E[R_i^{(2)}(\beta)|Z_i] = \exp(Z_i^T\beta)Z_i^{\otimes 2}$. Under Condition C1, C4 and C5, it follows from the uniform law of large numbers that, in a neighborhood $\Theta \subseteq \mathcal{B}$ of β_0, $n^{-1}\sum_{i=1}^n\{Y_i^{-1}R_i^{(2)}(\beta) + Y_iR_i^{(2)}(-\beta)\}$ converges to a non-random function $V(\beta)$ in probability, where $V(\beta) = E[\{Y^{-1}\exp(Z^T\beta) + Y\exp(-Z^T\beta)\}Z^{\otimes 2}]$.

Recall that $\hat{R}_i^{(2)}(\beta)$ can be obtained from $R_i^{(2)}(\beta)$ by replacing $\varphi_k(\gamma)$ with $\hat{\varphi}_k(\gamma)(k=0,1,2)$. And $E[h_i^{(2)}(\gamma)] = 2\varphi_0(\gamma)\varphi_2(\gamma) + 2\varphi_1^{\otimes 2}(\gamma)$. Thereafter, under Condition C5, it follows from the uniform law of large numbers that $\sup_{\gamma\in\mathcal{B}}\|\hat{\varphi}_k(\gamma) - \varphi_k(\gamma)\| \to 0$ in probability, where $k = 0,1,2$. Thus, we can obtain that $\sup_{\beta\in\Theta}\|V_n(\beta) - V(\beta)\| \to 0$ in probability. Condition C2 guarantees that the limit of $V_n(\beta)$ is nonnegative definite everywhere and positive definite at β_0. It follows from the proof of Theorem 2 of Foutz (1977) that $\hat{\beta}_{cms}$ exists and is unique in Θ with probability converging to 1 as $n \to \infty$ and $\hat{\beta}_{cms} \xrightarrow{p} \beta_0$. By the Taylor series expansion,

$$0 = n^{-1/2} U_{cms}(\hat{\beta}_{cms}) = n^{-1/2} U_{cms}(\beta_0) + V_n(\beta^*)\sqrt{n}(\hat{\beta}_{cms} - \beta_0),$$

where β^* lies on the line segment between β_0 and $\hat{\beta}_{cms}$. Therefore,

$$\sqrt{n}(\hat{\beta}_{cms} - \beta_0) = -V_n^{-1}(\beta^*) n^{-1/2} U_{cms}(\beta_0) \xrightarrow{D} N(0, \Gamma_{cms}).$$

Proof of Theorem 2.

For simplification, denote $U_{cee}(\beta) = \sum_{i=1}^n \left\{ Y_i^{-1} \check{R}_i^{(1)}(\beta) - Y_i \check{R}_i^{(1)}(-\beta) \right\}$. Thus, $U_{cee}(\beta)$ can also be written as $U_{cee}(\beta) = \sum_{k=1}^m U_{c,k}(\beta)$, where

$$U_{c,k}(\beta) = \sum_{i \in \mathcal{A}_k} \left\{ Y_i^{-1} \check{R}_i^{(1)}(\beta) - Y_i \check{R}_i^{(1)}(-\beta) \right\}.$$

Step 1: Proof of asymptotic normality of $n^{-1/2} U_{cee}(\beta_0)$

Note that

$$\begin{aligned} U_{c,k}(\beta_0) = & \sum_{i \in \mathcal{A}_k} Y_i^{-1} [\check{R}_i^{(1)}(\beta_0) - \tilde{R}_i^{(1)}(\beta_0)] - \sum_{i \in \mathcal{A}_k} Y_i [\check{R}_i^{(1)}(-\beta_0) - \tilde{R}_i^{(1)}(-\beta_0)] \\ & + \sum_{i \in \mathcal{A}_k} [Y_i^{-1} \tilde{R}_i^{(1)}(\beta_0) - Y_i \tilde{R}_i^{(1)}(-\beta_0)] \\ := & A_{k,1} - A_{k,2} + A_{k,3}. \end{aligned}$$

Similar to the proof of Theorem 1 and according to Condition C3, we have

$$\begin{aligned} & A_{k,1} - A_{k,2} \\ = & \rho_k J E(\varepsilon + \varepsilon^{-1}) \eta_0^{-2}(k, \gamma_0) \eta_1(k, \gamma_0) n \{ \hat{\eta}_0(k, \gamma_0) - \eta_0(k, \gamma_0) \} \end{aligned}$$

$$- \rho_k JE(\varepsilon + \varepsilon^{-1})\eta_0^{-1}(k,\gamma_0)n\{\hat{\eta}_1(k,\gamma_0) - \eta_1(k,\gamma_0)\} + o_p(n^{1/2}). \quad \text{(S2.7)}$$

