A Multivariate Latent Class Profile Analysis with Latent Group

Jung Wun Lee

April 4, 2019

Biometrics 64, 1–17
December 2008
DOI: 10.1111/j.1541-0420.2005.00454.x

A Multivariate Latent Class Profile Analysis with Latent Group

Summary: This paper suggests a new type of latent variable model which discovers the association between several categorical latent variables. A set of repeatedly measured categorical response variables forms a latent profile variable, while the other set of item variables identifies a latent group variable. Latent class profile analysis with group variable (GLCPA) explains an association between these two categorical latent variables as a form of two-dimensional contingency table. We applied GLCPA model to the NLSY 97 data to investigate the association between of depression process and the longitudinal behaviors of substance use development among adolescents who experienced an Authoritarian parental styles in their youth.
Key words: Latent class analysis, Longitudinal data, Logistic regression, Recursive EM algorithm.

1. Introduction

Latent class analysis (LCA) is one type of finite mixture model which can be applied for a set of discrete response random variable. It summarizes the structure of population distribution by defining several partitions of population (i.e., latent classes) which cannot be observed directly, but may be discovered with respect to the patterns for manifest response variables. This LCA framework has been expanded to be utilized for more complicated data structures such as a repeatedly measured longitudinal data in chung2011latent, a vectorized joint structure in Jeon et al. (2017), and the hierarchical group-outcome structure in Lee and Chung (2017).

In this article, a new type of LCPA with group variable has been proposed which consists of a typical multivariate latent class profile model and an additional categorical latent variable as a latent group variable. A set of repeatedly measured categorical response variables identifies a vector of categorical latent variables for each time points, and a latent profile variable is defined to divide the population into homogeneous subgroups whose sequential patterns of latent class memberships are common. In addition, another set of categorical response items defines a discrete latent variable using conventional LCA framework. Our proposed model allows the prevalence of latent profiles to be differed with respect to the latent group memberships. Namely, the prevalence of latent profiles are defined as the proportion of latent profiles given a certain latent group membership, and this conditional probability explain the existing association between latent profile variable and group variable in terms of condition probability.

The rest contents of this article are as follows. The description of the GLCPA and the estimation methods for the model parameters are presented in Section ModelSelection and EstimationSelection, respectively. In Section SimulSection, we examined the parameter estimation and inference procedure through empirical simulation, and the simulation results are available in Appendix. In Section NLSY, we illustrate the practical usefulness of our new model by analyzing the NLSY 97 data using discrete item variables that are related with adolescent depression and longitudinal behaviors in substance use including alcohol, cigarette, and marijuana. In Section Conclu, we summarize this paper and discuss about the further research area.

1.1 *Latent class analysis*

A latent class analysis is a classical methodology that divides the population into homogenous subgroups with respect to response patterns for manifest items. It postulates that a ditribution of a set of categorical random variable is a mixture of finite classes with their respective response patterns. Suppose there are P categorical manifest items $Z_1, \ldots, Z_P$. The responses of each manifest item for the ith individual are obtained as a P-dimensional vector $\mathbf{z}_i = [z_{i1}, \cdots, z_{iP}]^T$, where z_{ip} can take any value from $1, \ldots, r_p$ for $p = 1, \ldots, P$. Let the latent class variable D has G categories, then the observed-data likelihood of LCA can be written as follows:

$$\begin{aligned} P(\mathbf{Z}_i = \mathbf{z}_i) &= \sum_{d=1}^{G} P(D = d, \mathbf{Z}_i = \mathbf{z}_i) = \sum_{d=1}^{G} P(D = d) P(\mathbf{Z}_i = \mathbf{z}_i \mid D = d) \\ &= \sum_{d=1}^{G} P(D = d) \prod_{p=1}^{P} P(Z_{ip} = z_{ip} \mid D = d) \\ &= \sum_{d=1}^{G} \delta_d \prod_{p=1}^{P} \prod_{h=1}^{r_p} \phi_{ph|d}^{I(z_{ip}=h)} \end{aligned} \quad (1)$$

Here, $I(z_{ip} = h)$ is the indicator function which is 1 when $z_{ip} = h$ and 0 otherwise. The likelihood of LCA given in (1) is constructed under the local independence assumptions, implying that the manifest items are conditionally independent when a latent class membership is given. Here, $\phi_{ph|d} = P(Z_p = h \mid D = d)$, referred as the primary measurement parameter, explains the relationship between the latent class and the pth manifest item, and $\delta_d = P(D = d)$ represents the prevalence of latent class membership d. Since all parameters in (1) are conditional probabilities, the sum-to-one and non-negative constraints are explicit (i.e., $\sum_{d=1}^{G} \delta_d = 1$ and $\sum_{h=1}^{r_p} \phi_{ph|d} = 1$ for $p = 1, \ldots, P$, $d = 1, \ldots, G$).

