Supplemental Material to "Identifiability of Latent Class Models with Covariates"

This supplementary material contains two sections. Section A provides the proofs of propositions and theorems from Section 3 and Section 4 of the main article. Section B gives the proofs of lemmas introduced in Section A.

A Proofs of Propositions and Theorems

In this section, we first introduce a lemma adapted from Proposition 3 in Huang and Bandeen-Roche (2004), which is an important tool in later proofs to associate the identifiability of parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ with the identifiability of $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i) = \{\eta_c^i, \theta_{jrc}^i : j = 1, \ldots, J, r = 0, \ldots, M_j - 1, c = 0, \ldots, C-1\}$, for $i = 1, \ldots, N$.

Lemma 1. *For any subject $i = 1, \ldots, N$, we define transformed variables $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i) = \{\epsilon_c, \omega_{jrc} : j = 1, \ldots, J, r = 0, \ldots, M_j - 1, c = 0, \ldots, C-1\}$ such that $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ and $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i)$ are related through the following equations,*

$$\begin{aligned} \eta_c^i = \frac{\exp(\epsilon_c^i)}{1 + \sum_{s=1}^{C-1} \exp(\epsilon_s^i)}, \quad & c = 0, \ldots, C-1; \\ \theta_{jrc}^i = \frac{\exp(\omega_{jrc}^i)}{1 + \sum_{s=1}^{M_j-1} \exp(\omega_{jsc}^i)}, \quad & j = 1, \ldots, J; \\ & r = 0, \ldots, M_j - 1; \\ & c = 0, \ldots, C-1. \end{aligned}$$

Then $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are identifiable if and only if $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i)$ are identifiable.

The proof of Lemma 1 is presented in Section B.

Proof of Proposition 1. We first prove the second part of Proposition 1 that $(A3^*)$ is necessary for the identifiability of RegLCMs without covariates under $(A1)$ and $(A2^*)$. It is equivalent to show that if $\boldsymbol{\psi_0}, \cdots, \boldsymbol{\psi_{C-1}}$ are not linearly independent, $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are not identifiable. We prove it by the method of contradiction and assume the contrary that $\boldsymbol{\eta}$ are identifiable.

Recall the definitions in Section 2, $\boldsymbol{\eta} = (\eta_0, \cdots, \eta_{C-1})^T$ denotes the latent class membership probability, where $\eta_c = P(L = c)$ for $c = 0, \cdots, C-1$. And $\Psi = (\boldsymbol{\psi}_0, \cdots, \boldsymbol{\psi}_{C-1})$ denotes the marginal probability matrix, where each entry $\psi_{\boldsymbol{r}c}$ in $\boldsymbol{\psi}_c$ corresponding to a response pattern $\boldsymbol{r} \in \mathcal{S}'$ is

$$\psi_{\boldsymbol{r}c} = P(\boldsymbol{R} = \boldsymbol{r} \mid L = c) = \prod_{j=1}^{J} \theta_{jr_jc}, \quad c = 0, \cdots, C-1.$$

Based on the above definitions, we write the response probability vector as

$$\left(P(\boldsymbol{R} = \boldsymbol{r}) : \boldsymbol{r} \in \mathcal{S}'\right)^T = \Psi \cdot \boldsymbol{\eta}. \tag{1}$$

As we assume $\boldsymbol{\eta}$ are identifiable, there exist no $\boldsymbol{\eta'} \neq \boldsymbol{\eta}$ such that $P(\boldsymbol{R} = \boldsymbol{r} \mid \Psi, \boldsymbol{\eta}) = P(\boldsymbol{R} = \boldsymbol{r} \mid \Psi, \boldsymbol{\eta'})$. According to (1), $P(\boldsymbol{R} = \boldsymbol{r} \mid \Psi, \boldsymbol{\eta}) = P(\boldsymbol{R} = \boldsymbol{r} \mid \Psi, \boldsymbol{\eta'})$ implies $\Psi \cdot \boldsymbol{\eta} = \Psi \cdot \boldsymbol{\eta'}$. However, under the condition that $\boldsymbol{\psi_0}, \cdots, \boldsymbol{\psi_{C-1}}$ are not linearly independent, there could exist $\boldsymbol{\eta'} \neq \boldsymbol{\eta}$ such that $\Psi \cdot (\boldsymbol{\eta} - \boldsymbol{\eta'}) = \mathbf{0}$, and by the contradiction, $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are not identifiable.

Next, we prove the first part of Proposition 1, the necessity of $(A4)$ for the identifiability of RegLCMs under $(A1)$–$(A3)$. That is, if $\boldsymbol{\phi_0}, \cdots, \boldsymbol{\phi_{C-1}}$ are not linearly independent, then $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are not identifiable. This proof includes three steps.

Step 1: we prove if $\boldsymbol{\phi_0}, \cdots, \boldsymbol{\phi_{C-1}}$ are not linearly independent, then $\boldsymbol{\psi^i_0}, \cdots, \boldsymbol{\psi^i_{C-1}}$ are not linearly independent for $i = 1, \ldots, N$, where each $\boldsymbol{\psi^i_c}$ is an $(S-1)$-dimensional vector in which each element corresponds to a response pattern $\boldsymbol{r} = (r_1, \cdots, r_J) \in \mathcal{S}'$ and is defined as $\boldsymbol{\psi}^i_{\boldsymbol{r}c} = P(\boldsymbol{R}_i = \boldsymbol{r} \mid L_i = c, \boldsymbol{x}_i, \boldsymbol{z}_i)$. Equivalently, we need to prove if there exists subject i such that $\boldsymbol{\psi^i_0}, \cdots, \boldsymbol{\psi^i_{C-1}}$ are linearly independent, then $\boldsymbol{\phi_0}, \cdots, \boldsymbol{\phi_{C-1}}$ are linearly independent. We use similar techniques as in the *Proof of Proposition 2* in Huang and Bandeen-Roche (2004). First, we associate the linear combinations of $\boldsymbol{\phi}_c$'s with $\boldsymbol{\psi}_c$'s as follows. For any linear combination of $\boldsymbol{\phi}_c$'s with coefficients a_c's, there exist b_c's and $\boldsymbol{Y}^i$ such that the following equation holds,

$$\sum_{c=0}^{C-1} a_c \boldsymbol{\phi}_c = \left(\sum_{c=0}^{C-1} b_c \boldsymbol{\psi}^i_c\right) \odot \boldsymbol{Y}^i, \tag{2}$$

where $\odot$ denotes the element-wise multiplication and

$$\begin{aligned}\boldsymbol{Y}^i &= \left(\prod_{j=1}^J \frac{1}{\exp(\lambda_{1jr_j}z_{ij1}+\cdots+\lambda_{qjr_j}z_{ijq})} : \boldsymbol{r}=(r_1,\ldots,r_J)\in\mathcal{S}'\right)^T_{S\times 1}, \\ b_c &= a_c\prod_{j=1}^J \frac{1+\sum_{s=1}^{M_j-1}\exp(\gamma_{jsc}+\lambda_{1js}z_{ij1}+\cdots+\lambda_{qjs}z_{ijq})}{1+\sum_{s=1}^{M_j-1}e^{\gamma_{jsc}}}. \end{aligned} \quad (3)$$

Therefore to show $\boldsymbol{\phi}_c$'s are linearly independent, we need to show that $\sum_{c=0}^{C-1} a_c\boldsymbol{\phi}_c = 0$ implies $a_0 = \cdots = a_{C-1} = 0$. Based on (2), we have $\sum_{c=0}^{C-1} a_c\boldsymbol{\phi}_c = 0$ implies $\sum_{c=0}^{C-1} b_c\boldsymbol{\psi}_c^i = 0$. Under the condition that $\boldsymbol{\psi}_0^i, \ldots, \boldsymbol{\psi}_{C-1}^i$ are linearly independent, the equation

$$\sum_{c=0}^{C-1} b_c\boldsymbol{\psi}_c^i = b_0\boldsymbol{\psi}_0^i + \cdots + b_{C-1}\boldsymbol{\psi}_{C-1}^i = 0 \quad (4)$$

implies $b_0 = \cdots = b_{C-1} = 0$. And by (3), we have $a_0 = \cdots = a_{C-1} = 0$. Hence, $\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_{C-1}$ are linearly independent when $\boldsymbol{\psi}_0^i, \ldots, \boldsymbol{\psi}_{C-1}^i$ are linearly independent and we complete the proof for *Step 1*.