Recalling that $\eta_0(k,\gamma) = \varphi_0^k(\gamma/k)$ and $\hat{\eta}_0(k,\gamma) = \hat{\varphi}_0^k(\gamma/k)$, we have

$$\begin{aligned}
&n\{\hat{\eta}_0(k,\gamma) - \eta_0(k,\gamma)\} \\
=&n\{\hat{\varphi}_0^k(\gamma/k) - \varphi_0^k(\gamma/k)\} \\
=&n\{\hat{\varphi}_0(\gamma/k) - \varphi_0(\gamma/k)\} \cdot k\varphi_0^{k-1}(\gamma/k) + o_p(n^{1/2}).
\end{aligned} \quad \text{(S2.8)}$$

Since $\eta_1(k,\gamma) = \varphi_0^{k-1}(\gamma/k)\varphi_1(\gamma/k)$ and $\hat{\eta}_1(k,\gamma) = \hat{\varphi}_0^{k-1}(\gamma/k)\hat{\varphi}_1(\gamma/k)$, we have

$$\begin{aligned}
&n\{\hat{\eta}_1(k,\gamma) - \eta_1(k,\gamma)\} \\
=&n\{\hat{\varphi}_0^{k-1}(\gamma/k)\hat{\varphi}_1(\gamma/k) - \varphi_0^{k-1}(\gamma/k)\varphi_1(\gamma/k)\} \\
=&\hat{\varphi}_0^{k-1}(\gamma/k)n\{\hat{\varphi}_1(\gamma/k) - \varphi_1(\gamma/k)\} \\
&+ n\{\hat{\varphi}_0^{k-1}(\gamma/k) - \varphi_0^{k-1}(\gamma/k)\}\varphi_1(\gamma/k) \\
=&\varphi_0^{k-1}(\gamma/k)n\{\hat{\varphi}_1(\gamma/k) - \varphi_1(\gamma/k)\} \\
&+ (k-1)\varphi_0^{k-2}(\gamma/k)\varphi_1(\gamma/k)n\{\hat{\varphi}_0(\gamma/k) - \varphi_0(\gamma/k)\} + o_p(n^{1/2}).
\end{aligned} \quad \text{(S2.9)}$$

Plugging both (S2.8) and (S2.9) into (S2.7) yields

$$\begin{aligned}
&A_{k,1} - A_{k,2} \\
=&\rho_k JE(\varepsilon + \varepsilon^{-1})\varphi_0^{-2}(\gamma_0/k)\varphi_1(\gamma_0/k)n\{\hat{\varphi}_0(\gamma_0/k) - \varphi_0(\gamma_0/k)\} \\
&- \rho_k JE(\varepsilon + \varepsilon^{-1})\varphi_0^{-1}(\gamma_0/k)n\{\hat{\varphi}_1(\gamma_0/k) - \varphi_1(\gamma_0/k)\} + o_p(n^{1/2}).
\end{aligned} \quad \text{(S2.10)}$$

Combining (S2.10) with (S2.3) and (S2.6), we have

$$A_{k,1} - A_{k,2}$$

$$
\begin{aligned}
&= -\sum_{i=1}^{n}\Big[\xi_i\{2(1-\rho_1)\varphi_0^2(\gamma_0/k)\}^{-1}\rho_k J E(\varepsilon+\varepsilon^{-1}) \\
&\quad \times \big\{h_i^{(1)}(\gamma_0/k) - 2\varphi_0^{-1}(\gamma_0/k)\varphi_1(\gamma_0/k)h_i^{(0)}(\gamma_0/k)\big\}\Big] + o_p(n^{1/2}) \\
&= -J\sum_{i=1}^{n}\xi_i\rho_k\tilde{r}_{i,k} + o_p(n^{1/2}). \qquad \text{(S2.11)}
\end{aligned}
$$

Owing to the fact that $E[h_i^{(0)}(\gamma)] = \varphi_0^2(\gamma)$ and $E[h_i^{(1)}(\gamma)] = 2\varphi_0(\gamma)\varphi_1(\gamma)$ for any given γ, we have that $E(r_{i,k}) = 0$. Summarizing the preceding results, we have

$$
\begin{aligned}
U_{cee}(\beta_0) &= \sum_{k=1}^{m}(A_{k,1} - A_{k,2}) + \sum_{k=1}^{m}A_{k,3} \\
&= -J\sum_{i=1}^{n}\xi_i\sum_{k=1}^{m}\rho_k\tilde{r}_{i,k} + \sum_{i=1}^{n}\tilde{v}_i + o_p(n^{1/2}) \\
&= \sum_{i=1}^{n}\big\{\tilde{v}_i - \xi_i J\sum_{k=1}^{m}\rho_k\tilde{r}_{i,k}\big\} + o_p(n^{1/2}).
\end{aligned}
$$

From the above discussion, we have that $\tilde{v}_i - \xi_i J\sum_{k=1}^{m}\rho_k\tilde{r}_{i,k}$, $i \in \mathcal{A}_k$ is i.i.d with mean zero and the terms between $\mathcal{A}_k$ and $\mathcal{A}_l$ are independent for $k \neq l$. It then follows from the multivariate central limit theorem that

$$
n^{-1/2}U_{cee}(\beta_0) \xrightarrow{D} N(0, \Sigma_{cee}).
$$

Step 2: Proof of asymptotic normality of $n^{1/2}(\hat{\beta}_{cee} - \beta_0)$.