1.2 *Latent class profile analysis*

Latent class profile analysis (LCPA) has been introduced to explain the longitudinal patterns when the LCA is applied to the repeated measured responses Chung et al. (2011). In LCPA, each sets of manifest items measure a categorical latent variable, and the sequential patterns of identified latent classes are summarize by a latent profile variable. As a result, observations who share the same latent profile membership will have common sequential patterns of latent class memberships for each identified latent variables. In this manner, LCPA provides a statistical tool which allows researchers to discover the meaningful subgroup based on the representative sequential pattern of unobservable memberships for several latent classes.

Let C_{jt} denote the jth latent class variable having K_j nominal categories for $j = 1, \ldots, J$ at stage t, where $t = 1, \ldots T$. For each time stage, a vector of J latent variables $\mathbf{C}_t = [C_{1t}, \ldots, C_{Jt}]^T$ can be summarized as a contingency table with $\prod_{j=1}^{J} K_j$ cells, showing all possible combinations of class memberships. Thus, the T-sequences of J latent class variables will be written in a contingency table with $(\prod_{j=1}^{J} K_j)^T$ cells. Among all possible combinations of sequential patterns, LCPA discovers the representative sequential patterns and categorize them as latent profiles. Let the latent profile variable U have S nominal categories describing the most common stage-sequential patterns of J latent class memberships. Let $\mathbf{Y}_t = [\mathbf{Y}_{1t}, \ldots, \mathbf{Y}_{Jt}]$, where $\mathbf{Y}_{jt} = [Y_{1jt}, \ldots, Y_{M_j jt}]^T$ be a set of J vectors of discrete responses to M_j items to measure the jth latent class membership at stage t, where each variable $Y_{m_j jt}$ can take any value from 1 to r_{m_j} for $m_j = 1, \ldots, M_j$ and $j = 1, \ldots, J$. Then, the complete-data likelihood of the model of the probability of the latent profile $U = u$, the latent class memberships $\mathbf{C}_t$, and the responses $\mathbf{Y}_t$ for $t = 1, \ldots, T$ would be as follows:

$$\begin{aligned} L_i^* &= P(U = u, \mathbf{C}_1 = \mathbf{c}_1, \ldots, \mathbf{C}_T = \mathbf{c}_T, \mathbf{Y}_{i1} = \mathbf{y}_{i1}, \ldots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}) \\ &= P(U = u) P(\mathbf{C} = \mathbf{c} \mid U = u) P(\mathbf{Y}_i = \mathbf{y}_i \mid \mathbf{C} = \mathbf{c}) \\ &= P(U = u) \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} \left[P(C_{jt} = c_{jt} \mid U = u) \prod_{m_j=1}^{M_j} P(Y_{im_j jt} = y_{im_j jt} \mid C_{jt} = c_{jt}) \right] \right\} \\ &= \gamma_u \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} \left[\eta_{c_{jt}|u}^{(j,t)} \prod_{m_j=1}^{M_j} \prod_{k=1}^{k_{m_j}} \rho_{m_j k|c_{jt}}^{(j,t)}{}^{I(y_{im_j jt}=k)} \right] \right\}, \end{aligned} \tag{2}$$

where $I(y_{im_j jt} = k)$ is the indicator function which is 1 when $y_{im_j jt} = k$ and 0 otherwise.

1.3 *Latent class profile analysis with multiple latent group variables*

The LCPA with latent group variables (GLCPA) postulates that the distribution of latent profile variable can be affected by another latent class variable which can be identified through LCA structure. Combining the LCA structure as group variable and LCPA structure as an outcome, we propose GLCPA and illustrate the model in Figure 1.

[Figure 1 about here.]