Step 2: We next introduce parameters ϵ_c^i's and ω_{jrc}^i's and show that they are not identifiable when $\boldsymbol{\psi}^i_{\mathbf{0}}, \cdots, \boldsymbol{\psi}^i_{\boldsymbol{C-1}}$ are not linearly independent. By the similar arguments in proving the necessity of $(A3^*)$, we have $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are not identifiable when $\boldsymbol{\psi}^i_{\mathbf{0}}, \cdots, \boldsymbol{\psi}^i_{\boldsymbol{C-1}}$ are not linearly independent for any subject $i = 1, \ldots, N$. Recall in RegLCMs, $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are functionally dependent on the linear functions $\boldsymbol{x}_i^T\boldsymbol{\beta}$ and $\boldsymbol{\gamma}_{jc} + \boldsymbol{z}_{ij}^T\boldsymbol{\lambda}_j$, respectively. We follow the definitions of $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ and $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ from (3) and (4) in Section 2.1 and let

$$\epsilon_c^i = \boldsymbol{x_i}^T\boldsymbol{\beta}_c = \beta_{0c} + \beta_{1c}x_{i1} + \cdots + \beta_{pc}x_{ip}.$$

for $i = 1, \ldots, N$, $c = 0, \ldots, C-1$. And

$$\omega_{jrc}^i = \gamma_{jrc} + \boldsymbol{z}_{ij}^T\boldsymbol{\lambda}_{jr} = \gamma_{jrc} + \lambda_{1jr}z_{ij1} + \cdots + \lambda_{qjr}z_{ijq}.$$

for $i = 1, \ldots, N$, $j = 1, \ldots, J$, $r = 0, \cdots, M_j - 1$ and $c = 0, \ldots, C-1$. Then according to Lemma 1, ϵ_c^i's and ω_{jrc}^i's are not identifiable when $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are not identifiable. Hence, ϵ_c^i's and ω_{jrc}^i's are not identifiable when $\boldsymbol{\psi}^i_{\mathbf{0}}, \cdots, \boldsymbol{\psi}^i_{\boldsymbol{C-1}}$ are not linearly independent and we complete the proof for *Step 2*.

Step 3: Lastly, we prove that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are not identifiable when ϵ_c^i's and ω_{jrc}^i's are not identifiable by the method of contradiction. Assume to the contrary that $\boldsymbol{\beta}$ are identifiable given ϵ_c^i's and ω_{jrc}^i's are not identifiable. By the definition of identifiability, $P(\boldsymbol{R} \mid \boldsymbol{\beta}^*, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = P(\boldsymbol{R} \mid \boldsymbol{\beta}', \boldsymbol{\gamma}, \boldsymbol{\lambda})$ implies that $\boldsymbol{\beta}^* = \boldsymbol{\beta}'$. Because $\boldsymbol{X}$ has full column rank, according to the system of linear equations

$$\boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}^1 \\ \vdots \\ \boldsymbol{\epsilon}^N \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \cdots & x_{Np} \end{pmatrix} \begin{pmatrix} \beta_{00} & \cdots & \beta_{0(C-1)} \\ \vdots & \ddots & \vdots \\ \beta_{p0} & \cdots & \beta_{p(C-1)} \end{pmatrix} = \boldsymbol{X}\boldsymbol{\beta},$$

we have $\boldsymbol{\epsilon}^* = \boldsymbol{X}\boldsymbol{\beta}^*$ equivalent to $\boldsymbol{\epsilon}' = \boldsymbol{X}\boldsymbol{\beta}'$. So for all subject i, $P(\boldsymbol{R}_i \mid \boldsymbol{\epsilon}^{i*}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = P(\boldsymbol{R}_i \mid \boldsymbol{\epsilon}^{i\prime}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ would force $\boldsymbol{\epsilon}^{i*} = \boldsymbol{\epsilon}^{i\prime}$, which contradicts the non-identifiability of ϵ_c^i's. Therefore $\boldsymbol{\beta}$ is not identifiable. Using similar techniques, we can prove $\boldsymbol{\gamma}, \boldsymbol{\lambda}$ are not identifiable.

Combining the *Step 1–3*, we prove the first part of Proposition 1, and thus complete the proof of Proposition 1. □

Proof of Proposition 2. Theorem 4.4 (a) in Gu and Xu (2020) showed that binary-response CDMs are not generically identifiable if some attribute is required by only one item. We adapt their proof of Theorem 4.4 (a) to establish that for polytomous-response CDMs or RegCDMs, the parameters are not generically identifiable under $(P1)$ that some attribute is required by only one item. Consider polytomous-response CDMs first and let the Q-matrix to be

$$Q = \begin{pmatrix} 1 & \boldsymbol{u} \\ \boldsymbol{0} & Q^* \end{pmatrix}.$$

This Q-matrix implies that α_1 is required by the first item only. For any $(\boldsymbol{\eta}, \boldsymbol{\Theta})$, we can construct $(\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\Theta}}) \neq (\boldsymbol{\eta}, \boldsymbol{\Theta})$ such that $P(\boldsymbol{R} = \boldsymbol{r} \mid \boldsymbol{\eta}, \boldsymbol{\Theta}) = P(\boldsymbol{R} = \boldsymbol{r} \mid \bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\Theta}})$, and hence we show that $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are not identifiable. To better illustrate the idea, we next use $\boldsymbol{\alpha}$ to replace c in all parameter subscripts, i.e. $\eta_{\boldsymbol{\alpha}} = \eta_c$ and $\theta_{jr\boldsymbol{\alpha}} = \theta_{jrc}$ given $\boldsymbol{\alpha}^T \boldsymbol{v} = c$. When $j \neq 1$, we let $\eta_c = \bar{\eta}_c$, $\theta_{jrc} = \bar{\theta}_{jrc}$ for $r = 0, \ldots, M_j - 1$ and $c = 0, \ldots, C-1$. When $j = 1$, we denote $\boldsymbol{\alpha}' = (\alpha_2, \cdots, \alpha_K) \in \{0,1\}^{K-1}$ and for all $r_1 = 0, \ldots, M_1 - 1$, we let $\bar{\theta}_{1r_1(0,\boldsymbol{\alpha}')} = \theta_{1r_1(0,\boldsymbol{\alpha}')}$, and

$$\bar{\theta}_{1r_1(1,\boldsymbol{\alpha}')} = \frac{1}{E}\theta_{1r_1(1,\boldsymbol{\alpha}')} + (1-\frac{1}{E})\theta_{1r_1(0,\boldsymbol{\alpha}')},$$

where E is a constant in a small neighborhood of 1 and $E \neq 1$. So we have $\bar{\theta}_{1r_1(1,\boldsymbol{\alpha}')} \neq \theta_{1r_1(1,\boldsymbol{\alpha}')}$. We also let

$$\begin{aligned} \bar{\eta}_{(0,\boldsymbol{\alpha}')} &= \eta_{(0,\boldsymbol{\alpha}')} + (1-E)\cdot\eta_{(1,\boldsymbol{\alpha}')}, \\ \bar{\eta}_{(1,\boldsymbol{\alpha}')} &= E\cdot\eta_{(1,\boldsymbol{\alpha}')}. \end{aligned}$$