Denote $\tilde{V}_n(\beta) = n^{-1}\partial U_{cee}(\beta)/\partial\beta^T$ and recall that $\check{R}_i^{(2)}(\beta) = \partial\check{R}_i^{(1)}(\beta)/\partial\beta^T$, then

$$
\tilde{V}_n(\beta) = n^{-1}\sum_{i=1}^{n}\{Y_i^{-1}\check{R}_i^{(2)}(\beta) + Y_i\check{R}_i^{(2)}(-\beta)\}.
$$

Denote $\tilde{R}_i^{(2)}(\beta) = \partial \tilde{R}_i^{(1)}(\beta)/\partial \beta^T$. A simple calculation yields

$$\begin{aligned}\tilde{R}_i^{(2)}(\beta) =& \eta_0^{-2}(n_i,\gamma)\exp(\hat{Z}_i^T\beta)\big[\eta_0(n_i,\gamma)\hat{Z}_i^{\otimes 2} - J\eta_1(n_i,\gamma)\hat{Z}_i^T \\ & - \hat{Z}_i\eta_1^T(n_i,\gamma)J^T - J\eta_2(n_i,\gamma)J^T + 2\eta_0^{-1}(n_i,\gamma)\{J\eta_1(n_i,\gamma)\}^{\otimes 2}\big],\end{aligned}$$

where $\eta_2(n_i,\gamma) = \partial\eta_1(n_i,\gamma)/\partial\gamma^T$. Owning to the fact that

$$\begin{aligned}E[\exp(\hat{Z}_i^T\beta)\hat{Z}_i^{\otimes 2}|Z_i] =& \exp(Z_i^T\beta)\{\eta_0(n_i,\gamma)Z_i^{\otimes 2} + J\eta_1(n_i,\gamma)Z_i^T \\ & + Z_i\eta_1^T(n_i,\gamma)J^T + J\eta_2(n_i,\gamma)J^T\},\end{aligned}$$

we can obtain that $E[\tilde{R}_i^{(2)}(\beta)|Z_i] = \exp(Z_i^T\beta)Z_i^{\otimes 2}$. Under Condition C1, C4 and C5, it follows from the uniform law of large numbers that, in a neighborhood $\Theta \subseteq \mathcal{B}$ of β_0, $n^{-1}\sum_{i=1}^n\{Y_i^{-1}\tilde{R}_i^{(2)}(\beta) + Y_i\tilde{R}_i^{(2)}(-\beta)\}$ converges to a nonrandom function $V(\beta)$ in probability, where $V(\beta) = E[\{Y^{-1}\exp(Z^T\beta) + Y\exp(-Z^T\beta)\}Z^{\otimes 2}]$.

Take $\hat{\eta}_2(n_i,\gamma) = \partial\hat{\eta}_1(n_i,\gamma)/\partial\gamma^T$. Owning to $\hat{\eta}_1(n_i,\gamma) = \partial\hat{\eta}_0(n_i,\gamma)/\partial\gamma^T$ and $\hat{\eta}_2(n_i,\gamma) = \partial\hat{\eta}_1(n_i,\gamma)/\partial\gamma^T$, $\check{R}_i^{(2)}(\beta)$ can be obtained by replacing $\eta_k(n_i,\gamma)$ in $R_i^{(2)}(\beta)$ with $\hat{\eta}_k(n_i,\gamma)$, where $k = 0,1,2$. By a simple calculation, we have

$$\eta_2(n_i,\gamma) = (n_i-1)n_i^{-1}\varphi_0^{n_i-2}(\gamma/n_i)\varphi_1^{\otimes 2}(\gamma/n_i) + n_i^{-1}\varphi_0^{n_i-1}(\gamma/n_i)\varphi_2(\gamma/n_i).$$

Recall that $\eta_0(n_i,\gamma) = \varphi_0^{n_i}(\gamma/n_i)$ and $\eta_1(n_i,\gamma) = \varphi_0^{n_i-1}(\gamma/n_i)\varphi_1(\gamma/n_i)$. And $\hat{\eta}_k(n_i,\gamma)$ can be obtained by replacing the $\varphi_s(\cdot)$ in $\eta_k(n_i,\gamma)$ with $\hat{\varphi}_s(\cdot)$ where $k = 0,1,2$ and $s = 0,1,2$. In the proof of Theorem 1, we have known that under Condition C5, $\sup_{\gamma\in\mathcal{B}}\|\hat{\varphi}_k(\gamma)-\varphi_k(\gamma)\|\to 0$ in probability, where $k=0,1,2$. Thus, we have that $\sup_{\gamma\in\mathcal{B}}\|\hat{\eta}_k(n_i,\gamma)-\eta_k(n_i,\gamma)\|\to 0$ in probability, where $k=0,1,2$ and $n_i=1,\ldots,m$. Therefore, we can obtain that $\sup_{\beta\in\Theta}\|\tilde{V}_n(\beta)-V(\beta)\|\to 0$ in probability. Condition C2 guarantees that the limit of $\tilde{V}_n(\beta)$ is nonnegative definite everywhere and positive definite at β_0. It follows from the proof of Theorem 2 of Foutz (1977) that $\hat{\beta}_{cee}$ exists and is unique in Θ with probability converging to 1 as $n\to\infty$ and $\hat{\beta}_{cee}\xrightarrow{p}\beta_0$. By the Taylor series expansion,