A sequence of J latent variables $\boldsymbol{C}_t = [C_{1t}, \cdots, C_{Jt}]^T$ for $t = 1, \cdots, T$ in Figure 1 constitute the LCPA that are associated through latent profile variable U, and each latent variable C_{jt} is identified by the jth set of manifest items $\boldsymbol{Y}_{ijt} = [Y_{i1jt}, \ldots, Y_{iM_j jt}]^T$ at time stage t. Another latent variable D is the ordinary latent class model which can be identified through the manifest items $\boldsymbol{Z}_i = [Z_{i1}, \ldots, Z_{iP}]^T$, and the response variable $\boldsymbol{Z}$ can be measured in any time stage $t = 1, \cdots, T$, or in any other time stage. As discussed in Section **??**, the distribution of outcome latent profile variable U is affected by the latent group membership $D = d$.

(a) $\rho_{m_j k|c_{jt}}^{(j,t)} = P(Y_{m_j jt} = k \mid C_{jt} = c_{jt})$ denotes the probability of the response k to the m_jth item measuring the jth latent variable C_{jt}, for a given class c_{jt} of the jth latent variable C_{jt} at stage t.

(b) $\phi_{ph|d} = P(Z_{ip} = h \mid D = d)$ denotes the probability of the response h to the pth item measuring the latent variable D, for a given class d.

(c) $\eta^{(j,t)}_{c_{jt}|u} = P(C_{jt} = c_{jt} \mid U = u)$ denotes the conditional probability of belonging to class c_{jt}, the class membership of j latent variable C_{jt} at stage t, when a latent profile variable U has a profile membership u.

(d) $\gamma_{u|d} = P(U = u \mid D = d)$ denotes the probability that individual has a latent profile u among S latent profiles, given that its latent group membership is d.

(e) $\delta_d = P(D = d)$ denotes the probability that individual belongs to dth latent group D.

The primary measurement parameter ρ and ϕ idenfity the underlying categorical latent variables, depicting the probability of responding to the categorical response variable when the latent class memberships $C_{jt} = c_{jt}$, $D = d$ are given, respectively. The secondary measurement parameter η depicts the relationship between each latent class c_{jt} of C_{jt} and a latent profile u of U for $u = 1, \ldots, S$. Each identified latent profile can be explained through a set of estimated secondary measurement parameters as a individual's sequential patterns of changing latent class membership as stage flows.

The GLCPA assumes the following conditions: (1) the latent profile membership is related to the manifest items only through the class membership of each latent variable at each time wave, (2) the response variable Y_{im_jjt}, Z_{ip} are correlated only through the corresponding latent variable, (3) each latent variables are correlated only through the latent profile variable. (4) the group latent variable is only related with each identified latent variables in LCPA model only through latent profile varaible. Based on the condition (1), response variables $\boldsymbol{Y}_{ijt} = [Y_{im_jjt}, \cdots, Y_{im_jjt}]^T$ corresponding to jth latent variable at time t become mutually independent when the jth latent class membership at time t (i.e., c_{jt}) is known. Likewise, condition (2) allows each identified latent variables C_{jt} for $j = 1, \cdots, J, t = 1, \cdots, T$ be independent when the latent profile membership U is given. Using the notation given in (2) and (1.3), the complete-data likelihood of the GLCPA for the ith observation is written as follows:

$$\begin{aligned}
L_i^* &= P(U=u, D=d, \mathbf{C}_1=\mathbf{c}_1, \ldots, \mathbf{C}_T=\mathbf{c}_T, \mathbf{Y}_{i1}=\mathbf{y}_{i1}, \ldots, \mathbf{Y}_{iT}=\mathbf{y}_{iT}, \mathbf{Z}_i) \\
&= P(D=d)P(U=u \mid D=d)P(\mathbf{C}_1=\mathbf{c}_1, \ldots, \mathbf{C}_T=\mathbf{c}_T \mid U=u) \\
&\quad \times P(\mathbf{Y}_{i1}=\mathbf{y}_{i1}, \ldots, \mathbf{Y}_{iT}=\mathbf{y}_{iT}, \mathbf{Z}_i \mid \mathbf{C}_1=\mathbf{c}_1, \ldots, \mathbf{C}_T=\mathbf{c}_T, D=d) \qquad (3) \\
&= P(U=u \mid D=d) \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} \left[P(C_{jt}=c_{jt} \mid U=u) \prod_{m_j=1}^{M_j} P(Y_{m_j jt}=y_{m_j jt} \mid C_{jt}=c_{jt}) \right] \right\} \\
&\quad \times P(D=d) \prod_{p=1}^{P} P(Z_p=z_p \mid D=d) \\
&= \gamma_{u|d} \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} \left[\eta_{c_{jt}|u}^{(j,t)} \prod_{m_j=1}^{M_j} \prod_{k=1}^{k_{m_j}} \rho_{m_j k|c_{jt}}^{(j,t)}{}^{I(y_{m_j jt}=k)} \right] \right\} \delta_d \prod_{p=1}^{P} \prod_{h=1}^{r_p} \phi_{ph|d}^{I(Z_{ip}=h)}
\end{aligned}$$