Hence, we have

$$\bar{\eta}_{(1,\boldsymbol{\alpha}')} + \bar{\eta}_{(0,\boldsymbol{\alpha}')} = \eta_{(1,\boldsymbol{\alpha}')} + \eta_{(0,\boldsymbol{\alpha}')}, \tag{5}$$

$$\bar{\theta}_{1r_1(1,\boldsymbol{\alpha}')}\bar{\eta}_{(1,\boldsymbol{\alpha}')} + \bar{\theta}_{1r_1(0,\boldsymbol{\alpha}')}\bar{\eta}_{(0,\boldsymbol{\alpha}')} = \theta_{1r_1(1,\boldsymbol{\alpha}')}\eta_{(1,\boldsymbol{\alpha}')} + \theta_{1r_1(0,\boldsymbol{\alpha}')}\eta_{(0,\boldsymbol{\alpha}')}. \tag{6}$$

So for any $\boldsymbol{r} = (r_1, \cdots, r_J) \in \mathcal{S}'$,

$$\begin{aligned}
& P(\boldsymbol{R} = \boldsymbol{r} \mid \bar{\Psi}, \bar{\boldsymbol{\eta}}) = \bar{\Psi} \boldsymbol{\cdot} \bar{\boldsymbol{\eta}} \\
= & \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \bar{\theta}_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \bar{\eta}_{(\alpha_1,\boldsymbol{\alpha}')} [\bar{\theta}_{1r_1(1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_1=r_1\}} \bar{\eta}_{(1,\boldsymbol{\alpha}')} + \bar{\theta}_{jr_1(0,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_1=r_1\}} \bar{\eta}_{(0,\boldsymbol{\alpha}')}] \\
= & \begin{cases} \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \bar{\theta}_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \bar{\eta}_{(\alpha_1,\boldsymbol{\alpha}')} [\bar{\theta}_{1r_1(1,\boldsymbol{\alpha}')} \bar{\eta}_{(1,\boldsymbol{\alpha}')} + \bar{\theta}_{1r_1(0,\boldsymbol{\alpha}')} \bar{\eta}_{(0,\boldsymbol{\alpha}')}], & R_1 = r_1 \\ \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \bar{\theta}_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \bar{\eta}_{(\alpha_1,\boldsymbol{\alpha}')} [\bar{\eta}_{(1,\boldsymbol{\alpha}')} + \bar{\eta}_{(0,\boldsymbol{\alpha}')}], & R_1 \neq r_1 \end{cases} \\
= & \begin{cases} \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \theta_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \eta_{(\alpha_1,\boldsymbol{\alpha}')} [\theta_{1r_1(1,\boldsymbol{\alpha}')} \eta_{(1,\boldsymbol{\alpha}')} + \theta_{1r_1(0,\boldsymbol{\alpha}')} \eta_{(0,\boldsymbol{\alpha}')}], & R_1 = r_1 \\ \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \theta_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \eta_{(\alpha_1,\boldsymbol{\alpha}')} [\eta_{(1,\boldsymbol{\alpha}')} + \eta_{(0,\boldsymbol{\alpha}')}], & R_1 \neq r_1 \end{cases} \qquad (7) \\
= & \sum_{\substack{\boldsymbol{\alpha}' \in \{0,1\}^{K-1} \\ \alpha_1 \in \{0,1\}}} \prod_{j>1} \theta_{jr_j(\alpha_1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_j = r_j\}} \eta_{(\alpha_1,\boldsymbol{\alpha}')} [\theta_{1r_1(1,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_1=r_1\}} \eta_{(1,\boldsymbol{\alpha}')} + \theta_{1r_1(0,\boldsymbol{\alpha}')}^{\mathbb{I}\{R_1=r_1\}} \eta_{(0,\boldsymbol{\alpha}')}] \\
= & \Psi \boldsymbol{\cdot} \boldsymbol{\eta} = P(\boldsymbol{R} = \boldsymbol{r} \mid \Psi, \boldsymbol{\eta}).
\end{aligned}$$

Equation (7) is derived based on (6) as well as the assumption that $\eta_c = \bar{\eta}_c$, $\theta_{jrc} = \bar{\theta}_{jrc}$ for all $j = 2, \ldots, J$, $r = 0, \ldots, M_j - 1$ and $c = 0, \ldots, C-1$. This proves $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are not identifiable under $(P1)$ in Proposition 2.

For polytomous-response RegCDMs, we have similar results by following the above proof. That is, $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are not identifiable under $(P1)$ for $i = 1, \ldots, N$. Then following the same arguments in *Step 2–3* from the *Proof of Proposition 1*, we show that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegCDMs are not identifiable given $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are not identifiable.

Next we prove the remaining part, that is, the matrix $\boldsymbol{\Psi}$ in CDMs and the matrix $\boldsymbol{\Phi}$ in RegCDMs have full column ranks under $(P2)$. Before presenting the proof, we introduce another probability matrix T-matrix of size $S \times C$, where each row corresponds to one response pattern $\boldsymbol{r} \in \mathcal{S}$ and each column corresponds to one latent class $c = 0, \ldots, C-1$. Each entry of T-matrix is defined as $T_{\boldsymbol{r}c} = P(\boldsymbol{R} \succeq \boldsymbol{r} \mid L = c)$, where $\succeq$ denotes that for any item $j = 1, \ldots, J$, $R_j \geq r_j$. According to a similar argument in Appendix Section 4.2 in Xu (2017), T-matrix has full column rank under the condition that the corresponding Q-matrix contains an identity submatrix $\mathcal{I}_K$.

There exists a relation between the two probability matrices, T-matrix and $\boldsymbol{\Psi}$. Because $\boldsymbol{\Psi}$ excludes a reference response pattern, its size is $(S-1) \times C$. Denote $\boldsymbol{\Psi'} = (\boldsymbol{\Psi}^T, \boldsymbol{\Psi}_{ref}^T)^T$ where $\boldsymbol{\Psi}_{ref}$ is the row corresponding to the reference pattern. And $\boldsymbol{\Psi}_{ref}$ is linearly dependent on the rows in $\boldsymbol{\Psi}$ because $\sum_{\boldsymbol{r} \in \mathcal{S}} P(\boldsymbol{R} = \boldsymbol{r} \mid L = c) = 1$. So $\boldsymbol{\Psi}$ has full column rank if and only if $\boldsymbol{\Psi'}$ has full column rank. Further, $\boldsymbol{\Psi'}$ has full column rank if and only if T-matrix has full column rank, because $\boldsymbol{\Psi'}$ is bijectively corresponding to T-matrix according to their definitions. In conclusion, $\boldsymbol{\Psi}$ in the CDMs has full column rank when Q-matrix contains an identity submatrix $\mathcal{I}_K$. According to the *Proof of Proposition 2* in Huang and Bandeen-Roche (2004), the matrix $\boldsymbol{\Phi}$ has full column rank when the matrix $\boldsymbol{\Psi}$ has full column rank. So for RegCDMs, $\boldsymbol{\Phi}$ has full column rank when Q-matrix contains an identity submatrix $\mathcal{I}_K$. □

Proof of Theorem 1. Following the similar idea in Huang and Bandeen-Roche (2004) page 15, we let $f(\boldsymbol{R}; \boldsymbol{\eta}, \boldsymbol{\Theta})$ to be the likelihood function, and

$$f(\boldsymbol{R}; \boldsymbol{\eta}, \boldsymbol{\Theta}) = \prod_{\boldsymbol{r} \in \mathcal{S}} P(\boldsymbol{R} = \boldsymbol{r})^{\mathbb{I}\{\boldsymbol{R} = \boldsymbol{r}\}},$$

where $\mathcal{S} = \bigtimes_{j=1}^{J} \{0, \ldots, M_j - 1\}$. Let $\boldsymbol{\xi} = \{\eta_1, \cdots, \eta_{C-1}, \theta_{110}, \cdots, \theta_{1(M_1-1)0}, \cdots, \theta_{J1(C-1)}, \cdots,$