$$0=n^{-1/2}U_{cee}(\hat{\beta}_{cee})=n^{-1/2}U_{cee}(\beta_0)+\tilde{V}_n(\beta^\star)\sqrt{n}(\hat{\beta}_{cee}-\beta_0),$$

where $\beta^\star$ lies on the line segment between β_0 and $\hat{\beta}_{cee}$. Therefore,

$$\sqrt{n}(\hat{\beta}_{cee}-\beta_0)=-\tilde{V}_n^{-1}(\beta^\star)n^{-1/2}U_{cee}(\beta_0)\xrightarrow{D}N(0,\Gamma_{cee}).$$

Proof of Theorem 3.

The proof of Theorem 3 can be obtained by taking $\hat{R}_i^{(1)}(\beta_0)=R_i^{(1)}(\beta_0)$ in Theorem 1 and $\check{R}_i^{(1)}(\beta_0)=\tilde{R}_i^{(1)}(\beta_0)$ in Theorem 2.

Proof of Lemma 1.

Firstly, we compute $E\big[\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0)-Y_iR_{i,r}^{(1)}(-\beta_0)\}^{\otimes 2}\big]$. Some simple algebraic manipulation yields that

$$\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0)-Y_iR_{i,r}^{(1)}(-\beta_0)\}^{\otimes 2}$$

$$
\begin{aligned}
&=\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0)-Y_iR_{i,r}^{(1)}(-\beta_0)\}\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0)-Y_iR_{i,r}^{(1)}(-\beta_0)\}^T\\
&=Y_i^{-2}R_{i,r}^{(1)}(\beta_0)^{\otimes 2}-R_{i,r}^{(1)}(\beta_0)R_{i,r}^{(1)}(-\beta_0)^T\\
&\quad -R_{i,r}^{(1)}(-\beta_0)R_{i,r}^{(1)}(\beta_0)^T+Y_i^2R_{i,r}^{(1)}(-\beta_0)^{\otimes 2}\\
&=Y_i^{-2}\varphi_0^{-2}(\gamma_0)\exp(2\hat{Z}_{i,r}^T\beta_0)\{\hat{Z}_{i,r}-J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^{\otimes 2}\\
&\quad -\varphi_0^{-2}(\gamma_0)\{\hat{Z}_{i,r}-J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}\{\hat{Z}_{i,r}+J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^T\\
&\quad -\varphi_0^{-2}(\gamma_0)\{\hat{Z}_{i,r}+J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}\{\hat{Z}_{i,r}-J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^T\\
&\quad +Y_i^2\varphi_0^{-2}(\gamma_0)\exp(-2\hat{Z}_{i,r}^T\beta_0)\{\hat{Z}_{i,r}+J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^{\otimes 2}\\
&:=B_1-B_2-B_2^T+B_3.
\end{aligned}
$$

Owing to the fact that

$$
\begin{aligned}
E[\exp(\hat{Z}_{i,r}^T\beta)|Z_i]&=\exp(Z_i^T\beta)\varphi_0(\gamma),\\
E[\exp(\hat{Z}_{i,r}^T\beta)\hat{Z}_{i,r}|Z_i]&=\exp(Z_i^T\beta)\{\varphi_0(\gamma)Z_i+J\varphi_1(\gamma)\},\\
E[\exp(\hat{Z}_{i,r}^T\beta)\hat{Z}_{i,r}^{\otimes 2}|Z_i]&=\exp(Z_i^T\beta)\big\{\varphi_0(\gamma)Z_i^{\otimes 2}+J\varphi_1(\gamma)Z_i^T\\
&\qquad +Z_i(J\varphi_1(\gamma))^T+J\varphi_2(\gamma)J^T\big\},
\end{aligned}
$$

we can obtain that

$$
\begin{aligned}
E[B_1|Y_i,Z_i]=&\varphi_0^{-2}(\gamma_0)\varepsilon_i^{-2}\Big\{\varphi_0(2\gamma_0)Z_i^{\otimes 2}\\
&+J\big\{\varphi_1(2\gamma_0)-\varphi_0^{-1}(\gamma_0)\varphi_0(2\gamma_0)\varphi_1(\gamma_0)\big\}Z_i^T\\
&+Z_i\{\varphi_1(2\gamma_0)-\varphi_0^{-1}(\gamma_0)\varphi_0(2\gamma_0)\varphi_1(\gamma_0)\}^TJ^T\\
&+J\varphi_2(2\gamma_0)J^T-\varphi_0^{-1}(\gamma_0)J\varphi_1(2\gamma_0)\varphi_1(\gamma_0)^TJ^T
\end{aligned}
$$