The likelihood of the model that we actually observe (i.e., the observed-data likelihood) can be derived by the marginal summation of (3) with respect to all considered latent variables:

$$L_i = P(\mathbf{Z}_i=\mathbf{z}_i, \mathbf{Y}_{i1}=\mathbf{y}_{i1}, \cdots, \mathbf{Y}_{iT}=\mathbf{y}_{iT}) = \sum_{u=1}^{S} \sum_{c_{11}=1}^{K_1} \cdots \sum_{c_{JT}=1}^{K_J} \sum_{d=1}^{G} L_i^* \qquad (4)$$

The prevalence of the latent profile may also be affected by the individuals factors such as gender. As illustrated in Figure 1, we can construct the multinomial logistic regression model by treating the identified latent profile variables as a response variable. While the conventional multinomial logistic regression untilizes the observed values of response variables and covariates, the regression on unobservable latent profile memberships relates the covariates with posterior probabilities which will be discussed on Eq. 9. Suppose we have a vector of covariates $\boldsymbol{x}_i = [x_{i1}, \ldots, x_{ip}]^T$ for the ith observation, then the latent profile can be written as a function of covariates in multinomial logistic regression form.

$$\begin{aligned}
L^*(\boldsymbol{X}_i) &= \gamma_{u|d}(\boldsymbol{X}_i) \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} \left[\eta_{c_{jt}|u}^{(j,t)} \prod_{m_j=1}^{M_j} \prod_{k=1}^{r_{m_j}} \rho_{m_j k|c_{jt}}^{(j,t)}{}^{I(y_{im_j jt}=k)} \right] \right\} \delta_d \prod_{p=1}^{P} \prod_{h=1}^{r_p} \phi_{ph|d}^{I(Z_{ip}=h)} \qquad (5) \\
\gamma_{u|d}(\boldsymbol{X}_i) &= \frac{\exp(\mathbf{X}_i \boldsymbol{\beta}_{u|d})}{\sum_{d=1}^{G} \exp(\boldsymbol{X}_i \boldsymbol{\beta}_{u|d})}
\end{aligned}$$

Here, the vector of logistic regression coefficients $\boldsymbol{\beta}_{u|d} = [\beta_{1u|d}, \ldots, \beta_{pu|d}]^T$ is interpreted as the log-odds ratio that an individual belongs to a specific latent profile u versus to a baseline latent class, given the latent group membership $D = d$. Finally, the likelihood of the model that we actually observe (i.e., the observed-data likelihood) can be derived by the marginal summation of (3) with respect to all considered latent variables:

$$L_(\boldsymbol{X}_i) = P(\mathbf{Z}_i = \mathbf{z}_i, \mathbf{Y}_{i1} = \mathbf{y}_{i1}, \cdots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}) = \sum_{u=1}^{S} \sum_{c_{11}=1}^{K_1} \cdots \sum_{c_{JT}=1}^{K_J} \sum_{d=1}^{G} L^*(\boldsymbol{X}_i) \quad (6)$$

2. Parameter estimation and model selection

We adopted the Expectation-Maximization algorithm (Dempster et al., 1977) to estimate the ML estimates of parameters. To determine the number of classes for each latent variable, we investigate the model with various number of classes and chose the most appropriate one. We adopted AIC, BIC criteria.