$\theta_{J(M_J-1)(C-1)}\}$, so the Fisher information matrix is

$$
\begin{aligned}
& \mathbb{E}\left[\left(\frac{\partial \log f}{\partial \boldsymbol{\xi}}\right)\left(\frac{\partial \log f}{\partial \boldsymbol{\xi}}\right)^T\right] \\
= & \mathbb{E}\left[\left(\sum_{\boldsymbol{r}\in\mathcal{S}} \frac{\mathbb{I}\{\boldsymbol{R}=\boldsymbol{r}\}}{P(\boldsymbol{R}=\boldsymbol{r})}\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \boldsymbol{\xi}}\right)\left(\sum_{\boldsymbol{r}\in\mathcal{S}} \frac{\mathbb{I}\{\boldsymbol{R}=\boldsymbol{r}\}}{P(\boldsymbol{R}=\boldsymbol{r})}\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \boldsymbol{\xi}}\right)^T\right] \\
= & \sum_{\boldsymbol{r}\in\mathcal{S}} \frac{1}{P(\boldsymbol{R}=\boldsymbol{r})}\left(\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \boldsymbol{\xi}}\right)\left(\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \boldsymbol{\xi}}\right)^T \\
= & \mathbf{J}^T \begin{pmatrix} \frac{1}{P(\boldsymbol{R}=\boldsymbol{r}_1)} & 0 & \cdots & 0 \\ 0 & \frac{1}{P(\boldsymbol{R}=\boldsymbol{r}_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{P(\boldsymbol{R}=\boldsymbol{r}_S)} \end{pmatrix} \mathbf{J}.
\end{aligned}
$$

Hence the Fisher information matrix is non-singular if and only if $\mathbf{J}$ has full column rank. According to Theorem 1 of Rothenberg (1971), $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are locally identifiable if and only if the Fisher information matrix is non-singular when the true values of $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are regular point of the information matrix. Therefore $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are locally identifiable if and only if the Jacobian matrix $\mathbf{J}$ has full column rank. □

Proof of Theorem 2. As introduced in Section 4, we consider a hypothetical subject with all covariates being zeros and denote its Jacobian matrix as $\mathbf{J}^0$. We use the following three steps to prove that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are identifiable if and only if $\mathbf{J}^0$ has full column rank.

Step 1: We first show that for subject $i = 1, \ldots, N$, the Jacobian matrices $\mathbf{J}^i$, containing the derivatives of conditional response probabilities with respect to parameters in $\boldsymbol{\eta}^i$ and $\boldsymbol{\Theta}^i$, have full column rank if and only if $\mathbf{J}^0$ has full column rank. This proof is adapted from the *Proof of Proposition 1* in Huang and Bandeen-Roche (2004).

First, we need to set up a few notations. The Jacobian matrix $\mathbf{J}^i$ is written as

$$
\mathbf{J}^i = \left(\boldsymbol{J}^i_{\eta_1}, \cdots, \boldsymbol{J}^i_{\eta_{C-1}}, \boldsymbol{J}^i_{\theta_{110}}, \cdots, \boldsymbol{J}^i_{\theta_{1(M_1-1)0}}, \cdots, \boldsymbol{J}^i_{\theta_{J1(C-1)}}, \cdots, \boldsymbol{J}^i_{\theta_{J(M_J-1)(C-1)}}\right).
$$

Each entry in $\boldsymbol{J}^i_{\eta_c}$ is a partial derivative of response probability $P(\boldsymbol{R} = \boldsymbol{r})$ with respect to η^i_c

at true value of η^i_c, which is computed to be

$$\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \eta^i_c} = \prod_{j=1}^{J} \theta^i_{jr_jc} - \prod_{j=1}^{J} \theta^i_{jr_j0} = \boldsymbol{\psi}^i_{\boldsymbol{r}c} - \boldsymbol{\psi}^i_{\boldsymbol{r}0}.$$

And each entry in $\boldsymbol{J}^i_{\theta_{jrc}}$ is a partial derivative of response probability $P(\boldsymbol{R}=\boldsymbol{r})$ with respect to θ^i_{jrc} at true value of θ^i_{jrc}, which is computed to be

$$\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \theta^i_{jrc}} = \begin{cases} \eta^i_c \prod_{d\neq j} \theta^i_{dr_dc}, & \text{if } r_j = r, \\ -\eta^i_c \prod_{d\neq j} \theta^i_{dr_dc}, & \text{if } r_j = 0, \\ 0, & \text{otherwise.} \end{cases}$$

or summarized as

$$\frac{\partial P(\boldsymbol{R}=\boldsymbol{r})}{\partial \theta^i_{jrc}} = \eta^i_c \boldsymbol{\psi}^i_{\boldsymbol{r}c}\left(\frac{\mathbb{I}\{r_j = r\}}{\theta^i_{jrc}} - \frac{\mathbb{I}\{r_j = 0\}}{\theta^i_{j0c}}\right).$$

In addition to $\mathbf{J}^i$, we also define the following two sets of vectors for this proof. Denote $\overline{\mathbf{J}}^0 = \{\boldsymbol{\psi}^0_0, \ldots, \boldsymbol{\psi}^0_{C-1}\} \cup \{\eta^0_c \left(\mathbf{I}\{r_j = r\}/\theta^0_{jrc}\right) \odot \boldsymbol{\psi}^0_c : j = 1, \ldots, J, r = 0, \ldots, M_j - 1, c = 0, \ldots, C-1\}$ and $\overline{\mathbf{J}}^i = \{\boldsymbol{\psi}^i_0, \ldots, \boldsymbol{\psi}^i_{C-1}\} \cup \{\eta^i_c \left(\mathbf{I}\{r_j = r\}/\theta^i_{jrc}\right) \odot \boldsymbol{\psi}^i_c : j = 1, \ldots, J, r = 0, \ldots, M_j - 1, c = 0, \ldots, C-1\}$, where $\mathbf{I}\{r_j = r\}$ is a $(S-1)$-dimensional vector containing all $\mathbb{I}\{r_j = r\}$ for $\boldsymbol{r} = (r_1, \ldots, r_J) \in \mathcal{S}'$. With the notations defined, we then introduce a useful lemma which simplify the arguments in proving the linear independence of the columns in $\mathbf{J}^0$ and $\mathbf{J}^i$.

Lemma 2. *The Jacobian matrix* $\mathbf{J}^0$ *has full column rank if and only if* $\overline{\mathbf{J}}^0$ *are linearly independent. The Jacobian matrix* $\mathbf{J}^i$ *has full column rank if and only if* $\overline{\mathbf{J}}^i$ *are linearly independent.*

The proof of Lemma 2 is presented in Section B. According to Lemma 2, to prove $\mathbf{J}^i$ has full column rank if and only if $\mathbf{J}^0$ has full column rank, we can equivalently show that $\overline{\mathbf{J}}^i$ are linearly independent if and only if $\overline{\mathbf{J}}^0$ are linearly independent. First, we associate the linear combinations of $\overline{\mathbf{J}}^i$ to that of $\overline{\mathbf{J}}^0$ as follows. For any linear combinations of $\overline{\mathbf{J}}^i$ with coefficients t_c^i, u_{jrc}^i, there exist t_c^0, u_{jrc}^0 and $\boldsymbol{W}^i$ such that the following equation holds