$$
\begin{aligned}
&- \varphi_0^{-1}(\gamma_0) J\varphi_1(\gamma_0)\varphi_1(2\gamma_0)^T J^T \\
&+ \varphi_0(2\gamma_0)\varphi_0^{-2}(\gamma_0) J\varphi_1(\gamma_0)^{\otimes 2} J^T \Big\}, \\
E[B_2|Y_i, Z_i] =& \varphi_0^{-2}(\gamma_0)\Big\{ Z_i^{\otimes 2} + J\Sigma_u J^T + Z_i\{J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^T \\
&- J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0) Z_i^T - \{J\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0)\}^{\otimes 2} \Big\}, \\
E[B_3|Y_i, Z_i] =& \varphi_0^{-2}(\gamma_0)\varepsilon_i^2 \Big\{ \varphi_0(2\gamma_0) Z_i^{\otimes 2} \\
&- J\big\{\varphi_1(2\gamma_0) - \varphi_0^{-1}(\gamma_0)\varphi_0(2\gamma_0)\varphi_1(\gamma_0)\big\} Z_i^T \\
&- Z_i\{\varphi_1(2\gamma_0) - \varphi_0^{-1}(\gamma_0)\varphi_0(2\gamma_0)\varphi_1(\gamma_0)\}^T J^T \\
&+ J\varphi_2(2\gamma_0) J^T - \varphi_0^{-1}(\gamma_0) J\varphi_1(2\gamma_0)\varphi_1(\gamma_0)^T J^T \\
&- \varphi_0^{-1}(\gamma_0) J\varphi_1(\gamma_0)\varphi_1(2\gamma_0)^T J^T \\
&+ \varphi_0(2\gamma_0)\varphi_0^{-2}(\gamma_0) J\varphi_1(\gamma_0)^{\otimes 2} J^T \Big\}.
\end{aligned}
$$

Under the assumptions that ε_i and Z_i are independent and $E(Z_i) = 0$, we have

$$
\begin{aligned}
& E\big[\{Y_i^{-1} R_{i,r}^{(1)}(\beta_0) - Y_i R_{i,r}^{(1)}(-\beta_0)\}^{\otimes 2}\big] \\
=& E\big\{E[B_1 - B_2 - B_2^T + B_3|Y_i, Z_i]\big\} \\
=& E(\varepsilon_i^{-2} + \varepsilon_i^2)\varphi_0^{-2}(\gamma_0)\Big\{ \varphi_0(2\gamma_0) E Z_i^{\otimes 2} + J\varphi_2(2\gamma_0) J^T \\
& - J\varphi_1(2\gamma_0)\varphi_1^T(\gamma_0)\varphi_0^{-1}(\gamma_0) J^T - J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\varphi_1(2\gamma_0)^T J^T \\
& + \varphi_0(2\gamma_0) J\{\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2} J^T \Big\} \\
& - 2\varphi_0^{-2}(\gamma_0)\Big\{ E Z_i^{\otimes 2} + J\Sigma_u J^T - \{J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2} \Big\}.
\end{aligned}
$$

When $r \neq s$, it follows from the independence of $U_{i,r}$ and $U_{i,s}$ given the condition Y_i and Z_i that

$$\begin{aligned}
&E\big[\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0) - Y_iR_{i,r}^{(1)}(-\beta_0)\}\{Y_i^{-1}R_{i,s}^{(1)}(\beta_0) - Y_iR_{i,s}^{(1)}(-\beta_0)\}^T\big] \\
=&E\Big\{E[Y_i^{-1}R_{i,r}^{(1)}(\beta_0) - Y_iR_{i,r}^{(1)}(-\beta_0)|Y_i, Z_i]^{\otimes 2}\Big\} \\
=&E\Big\{(\varepsilon_i - 1/\varepsilon_i)^2 Z_i^{\otimes 2}\Big\} \\
=&E(\varepsilon_i - 1/\varepsilon_i)^2 EZ_i^{\otimes 2}.
\end{aligned}$$

Summarizing the above results, we have that

$$\begin{aligned}
&E[v_i^{\otimes 2}] \\
=&E\Big\{k^{-2}\big[\sum_{r=1}^{k}\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0) - Y_iR_{i,r}^{(1)}(-\beta_0)\}^{\otimes 2} \\
&+ \sum_{r\neq s}\{Y_i^{-1}R_{i,r}^{(1)}(\beta_0) - Y_iR_{i,r}^{(1)}(-\beta_0)\}\{Y_i^{-1}R_{i,s}^{(1)}(\beta_0) - Y_iR_{i,s}^{(1)}(-\beta_0)\}^T\big]\Big\} \\
=&k^{-1}\Big\{E(\varepsilon_i^{-2} + \varepsilon_i^2)\varphi_0(2\gamma_0)\varphi_0^{-2}(\gamma_0)\big[EZ_i^{\otimes 2} + J\varphi_2(2\gamma_0)\varphi_0^{-1}(2\gamma_0)J^T \\
&- J\varphi_1(2\gamma_0)\varphi_0^{-1}(2\gamma_0)\varphi_1^T(\gamma_0)\varphi_0^{-1}(\gamma_0)J^T \\
&- J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\varphi_1^T(2\gamma_0)\varphi_0^{-1}(2\gamma_0)J^T + \{J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2}\big] \\
&- 2\varphi_0^{-2}(\gamma_0)\big[EZ_i^{\otimes 2} + J\Sigma_u J^T - \{J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2}\big]\Big\} \\
&+ (k-1)k^{-1}E(\varepsilon_i - 1/\varepsilon_i)^2 E(Z_i^{\otimes 2}).
\end{aligned}$$