2.1 *Recursive Expectation-Maximization Algorithm*

The typical EM algorithm implements Expectation step and Maximization steps for each iteration, and repeats these steps until the solutions satisfies the convergence threshold.

E-step. The expectation of the complete log-likelihood is computed using the formula in Eq. (3) as follows:

$$\begin{aligned} E\left(\sum_{i=1}^{n} \log L^*(\boldsymbol{x}_i)\right) \quad = \quad & \sum_{i=1}^{n} E[I(D=d)] \log \delta_d + \sum_{i=1}^{n} E[I(U=u, D=d)] \log \gamma_{u|d}(\boldsymbol{x}_i) \\ & + \sum_{i=1}^{n} \sum_{p=1}^{P} \sum_{h=1}^{r_p} E[I(D=d \mid Z_{ip}=h)] \log \phi_{ph|d} \\ & + \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{j=1}^{J} E[I(C_{jt}=c_{jt}, U=u)] \log \eta^{(j,t)}_{c_{jt}|u} \\ & + \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{j=1}^{J} \left\{ \sum_{m_j=1}^{M_j} \sum_{k=1}^{r_{m_j}} E\big[I(C_{jt}=c_{jt} \mid Y_{im_jjt}=k)\big] \log \rho^{(j,t)}_{m_jk|c_{jt}} \right\} \end{aligned} \quad (7)$$

To obtain the expectations of indicator functions in Eq. (7), we define the joint posterior probability of latent variables given the ith observed responses and covariates as follows:

$$\theta_{i(u,d,\boldsymbol{c}_1,\ldots,\boldsymbol{c}_T)} = P(U = u, D = d, \boldsymbol{C}_1 = \boldsymbol{c}_1, \ldots, \boldsymbol{C}_T = \boldsymbol{c}_T \mid \boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{iT}, \boldsymbol{z}_i, \boldsymbol{x}_i) = \frac{L^*(\boldsymbol{x}_i)}{L(\mathbf{x}_i)} \quad (8)$$

for $i = 1, \ldots, n$, $u = 1, \ldots, S$, $d = 1, \ldots, G$, $c_{jt} = K_j$, $j = 1, \ldots, J$, and $t = 1, \ldots, T$. Since the conditional distributions of latent variables given response variables follow the multinomial distribution respectively, the expectations of indicator functions can be expressed in terms of marginal posterior probabilities $\theta_{i(u)}$, $\theta_{i(d)}$, $\theta_{i(u,c_{jt})}$, and $\theta_{i(c_{jt})}$, respectively:

$$\begin{aligned}
E\left[I(D = d \mid \boldsymbol{Y}_i, \boldsymbol{z}_i)\right] &= \theta_{i(u)} = \sum_{u=1}^{S}\prod_{j=1}^{J}\prod_{t=1}^{T}\left\{\sum_{c_{jt}=1}^{K_j}\right\}\theta_{i(u,d,\boldsymbol{c}_1,\ldots,\boldsymbol{c}_T)} \\
E\left[I(U = u \mid \boldsymbol{Y}_i, \boldsymbol{z}_i)\right] &= \theta_{i(u)} = \sum_{d=1}^{G}\prod_{j=1}^{J}\prod_{t=1}^{T}\left\{\sum_{c_{jt}=1}^{K_j}\right\}\theta_{i(u,d,\boldsymbol{c}_1,\ldots,\boldsymbol{c}_T)} \qquad (9) \\
E\left[I(C_{jt} = c_{jt}, U = u \mid \boldsymbol{Y}_i, \boldsymbol{z}_i)\right] &= \theta_{i(u,c_{jt})} = \sum_{d=1}^{G}\prod_{j'\neq j}^{K}\prod_{t'\neq t}^{T}\left\{\sum_{c_{j't'}=1}^{K_{j'}}\right\}\theta_{i(u,d,\boldsymbol{c}_1,\ldots,\boldsymbol{c}_T)} \\
E\left[I(C_{jt} = c_{jt} \mid \mathbf{Y}_i, \boldsymbol{z}_i)\right] &= \theta_{i(c_{jt})} = \sum_{u=1}^{S}\sum_{u=1}^{S}\theta_{i(u,c_{jt})}
\end{aligned}$$