$$\begin{aligned}&\sum_{c=0}^{C-1} t_c^i \boldsymbol{\psi}_c^i + \sum_{j=1}^{J}\sum_{r=0}^{M_j-1}\sum_{c=0}^{C-1}\left\{u_{jrc}^i \eta_c^i \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^i}\right) \odot \boldsymbol{\psi}_c^i\right\} \\ &= \left(\sum_{c=0}^{C-1} t_c^0 \boldsymbol{\psi}_c^0 + \sum_{j=1}^{J}\sum_{r=0}^{M_j-1}\sum_{c=0}^{C-1}\left\{u_{jrc}^0 \eta_c^0 \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^0}\right) \odot \boldsymbol{\psi}_c^0\right\}\right) \odot \boldsymbol{W}^i, \qquad (8)\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{W}^i &= \left(\textstyle\prod_{j=1}^{J} \exp(\lambda_{1jr_j} z_{ij1} + \cdots + \lambda_{qjr_j} z_{ijq}) : \boldsymbol{r} = (r_1, \ldots, r_J) \in \mathcal{S}'\right)_{S\times 1}^{T}, \\ t_c^0 &= t_c^i \prod_{j=1}^{J} \frac{1 + \sum\limits_{s=1}^{M_j-1} e^{\gamma_{jsc}}}{1 + \sum\limits_{s=1}^{M_j-1} \exp(\gamma_{jsc} + \lambda_{1js} z_{ij1} + \cdots + \lambda_{qjs} z_{ijq})}, \qquad (9)\\ u_{jrc}^0 &= u_{jrc}^i \frac{\exp(\beta_{1c} x_{i1} + \cdots + \beta_{pc} x_{ip})}{\exp(\lambda_{1jr_j} z_{ij1} + \cdots + \lambda_{qjr_j} z_{ijq})} \\ &\quad \times \frac{\{1 + \sum\limits_{l=1}^{C-1} e^{\beta_{0l}}\}\{1 + \sum\limits_{s=1}^{M_j-1} \exp(\gamma_{jsc} + \lambda_{1js} z_{ij1} + \cdots + \lambda_{qjs} z_{ijq})\}}{\{1 + \sum\limits_{l=1}^{C-1} \exp(\beta_{0l} + \beta_{1l} x_{i1} + \cdots + \beta_{pl} x_{ip})\}\{1 + \sum\limits_{s=1}^{M_j-1} e^{\gamma_{jsc}}\}} \\ &\quad \times \prod_{j=1}^{J} \frac{1 + \sum\limits_{s=1}^{M_j-1} e^{\gamma_{jsc}}}{1 + \sum\limits_{s=1}^{M_j-1} \exp(\gamma_{jsc} + \lambda_{1js} z_{ij1} + \cdots + \lambda_{qjs} z_{ijq})}. \qquad (10)\end{aligned}$$

The next two parts prove that $\overline{\mathbf{J}}^i$ are linearly independent if and only if $\overline{\mathbf{J}}^0$ are linearly independent in two directions.

Part (i): We prove $\overline{\mathbf{J}}^i$ are linearly independent if $\overline{\mathbf{J}}^0$ are linearly independent. To show $\overline{\mathbf{J}}^i$ are linearly independent, we need to show that

$$\sum_{c=0}^{C-1} t_c^i \boldsymbol{\psi}_c^i + \sum_{j=1}^{J}\sum_{r=0}^{M_j-1}\sum_{c=0}^{C-1}\left\{u_{jrc}^i \eta_c^i \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^i}\right) \odot \boldsymbol{\psi}_c^i\right\} = 0, \qquad (11)$$

implies $t_c^i = 0$ and $u_{jrc}^i = 0$. By (8), for any t_c^i, u_{jrc}^i such that (11) holds, we have

$$\sum_{c=0}^{C-1} t_c^0 \boldsymbol{\psi}_c^0 + \sum_{j=1}^{J}\sum_{r=0}^{M_j-1}\sum_{c=0}^{C-1}\left\{u_{jrc}^0 \eta_c^0 (\frac{\mathbb{I}\{\boldsymbol{r}_j = r\}}{\theta_{jrc}^0}) \odot \boldsymbol{\psi}_c^0\right\} = 0.$$

Under the condition that $\overline{\mathbf{J}}^0$ are linearly independent, $t_c^0 = 0$ and $u_{jrc}^0 = 0$. Then by (9) and (10), we have $t_c^i = 0$ and $u_{jrc}^i = 0$ for $j = 1, \ldots, J$, $r = 0, \ldots, M_j - 1$ and $c = 0, \ldots, C-1$. So $\overline{\mathbf{J}}^i$ are linearly independent.

Part (ii): We prove $\overline{\mathbf{J}}^0$ are linearly independent if $\overline{\mathbf{J}}^i$ are linearly independent. This part is similar to *Part (i)*. To show $\overline{\mathbf{J}}^0$ are linearly independent, we need to show that

$$\sum_{c=0}^{C-1} t_c^0 \boldsymbol{\psi}_c^0 + \sum_{j=1}^{J} \sum_{r=0}^{M_j-1} \sum_{c=0}^{C-1} \left\{ u_{jrc}^0 \eta_c^0 \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^0} \right) \odot \boldsymbol{\psi}_c^0 \right\} = 0 \tag{12}$$

implies $t_c^0 = 0$ and $u_{jrc}^0 = 0$. By (8), for any t_c^0, u_{jrc}^0 such that (12) holds, we have

$$\sum_{c=0}^{C-1} t_c^i \boldsymbol{\psi}_c^i + \sum_{j=1}^{J} \sum_{r=0}^{M_j-1} \sum_{c=0}^{C-1} \left\{ u_{jrc}^i \eta_c^i \left(\frac{\mathbb{I}\{\boldsymbol{r}_j = r\}}{\theta_{jrc}^i} \right) \odot \boldsymbol{\psi}_c^i \right\} = 0.$$

Under the condition that $\overline{\mathbf{J}}^i$ are linearly independent, $t_c^i = 0$ and $u_{jrc}^i = 0$, and hence $t_c^0 = 0$ and $u_{jrc}^i = 0$ by (9) and (10), for $j = 1, \ldots, J$, $r = 0, \ldots, M_j - 1$ and $c = 0, \ldots, C-1$. So $\overline{\mathbf{J}}^0$ are linearly independent.

Combining *Part (i)* and *Part (ii)*, we show $\overline{\mathbf{J}}^i$ are linearly independent if and only if $\overline{\mathbf{J}}^0$ are linearly independent. And therefore $\mathbf{J}^i$ has full column rank if and only if $\mathbf{J}^0$ has full column rank.

Step 2: We introduce $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i)$ and prove that they are identifiable if and only if $\mathbf{J}^i$ has full column rank. By following similar arguments in the *Proof of Theorem 1*, we have $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are identifiable if and only if $\mathbf{J}^i$ has full column rank, for $i = 1, \ldots, N$. Next, we define $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i)$ and the remaining is to show that they are identifiable if and only if $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are identifiable. Following the same arguments as *Step 2* in *Proof of Proposition 1*, we let

$$\epsilon_c^i = \boldsymbol{x}_{\boldsymbol{i}}^T \boldsymbol{\beta}_c = \beta_{0c} + \beta_{1c} x_{i1} + \cdots + \beta_{pc} x_{ip}.$$

for $c = 0, \ldots, C-1$. And

$$\omega_{jrc}^i = \gamma_{jrc} + \boldsymbol{z}_{ij}^T \boldsymbol{\lambda}_{jr} = \gamma_{jrc} + \lambda_{1jr} z_{ij1} + \cdots + \lambda_{qjr} z_{ijq}.$$

for $j = 1, \ldots, J$, $r = 0, \cdots, M_j - 1$ and $c = 0, \ldots, C-1$. Then according to Lemma 1, $(\boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i)$ are identifiable if and only if $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are identifiable. Hence the proof for *Step 2* is complete.