Similarly, we can obtain that

$$E(\tilde{v}_i^{\otimes 2}) = E(\varepsilon_i^{-2} + \varepsilon_i^2)\varphi_0^k(2\gamma_0/k)\varphi_0^{-2k}(\gamma_0/k)\big[EZ_i^{\otimes 2}$$

$$+ k^{-1}J\varphi_2(2\gamma_0/k)\varphi_0^{-1}(2\gamma_0/k)J^T$$

$$+ (k-1)k^{-1}\{J\varphi_1(2\gamma_0/k)\varphi_0^{-1}(2\gamma_0/k)\}^{\otimes 2}$$

$$- J\varphi_1(2\gamma_0/k)\varphi_0^{-1}(2\gamma_0/k)\varphi_1^T(\gamma_0/k)\varphi_0^{-1}(\gamma_0/k)J^T$$

$$- J\varphi_1(\gamma_0/k)\varphi_0^{-1}(\gamma_0/k)\varphi_1^T(2\gamma_0/k)\varphi_0^{-1}(2\gamma_0/k)J^T$$

$$+ \{J\varphi_1(\gamma_0/k)\varphi_0^{-1}(\gamma_0/k)\}^{\otimes 2}]$$

$$- 2\varphi_0^{-2k}(\gamma_0/k)\big[EZ_i^{\otimes 2} + k^{-1}J\Sigma_u J^T - \{J\varphi_1(\gamma_0/k)\varphi_0^{-1}(\gamma_0/k)\}^{\otimes 2}\big].$$

Proof of Theorem 4.

For later convenience, we decompose $\Sigma^*_{cms} - \Sigma^*_{cee}$ as the sum of the following three terms:

$$\begin{aligned} D_1 =& \Big\{k^{-1}[E(\varepsilon^2+\varepsilon^{-2})\exp(\gamma_0^T\Sigma\gamma_0) - 2\exp(-\gamma_0^T\Sigma\gamma_0)] \\ &+ (k-1)k^{-1}[E(\varepsilon^2+\varepsilon^{-2}) - 2] - [E(\varepsilon^2+\varepsilon^{-2})\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) \\ &- 2\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\}E(Z^{\otimes 2}), \end{aligned}$$

which involves the term $E(Z^{\otimes 2})$,

$$\begin{aligned} D_2 =& \Big\{k^{-1}[E(\varepsilon^2+\varepsilon^{-2})\exp(\gamma_0^T\Sigma\gamma_0) + 2\exp(-\gamma_0^T\Sigma\gamma_0)] \\ &- k^{-2}[E(\varepsilon^2+\varepsilon^{-2})\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) + 2\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\}(J\Sigma\gamma_0)^{\otimes 2}, \end{aligned}$$

which involves the term $(J\Sigma\gamma_0)^{\otimes 2}$, and

$$D_3 = \Big\{k^{-1}[E(\varepsilon^2+\varepsilon^{-2})\exp(\gamma_0^T\Sigma\gamma_0) - 2\exp(-\gamma_0^T\Sigma\gamma_0)]$$

$$- k^{-1}[E(\varepsilon^2+\varepsilon^{-2})\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) - 2\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\} J\Sigma J^T,$$

which involves the term $J\Sigma J^T$.

We divide our proof into three steps. First, we need to verify $D_1 \geq 0$. Set $g_0(x) = k^{-1}x^k + (k-1)k^{-1} - x$. It is easy to see that $g_0'(x) = x^{k-1} - 1 \geq 0$ for any $x \geq 1$. Thanks to the fact that $g_0(1) = 0$, we obtain that $g_0(x) \geq 0$ for any $x \geq 1$. Thus, $k^{-1}\exp(\gamma_0^T\Sigma\gamma_0) + (k-1)k^{-1} - \exp(k^{-1}\gamma_0^T\Sigma\gamma_0) = g_0(\exp(k^{-1}\gamma_0^T\Sigma\gamma_0)) \geq 0$. Therefore, we have

$$\begin{aligned}
D_1 =& \Big\{ E(\varepsilon^2+\varepsilon^{-2})[k^{-1}\exp(\gamma_0^T\Sigma\gamma_0) + (k-1)k^{-1} - \exp(k^{-1}\gamma_0^T\Sigma\gamma_0)] \\
&- 2[k^{-1}\exp(-\gamma_0^T\Sigma\gamma_0) + (k-1)k^{-1} - \exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\} E(Z^{\otimes 2}) \\
\geq& \Big\{ 2[k^{-1}\exp(\gamma_0^T\Sigma\gamma_0) + (k-1)k^{-1} - \exp(k^{-1}\gamma_0^T\Sigma\gamma_0)] \\
&- 2[k^{-1}\exp(-\gamma_0^T\Sigma\gamma_0) + (k-1)k^{-1} - \exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\} E(Z^{\otimes 2}) \\
=& 2\Big\{ k^{-1}[\exp(\gamma_0^T\Sigma\gamma_0) - \exp(-\gamma_0^T\Sigma\gamma_0)] - [\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) - \exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\} \\
&\times E(Z^{\otimes 2}).
\end{aligned}$$