Step 3: The final step is to show $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are identifiable if and only if $(\boldsymbol{\epsilon}^i,\ \boldsymbol{\omega}^i)$ are identifiable. We have shown that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are not identifiable when $(\boldsymbol{\epsilon}^i,\ \boldsymbol{\omega}^i)$ are not identifiable in the *Proof of Proposition 1.* So all left to show is the necessary part that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ are identifiable when $(\boldsymbol{\epsilon}^i,\ \boldsymbol{\omega}^i)$ are identifiable. We prove this result by the method of contradiction. Assuming the contrary that $\boldsymbol{\beta}$ is not identifiable, there exist $\boldsymbol{\beta} \neq \boldsymbol{\beta}'$ such that $P(\boldsymbol{R}_i \mid \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = P(\boldsymbol{R}_i \mid \boldsymbol{\beta}', \boldsymbol{\gamma}, \boldsymbol{\lambda})$. According to the system of linear equations

$$\boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}^1 \\ \vdots \\ \boldsymbol{\epsilon}^N \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \cdots & x_{Np} \end{pmatrix} \begin{pmatrix} \beta_{00} & \cdots & \beta_{0(C-1)} \\ \vdots & \ddots & \vdots \\ \beta_{p0} & \cdots & \beta_{p(C-1)} \end{pmatrix} = \boldsymbol{X}\boldsymbol{\beta},$$

and because the full rank $\boldsymbol{X}$ is an injective mapping, we have $\boldsymbol{\beta} \neq \boldsymbol{\beta}'$ implies that $\boldsymbol{\epsilon} = \boldsymbol{X}\boldsymbol{\beta}$ is different from $\boldsymbol{\epsilon}' = \boldsymbol{X}\boldsymbol{\beta}'$ for at least one $\boldsymbol{\epsilon}^i \neq \boldsymbol{\epsilon}'^i$. However, since $\boldsymbol{\epsilon}^i$'s are identifiable, there exist no $\boldsymbol{\epsilon}^i \neq \boldsymbol{\epsilon}'^i$ such that $P(\boldsymbol{R}_i \mid \boldsymbol{\epsilon}^i, \boldsymbol{\omega}^i) = P(\boldsymbol{R}_i \mid \boldsymbol{\epsilon}'^i, \boldsymbol{\omega}^i)$. By this contradiction, we prove $\boldsymbol{\beta}$ is identifiable. Using similar arguments, we can show $\boldsymbol{\gamma}, \boldsymbol{\lambda}$ are also identifiable and hence complete the proof.

Combining *Steps 1–3*, we prove that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegCDMs are identifiable if and only if $\mathbf{J}^0$ has full column rank under $(A1)$–$(A3)$. □

To prove the main results in Section 4, we next introduce other useful lemmas and corollaries from existing works in literature. Lemma 3 and Corollaries 1–2 summarize the conditions for the global identifiability of general restricted latent class models proposed by Allman, Matias, and Rhodes (2009), which is based on the algebraic results in Kruskal (1977).

Before presenting these lemmas and corollaries, we introduce the decomposition of $\boldsymbol{\Psi}$ and some notation definitions. The decomposition of $\boldsymbol{\Psi}$ is similar as the decomposition of $\boldsymbol{\Phi}$ defined in Section 4 in the main text. We divide the total of J items into three mutually exclusive item sets $\mathcal{J}_1, \mathcal{J}_2$ and $\mathcal{J}_3$ containing J_1, J_2 and J_3 items respectively, with $J_1 + J_2 + J_3 = J$. For $t = 1, 2$ and 3, let $\mathcal{S}_{J_t}$ be the set containing the response patterns from items in $\mathcal{J}_t$ with cardinality of $\mathcal{S}_{J_t}$ to be $\kappa_t = |\mathcal{S}_{J_t}| = \prod_{j \in \mathcal{J}_t} M_j$. The submatrix $\boldsymbol{\Psi}_t$ has dimension $\kappa_t \times C$. The definition for the entries in $\boldsymbol{\Psi}_t$ is the same as in (10), except that each row of $\boldsymbol{\Psi}_t$ corresponds to one response patterns $\boldsymbol{r} \in \mathcal{S}_{J_t}$ while each row of $\boldsymbol{\Psi}$ corresponds to $\boldsymbol{r} \in \mathcal{S}'$.

Lemma 3. *(Kruskal, 1977) For $t = 1, 2$ and 3, denote $O_t = rank_K(\boldsymbol{\Psi}_t)$ as the Kruskal rank of $\boldsymbol{\Psi}_t$, where $\boldsymbol{\Psi}_t$ is a decomposed matrix of $\boldsymbol{\Psi}$. If*

$$O_1 + O_2 + O_3 \geq 2C + 2,$$

then $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2$ and $\boldsymbol{\Psi}_3$ uniquely determines the decomposition of $\boldsymbol{\Psi}$ up to simultaneous permutation and rescaling of columns.

Corollary 1. *(Allman et al., 2009) Consider the restricted latent class models with C classes. For $t = 1, 2$ and 3, let $\boldsymbol{\Psi}_t$ denote a decomposed matrix of $\boldsymbol{\Psi}$ and O_t denote its Kruskal rank. If*

$$O_1 + O_2 + O_3 \geq 2C + 2,$$

then the parameters of the model are uniquely identifiable, up to label swapping.

Corollary 2. *(Allman et al., 2009) Continue with the setting in Corollary 1. For $t = 1, 2, 3$, let $\boldsymbol{\Psi}_t$ denote a decomposed matrix of $\boldsymbol{\Psi}$ and κ_t denote its row dimension. If*

$$\min\{C, \kappa_1\} + \min\{C, \kappa_2\} + \min\{C, \kappa_3\} \geq 2C + 2,$$

Then the parameters of the restricted latent class models are generically identifiable up to label swapping.

Combining all these results as well as Proposition 2 in Huang and Bandeen-Roche (2004), we present Lemma 4, which is the key in the proof of Theorem 3.

Lemma 4. *For polytomous-response RegLCMs, $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are strictly identifiable if (A1) ,(A2) and (B3.a) hold, and are generically identifiable if (A1), (A2) and (B3.b) hold.*

(B3) The matrix $\boldsymbol{\Phi}$ can be decomposed into $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$, with Kruskal rank of each $\boldsymbol{\Phi}_t$ to be I_t and the dimension of each $\boldsymbol{\Phi}_t$ to be $\kappa_t \times C$. We have either

(B3.a) $I_1 + I_2 + I_3 \geq 2C + 2$; or

(B3.b) $\min\{C, \kappa_1\} + \min\{C, \kappa_2\} + \min\{C, \kappa_3\} \geq 2C + 2.$

The proof of Lemma 4 is provided in Section B.

Proof of Theorem 3. From condition $(C4)$, the Kruskal rank I_t of $\mathbf{\Phi}_t$ fulfill the arithmetic condition of (B3.a) in Lemma 4. Given $(A1)$ and $(A2)$, $(A1)$ and $(A2)$ also hold. According to Lemma 4, the RegLCMs are strictly identifiable at $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ for $i = 1, \ldots, N$. Following the similar arguments in *Step 2–3* from the *Proof of Theorem 2*, we show that $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegLCMs are identifiable given $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are identifiable under $(A3)$. Hence we complete the proof. □

Proof of Proposition 3. As mentioned in Section 4, $(C4^{*})$ is the sufficient condition for the identifiability of general restricted latent class models with binary responses according to Theorem 1 in Xu (2017). This condition is further extended to restricted latent class models with polytomous responses by Theorem 2 in Culpepper (2019). So for RegCDMs, $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are strictly identifiable given $(C4^{*})$ for $i = 1, \ldots, N$. Then based on the the similar arguments in *Step 2–3* from the *Proof of Theorem 2*, $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegCDMs are identifiable given $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are identifiable. □

Proof of Theorem 4. For $t = 1, 2$ and 3, the decomposed matrix $\mathbf{\Phi}_t$ and the decomposed matrix $\mathbf{\Psi}_t$ have the same row dimension κ_t. So given $(C4')$, condition (B3.b) in Lemma 4 holds. According to Lemma 4, RegLCMs are generically identifiable at $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ for $i = 1, \ldots, N$. Based on the similar arguments in *Step 2–3* from the *Proof of Theorem 2*, $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegLCMs are generically identifiable given $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are generically identifiable. □

Proof of Proposition 4. In Proposition 5.1(b) of Gu and Xu (2020), the condition $(C4'')$ is sufficient for the generic identifiability of CDMs. So for RegCDMs, $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are generically identifiable under $(C4'')$ for $i = 1, \ldots, N$. Based on the the similar arguments in *Step 2–3* from the *Proof of Theorem 2*, $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ in RegCDMs are generically identifiable given $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are generically identifiable. □