Take $g(x) = k^{-1}[\exp(x) - \exp(-x)] - [\exp(k^{-1}x) - \exp(-k^{-1}x)]$, $g_1(x) = \exp(x) - \exp(-x)$ and $g_2(x) = \exp(x) + \exp(-x)$. Then, $g'(x) = k^{-1}[g_2(x) - g_2(x/k)]$. We have $g_1'(x) = g_2(x) \geq 0$ and $g_2'(x) = g_1(x) \geq 0$ for any $x \geq 0$. Thus, $g_1(x)$ and $g_2(x)$ are increasing when $x > 0$. Thus, $g'(x) \geq 0$. In addition, $g(0) = 0$. Thus, $g(x) \geq 0$ for any $x \geq 0$. Recalling that $D_1 \geq 2g(\gamma_0^T\Sigma\gamma_0)E(Z^{\otimes 2})$ and $E(Z^{\otimes 2}) \geq 0$, we see that $D_1 \geq 0$.

Next, we show that $D_2 \geq 0$. Since $g_2(x)$ is an increasing function for $x > 0$, we have

$$\begin{aligned} D_2 =& k^{-1}\Big\{E(\varepsilon^2+\varepsilon^{-2})[\exp(\gamma_0^T\Sigma\gamma_0) - k^{-1}\exp(k^{-1}\gamma_0^T\Sigma\gamma_0)] \\ &+ 2\exp(-\gamma_0^T\Sigma\gamma_0) - 2k^{-1}\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)\Big\}(J\Sigma\gamma_0)^{\otimes 2} \\ \geq& k^{-1}\Big\{2\exp(\gamma_0^T\Sigma\gamma_0) - 2k^{-1}\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) \\ &+ 2\exp(-\gamma_0^T\Sigma\gamma_0) - 2k^{-1}\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)\Big\}(J\Sigma\gamma_0)^{\otimes 2} \\ =& 2k^{-1}\Big\{\exp(\gamma_0^T\Sigma\gamma_0) + \exp(-\gamma_0^T\Sigma\gamma_0) \\ &- k^{-1}[\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) + \exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\}(J\Sigma\gamma_0)^{\otimes 2} \\ \geq& 0, \end{aligned}$$

where the first inequality is due to $\exp(\gamma_0^T\Sigma\gamma_0) - k^{-1}\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) \geq 0$, and the second inequality is a consequence of $g_2(\gamma_0^T\Sigma\gamma_0) - k^{-1}g_2(k^{-1}\gamma_0^T\Sigma\gamma_0) \geq g_2(\gamma_0^T\Sigma\gamma_0) - g_2(k^{-1}\gamma_0^T\Sigma\gamma_0) \geq 0$.

Finally, we need to prove that $D_3 \geq 0$. Note that $g_1(x)$ is increasing when $x > 0$, we deduce

$$\begin{aligned} D_3 =& k^{-1}\Big\{E(\varepsilon^2+\varepsilon^{-2})[\exp(\gamma_0^T\Sigma\gamma_0) - \exp(k^{-1}\gamma_0^T\Sigma\gamma_0)] \\ &- 2[\exp(-\gamma_0^T\Sigma\gamma_0) - \exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\}J\Sigma J^T \\ \geq& k^{-1}\Big\{2\exp(\gamma_0^T\Sigma\gamma_0) - 2\exp(k^{-1}\gamma_0^T\Sigma\gamma_0) \\ &- 2\exp(-\gamma_0^T\Sigma\gamma_0) + 2\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)\Big\}J\Sigma J^T \end{aligned}$$

$$
\begin{aligned}
=&2k^{-1}\Big\{[\exp(\gamma_0^T\Sigma\gamma_0)-\exp(-\gamma_0^T\Sigma\gamma_0)]\\
&-[\exp(k^{-1}\gamma_0^T\Sigma\gamma_0)-\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)]\Big\}J\Sigma J^T\\
\geq&0.
\end{aligned}
$$

Summarizing the above results, we have $\Sigma^*_{cms}-\Sigma^*_{cee}\geq 0$.

References

Foutz, R. V. (1977). On the unique consistent solution to the likelihood equations. *Journal of the American Statistical Association*, 72(357), pp. 147–148.

Table S1. Simulation results when the assumption of model error ε is violated. LPREf, LSf, LPREnv, LSnv, CLS, CMS and CEE stand for the full LPRE, full least square(LS), naive LPRE, naive LS, classical corrected LS, proposed conditional mean score and proposed corrected estimating equation estimators. The measurement error $U \sim Uniform(-\sqrt{3}/2, \sqrt{3}/2)$ and the sample size $n = 200$. All entries are multiplied by 100.

	$\hat{\alpha}$			$\hat{\gamma}$			$\hat{c}$		
method	Bias	SE	SEE	Bias	SE	SEE	Bias	SE	SEE
				$\log \varepsilon \sim Beta(2,4)$					
LPREf	-0.02	1.50	1.45	-0.03	1.47	1.44	33.41	1.25	1.25
LSf	-0.02	1.50	1.46	-0.03	1.48	1.46	33.37	1.25	1.26
LPREnv	9.97	4.60	4.52	-19.84	4.34	4.30	33.37	4.10	4.00
LSnv	10.11	4.65	4.64	-20.13	4.36	4.46	33.33	4.13	4.05
CLS	-0.08	5.15	5.13	0.32	5.70	5.62	33.27	4.32	4.30
CMS	0.04	4.36	4.32	0.03	4.59	4.53	33.34	3.72	3.69
CEE	-0.08	5.05	4.96	0.33	5.55	5.39	33.31	4.26	4.19
				$\log \varepsilon \sim 0.5 \times t(5)$					
LPREf	0.22	11.19	6.48	0.14	10.52	6.47	0.21	11.76	6.12
LSf	0.14	5.19	5.24	0.12	5.22	5.25	-0.04	4.57	4.53
LPREnv	10.02	11.52	7.78	-19.67	10.27	7.43	0.25	13.00	7.35
LSnv	10.10	6.59	6.82	-20.00	6.42	6.55	0.03	6.13	5.96
CLS	-0.15	7.06	7.26	0.51	7.80	7.75	0.04	6.24	6.16
CMS	-0.06	11.78	7.78	0.52	11.47	8.05	0.30	12.39	7.17
CEE	-0.28	11.96	8.17	0.96	11.87	8.67	0.25	12.87	7.43

Table S2. Simulation result when the assumption of U is violated. LPREf, LSf, LPREnv, LSnv, CLS, CMS and CEE stand for the full LPRE, full least square(LS), naive LPRE, naive LS, classical corrected LS, proposed conditional mean score and proposed corrected estimating equation estimators. $\log \varepsilon \sim N(0, 0.25)$ and the sample size $n = 200$. All entries are multiplied by 100.

	$\hat{\alpha}$			$\hat{\gamma}$			$\hat{c}$		
method	Bias	SE	SEE	Bias	SE	SEE	Bias	SE	SEE
				$U \sim 0.5 \times Beta(2,4)$					
LPREf	0.02	4.19	4.05	0.04	4.12	4.05	0.06	3.60	3.53
LSf	0.03	4.18	4.09	0.03	4.10	4.09	0.06	3.58	3.53
LPREnv	10.10	6.12	5.93	-19.86	5.94	5.69	-0.62	5.44	5.26
LSnv	10.14	6.05	5.99	-19.96	5.81	5.75	-0.05	5.35	5.23
CLS	-0.09	6.53	6.45	0.50	7.04	6.98	-0.06	5.55	5.44
CMS	-0.09	6.57	6.38	0.44	7.13	6.93	-6.63	5.63	5.44
CEE	-0.15	6.63	6.37	0.61	7.20	6.85	-0.85	5.64	5.43
				$U \sim 0.5\times$ standardized χ_1^2 truncated at 5					
LPREf	-0.18	4.24	4.06	0.01	4.07	4.06	-0.01	3.68	3.53
LSf	-0.17	4.22	4.09	0.03	4.04	4.08	-0.01	3.68	3.52
LPREnv	10.14	6.50	6.08	-20.63	6.00	5.93	-2.09	5.15	5.35
LSnv	9.89	6.26	5.98	-20.08	5.71	5.74	0.02	4.95	5.22
CLS	-0.37	6.63	6.48	0.36	6.83	7.13	-0.01	5.01	5.43
CMS	-0.89	8.60	7.86	1.13	9.32	8.60	-25.30	6.86	7.19
CEE	-0.56	7.05	6.59	0.66	7.54	7.22	-2.93	5.33	5.59
				$U \sim 0.5\times$ standardized $t(5)$					
LPREf	0.01	4.22	4.06	0.20	4.23	4.04	-0.20	3.63	3.52
LSf	0.01	4.20	4.09	0.20	4.21	4.08	-0.19	3.62	3.52
LPREnv	10.79	7.13	6.23	-21.10	7.76	6.17	0.00	5.87	5.46
LSnv	10.15	6.30	6.00	-19.88	6.19	5.75	-0.05	5.32	5.24
CLS	-0.10	6.63	6.56	0.60	7.13	7.27	-0.02	5.47	5.45
CMS	0.54	29.96	11.84	0.27	40.41	13.45	0.32	31.44	11.29
CEE	-0.25	7.52	6.84	0.99	8.67	7.60	0.03	6.16	5.75