B Proofs of Lemmas

Proof of Lemma 1. For notational convenience, we use $\boldsymbol{\eta}$, $\boldsymbol{\Theta}$, $\boldsymbol{\epsilon}$ and $\boldsymbol{\omega}$ to denote the parameters $\boldsymbol{\eta}^i$, $\boldsymbol{\Theta}^i$, $\boldsymbol{\epsilon}^i$ and $\boldsymbol{\omega}^i$ of a general subject i. According to the definition of identifiability, $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are identifiable means that there exist no $(\boldsymbol{\eta}, \boldsymbol{\Theta}) \neq (\boldsymbol{\eta}', \boldsymbol{\Theta}')$ such that $P(\boldsymbol{R} = \boldsymbol{r} \mid \boldsymbol{\eta}, \boldsymbol{\Theta}) = P(\boldsymbol{R} = \boldsymbol{r} \mid \boldsymbol{\eta}', \boldsymbol{\Theta}')$. To prove Lemma 1 that $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are identifiable if and only if $(\boldsymbol{\epsilon}, \boldsymbol{\omega})$ are identifiable, we need to show that the transformation from $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ to $(\boldsymbol{\epsilon}, \boldsymbol{\omega})$ is bijective. We next illustrate this bijective mapping from $\boldsymbol{\eta}$ to $\boldsymbol{\epsilon}$ holds by showing $(\eta_0, \cdots, \eta_{C-1}) = (\eta'_0, \cdots, \eta'_{C-1})$ if and only if $(\epsilon_0, \cdots, \epsilon_{C-1}) = (\epsilon'_0, \cdots, \epsilon'_{C-1})$.

First, we show that $(\eta_0, \cdots, \eta_{C-1}) = (\eta'_0, \cdots, \eta'_{C-1})$ implies $(\epsilon_0, \cdots, \epsilon_{C-1}) = (\epsilon'_0, \cdots, \epsilon'_{C-1})$. For $c = 0, \ldots, C-1$, under the condition that

$$\eta_c = \frac{e^{\epsilon_c}}{1 + \sum_{s=1}^{C-1} e^{\epsilon_s}} = \frac{e^{\epsilon'_c}}{1 + \sum_{s=1}^{C-1} e^{\epsilon'_s}} = \eta'_c,$$

we can write

$$e^{\delta} = \frac{e^{\epsilon_0}}{e^{\epsilon'_0}} = \cdots = \frac{e^{\epsilon_c}}{e^{\epsilon'_c}} = \cdots = \frac{e^{\epsilon_{C-1}}}{e^{\epsilon'_{C-1}}} = \frac{1 + \sum_{s=1}^{C-1} e^{\epsilon_s}}{1 + \sum_{s=1}^{C-1} e^{\epsilon'_s}},$$

where e^{δ} denotes the common ratio among all $e^{\epsilon_c}/e^{\epsilon'_c}$. Hence

$$\delta = \epsilon_c - \epsilon'_c, \quad c = 0, \cdots, C-1. \tag{13}$$

Substituting every ϵ'_c with $\epsilon_c - \delta$ into the equation $\eta_0 = \eta'_0$, we have

$$\frac{e^{\epsilon_0}}{1 + \sum_{s=1}^{C-1} e^{\epsilon_s}} = \frac{e^{\epsilon_0 - \delta}}{1 + \sum_{s=1}^{C-1} e^{\epsilon_s - \delta}},$$

Further simplifying the above equation gives

$$\frac{1}{1 + \sum_{s=1}^{C-1} e^{\epsilon_s}} = \frac{1}{e^{\delta} + \sum_{s=1}^{C-1} e^{\epsilon_s}},$$

and then we have

$$e^{\delta} + \sum_{s=1}^{C-1} e^{\epsilon_s} = 1 + \sum_{s=1}^{C-1} e^{\epsilon_s},$$

which has unique solution $\delta = 0$. Taking $\delta = 0$ back into (13), we have $\epsilon_c = \epsilon'_c$ for all $c = 0, \ldots, C-1$. Therefore $\boldsymbol{\epsilon} = (\epsilon_0, \cdots, \epsilon_{C-1})$ is equivalent to $\boldsymbol{\epsilon}' = (\epsilon'_0, \cdots, \epsilon'_{C-1})$.

Next we prove $(\epsilon_0, \cdots, \epsilon_{C-1}) = (\epsilon'_0, \cdots, \epsilon'_{C-1})$ implies $(\eta_0, \cdots, \eta_{C-1}) = (\eta'_0, \cdots, \eta'_{C-1})$. This part is straightforward as $(\epsilon_0, \cdots, \epsilon_{C-1}) = (\epsilon'_0, \cdots, \epsilon'_{C-1})$ implies that for any $c = 0, \ldots, C-1$, we have

$$\frac{\exp(\epsilon_c)}{1 + \sum_{s=1}^{C-1} \exp(\epsilon_s)} = \frac{\exp(\epsilon'_c)}{1 + \sum_{s=1}^{C-1} \exp(\epsilon'_s)}.$$

Equivalently, we show $\eta_c = \eta'_c$ for any $c = 0, \ldots, C-1$. So $(\eta_0, \cdots, \eta_{C-1}) = (\eta'_0, \cdots, \eta'_{C-1})$. Combining the above arguments, we prove $\boldsymbol{\eta} = \boldsymbol{\eta'}$ if and only if $\boldsymbol{\epsilon} = \boldsymbol{\epsilon'}$.

Similar arguments can be applied to show $\boldsymbol{\Theta} = \boldsymbol{\Theta'}$ if and only if $\boldsymbol{\omega} = \boldsymbol{\omega'}$. Hence $(\boldsymbol{\eta}, \boldsymbol{\Theta})$ are identifiable if and only if $(\boldsymbol{\epsilon}, \boldsymbol{\omega})$ are identifiable. □

Proof of Lemma 2. We prove the the first part, that is, $\mathbf{J}^0$ has full column rank if and only if $\overline{\mathbf{J}}^0$ are linearly independent. The second part regarding $\mathbf{J}^i$ can be similarly proved.

To show the linear independence of $\mathbf{J}^0$ or $\overline{\mathbf{J}}^0$, we need to establish the relationship between the two linear combinations as follows. For any linear combinations of the columns in $\mathbf{J}^0$ with coefficients h_c^0's and l_{jrc}^0's, there exist a_c^0's and b_{jrc}^0's such that the following equation holds.

$$\sum_{c=1}^{C-1} h_c^0(\boldsymbol{\psi}_c^0 - \boldsymbol{\psi}_0^0) + \sum_{j=1}^{J} \sum_{r=1}^{M_j-1} \sum_{c=0}^{C-1} \left\{ l_{jrc}^0 \eta_c^0 \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^0} - \frac{\mathbf{I}\{r_j = 0\}}{\theta_{j0c}^0}\right) \odot \boldsymbol{\psi}_c^0 \right\} \tag{14}$$

$$= \sum_{c=0}^{C-1} a_c^0 \boldsymbol{\psi}_c^0 + \sum_{j=1}^{J} \sum_{r=0}^{M_j-1} \sum_{c=0}^{C-1} \left\{ b_{jrc}^0 \eta_c^0 \left(\frac{\mathbf{I}\{r_j = r\}}{\theta_{jrc}^0}\right) \odot \boldsymbol{\psi}_c^0 \right\}, \tag{15}$$

where

$$a_c^0 = \begin{cases} h_c^0, & \text{if } c \neq 0, \\ -(h_1^0 + \cdots + h_{C-1}^0), & \text{if } c = 0, \end{cases} \tag{16}$$

and for any $j = 1, \ldots, J$, $c = 0, \ldots, C-1$,

$$b_{jrc}^0 = \begin{cases} l_{jrc}^0, & \text{if } r \neq 0, \\ -(l_{j1c}^0 + \cdots + l_{j(M_j-1)c}^0), & \text{if } r = 0. \end{cases} \tag{17}$$

With the above relationship established, we next show that $\mathbf{J}^0$ has full column rank if and only if $\overline{\mathbf{J}}^0$ are linearly independent. When $\overline{\mathbf{J}}^0$ are linearly independent, (15) = 0 implies

$a_c^0 = 0$ and $b_{jrc}^0 = 0$, which further implies $h_c^0 = 0$ and $l_{jrc}^0 = 0$ by (16) and (17). So (14) $= 0$ implies $h_c^0 = 0$ and $l_{jrc}^0 = 0$. Hence, $\mathbf{J}^0$ has full column ranks. Similarly, when $\mathbf{J}^0$ has full column ranks, (14) $= 0$ implies $h_c^0 = 0$ and $l_{jrc}^0 = 0$ which further implies $a_c^0 = 0$ and $b_{jrc}^0 = 0$ by (16) and (17). So (15) $= 0$ implies $a_c^0 = 0$ and $b_{jrc}^0 = 0$. Hence, $\overline{\mathbf{J}}^0$ are linearly independent. □

Proof of Lemma 4. This proof is adapted from the *Proof of Proposition 2* in Huang and Bandeen-Roche (2004). Before presenting the proof, we set up a few notations. In Section 4, $\mathbf{\Phi}$ can be decomposed into $\mathbf{\Phi}_1, \mathbf{\Phi}_2$ and $\mathbf{\Phi}_3$, where each $\mathbf{\Phi}_t$ has Kruskal rank I_t and row dimension κ_t. And in Appendix A, $\mathbf{\Psi}^i$ can be decomposed into $\mathbf{\Psi}_1^i$, $\mathbf{\Psi}_2^i$ and $\mathbf{\Psi}_3^i$, where each $\mathbf{\Psi}_t^i$ has Kruskal rank O_t and the same row dimension κ_t as $\mathbf{\Phi}_t$. Denote the columns in $\mathbf{\Phi}_t$ to be $\boldsymbol{\phi}_{t0}, \cdots, \boldsymbol{\phi}_{t(C-1)}$ and the columns in $\mathbf{\Psi}_t^i$ to be $\boldsymbol{\psi}_{t0}^i, \cdots, \boldsymbol{\psi}_{t(C-1)}^i$. Conditions $(A1)$ and $(A2)$ are shown to be necessary in Section 2 and assumed to hold. To prove (B3.a) is sufficient for the strict identifiability of $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$, we first need to show that for $t = 1, 2$ and 3, given $\mathbf{\Phi}_t$ has Kruskal rank I_t, the equation $O_t \geq I_t$ holds, so that $I_1 + I_2 + I_3 \geq 2C + 2$ from (B3.a) implies $O_1 + O_2 + O_3 \geq 2C + 2$. Then based on Corollary 1 that $(\boldsymbol{\eta}^i, \mathbf{\Theta}^i)$ are strictly identifiable under the condition that $O_1 + O_2 + O_3 \geq 2C + 2$, we complete the proof of strict identifiability.

The remaining part is to show $O_t \geq I_t$ for $t = 1, 2$ and 3. Without loss of generality, we only show $O_1 \geq I_1$, then $O_2 \geq I_2$ and $O_3 \geq I_3$ can be similarly proved. Under the condition that any set of I_1 columns in $\mathbf{\Phi}_1$ are linearly independent, $\boldsymbol{\phi}_{1\sigma(1)}, \cdots, \boldsymbol{\phi}_{1\sigma(I_1)}$ are linearly independent for any permutation σ on $\{1, \ldots, I_1\}$ such that $\{\sigma(1),\ \sigma(2), \cdots,\ \sigma(I_1)\} \subseteq \{0, \cdots, C-1\}$. To show $O_t \geq I_t$, we need $\boldsymbol{\psi}_{1\sigma(1)}, \cdots, \boldsymbol{\psi}_{1\sigma(I_1)}$ to be linearly independent for any permutation set $\{\sigma(1),\ \sigma(2), \cdots,\ \sigma(I_1)\}$. The linear combinations of $\boldsymbol{\phi}_{1\sigma(1)}, \cdots, \boldsymbol{\phi}_{1\sigma(I_1)}$ can be associated with the linear combinations of $\boldsymbol{\psi}_{1\sigma(1)}, \cdots, \boldsymbol{\psi}_{1\sigma(I_1)}$ as follows. For any permutation σ and $a_{\sigma(c)}$, there exists $b_{\sigma(c)}$ and $\boldsymbol{Y}_\mathbf{1}^i$ such that

$$\sum_{c=1}^{I_1} a_{\sigma(c)} \boldsymbol{\psi}_{1\sigma(c)}^i = (\sum_{c=1}^{I_1} b_{\sigma(c)} \boldsymbol{\phi}_{1\sigma(c)}) \odot \boldsymbol{Y}_\mathbf{1}^i \tag{18}$$

where

$$\boldsymbol{Y}_{\mathbf{1}}^{i} = \left(\prod_{j\in\mathcal{J}_1} \exp(\lambda_{1jr_j}z_{ij1} + \cdots + \lambda_{qjr_j}z_{ijq}) : \boldsymbol{r} = (r_1,\ldots,r_J) \in \mathcal{S}_{J_1}\right)_{\kappa_1\times 1},$$

$$b_{\sigma(c)} = a_{\sigma(c)} \prod_{j\in\mathcal{J}_1} \frac{1+\sum_{s=1}^{M_j-1} e^{\gamma_{js\sigma(c)}}}{1+\sum_{s=1}^{M_j-1}\exp(\gamma_{js\sigma(c)} + \lambda_{1js}z_{ij1} + \cdots + \lambda_{qjs}z_{ijq})}. \tag{19}$$

To show $\boldsymbol{\psi}_{1\sigma(1)}, \cdots, \boldsymbol{\psi}_{1\sigma(I_1)}$ to be linearly independent, we need to show $\sum_{c=1}^{I_1} a_{\sigma(c)}\boldsymbol{\psi}_{1\sigma(c)}^{i} = 0$ implies $a_{\sigma(c)} = 0$ for any σ. Based on (18), we have $\sum_{c=1}^{I_1} a_{\sigma(c)}\boldsymbol{\psi}_{1\sigma(c)}^{i} = 0$ implies $\sum_{c=1}^{I_1} b_{\sigma(c)}\boldsymbol{\phi}_{1\sigma(c)} = 0$. Under the condition that $\boldsymbol{\phi}_{1\sigma(1)}, \cdots, \boldsymbol{\phi}_{1\sigma(I_1)}$ are linear independent, $\sum_{c=1}^{I_1} b_{\sigma(c)}\boldsymbol{\phi}_{1\sigma(c)} = 0$ implies $b_{\sigma(1)} = \cdots = b_{\sigma(I_1)} = 0$. And by (19), $a_{\sigma(1)} = \cdots = a_{\sigma(I_1)} = 0$. Hence $\boldsymbol{\psi}_{1\sigma(1)}, \cdots, \boldsymbol{\psi}_{1\sigma(I_1)}$ are linearly independent for any σ. Hence we show $O_1 \geq I_1$ and complete the proof for strict identifiability.

For condition (B3.b), because each Ψ_t^i has row dimension κ_t the same as Φ_t does and we have $\min\{C,\kappa_1\} + \min\{C,\kappa_2\} + \min\{C,\kappa_3\} \geq 2C+2$, according to Corollary 2, $(\boldsymbol{\eta}^i, \boldsymbol{\Theta}^i)$ are generically identifiable under (B3.b) for all $i = 1,\ldots,N$. □

References

Allman, E. S., Matias, C., & Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables. *Annals of Statistics*, *37*(6A), 3099–3132.

Culpepper, S. A. (2019). An exploratory diagnostic model for ordinal responses with binary attributes: Identifiability and estimation. *Psychometrika*, *84*(4), 921–940.

Gu, Y., & Xu, G. (2020). Partial identifiability of restricted latent class models. *Annals of Statistics*, *48*(4), 2082–2107.

Huang, G.-H., & Bandeen-Roche, K. (2004). Building an identifiable latent class model with covariate effects on underlying and measured variables. *Psychometrika*, *69*(1), 5–32.

Kruskal, J. B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, *18*(2), 95–138.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, *39*(3), 577–591.

Xu, G. (2017). Identifiability of restricted latent class models with binary responses. *Annals of Statistics*, *45*(2), 675–707.