Once the overall posterior probability in (8) is obtained, the marginal posterior probabilities can be easily calculated. We adopt the recursive formula to the E-step using the forward and backward probabilities introduced in Chang and Chung (2013). Let α and λ represent the forward and backward probabilities, respectively:

$$\begin{aligned}
\alpha_{it}\left(u, \mathbf{c}_t\right) &= P(\boldsymbol{Y}_1 = \boldsymbol{y}_1, \ldots, \boldsymbol{Y}_t = \boldsymbol{y}_t, \boldsymbol{C}_t = \boldsymbol{c}_t \mid u) \\
&= \sum_{c_{1(t-1)}=1}^{K_1}\cdots\sum_{c_{J(t-1)}=1}^{K_J}\alpha_{i(t-1)}\left(u, \boldsymbol{c}_{(t-1)}\right)\prod_{j=1}^{J}\left[\eta_{c_{jt}|u}^{(j,t)}\prod_{m_j=1}^{M_j}\prod_{k=1}^{r_{m_j}}\rho_{m_jk|c_{jt}}^{(j,t)}{}^{I(y_{im_jjt}=k)}\right] \\
\lambda_{it}\left(u, \mathbf{c}_t\right) &= P(\boldsymbol{Y}_{t+1} = \boldsymbol{y}_{t+1}, \ldots, \boldsymbol{Y}_T = \boldsymbol{y}_T \mid \mathbf{c}_t, u) \qquad (10) \\
&= \sum_{c_{1(t+1)}=1}^{K_1}\cdots\sum_{c_{J(t+1)}=1}^{K_J}\lambda_{i(t+1)}\left(u, \boldsymbol{c}_{(t+1)}\right)\prod_{j=1}^{J}\left[\eta_{c_{j(t+1)}|u}^{(j,t+1)}\prod_{m_j=1}^{M_j}\prod_{k=1}^{r_{m_j}}\rho_{m_jk|c_{j(t+1)}}^{(j,t+1)}{}^{I(y_{im_jj(t+1)}=k)}\right]
\end{aligned}$$

The posterior probability of latent class memberships at stage t can be obtained as follows:

$$\theta_{i(u,d,\boldsymbol{c}_t)} = \frac{\delta_d \prod_{p=1}^{p} \prod_{h=1}^{r_p} \phi_{ph|d}^{I(z_{ip}=h)} \gamma_{u|d}(\boldsymbol{x}_i)\alpha_{it}(u,\boldsymbol{c}_t)\lambda_{it}(u,\boldsymbol{c}_t)}{\sum_{g=1}^{G} \delta_g \prod_{p=1}^{p} \prod_{h=1}^{r_p} \phi_{ph|g}^{I(z_{ip}=h)} \sum_{s=1}^{S} \gamma_{s|g}(\boldsymbol{x}_i) \sum_{c_{1T}=1}^{K_1} \cdots \sum_{c_{JT}=1}^{K_J} \alpha_{iT}(s,\boldsymbol{c}_T)} \quad (11)$$

M-step. The M-step maximizes the expected complete-data likelihood of the GLCPA with respect to the model parameters. Since the sum of parameters that are used in measuring each latent variables are constrained to be one (for instance, $\sum_{d=1}^{G} \delta_d = 1$, $\sum_{u=1}^{S} \gamma_{u|d} = 1, d = 1, \cdots, G$), we adopted Lagrange multiplier to obtain the ML estimator under such constraints.

$$\hat{\gamma}_{(u|d)} = \frac{\sum_{i=1}^{n} \theta_{i(u,d)}}{\sum_{i=1}^{n} \theta_{i(d)}}, \quad \hat{\eta}_{c_{jt}|u}^{(j,t)} = \frac{\sum_{i=1}^{n} \theta_{i(u,c_{jt})}}{\sum_{i=1}^{n} \theta_{i(u)}}, \quad \hat{\delta}_d = \frac{\sum_{i=1}^{n} \theta_{i(d)}}{n} \quad (12)$$

$$\hat{\phi}_{ph|d} = \frac{\sum_{i=1}^{n} \theta_{i(d)} I(z_{ip}=h)}{\sum_{i=1}^{n} \theta_{i(d)}}, \quad \hat{\rho}_{m_j k|c_{jt}}^{(j,t)} = \frac{\sum_{i=1}^{n} \theta_{i(c_{jt})} I(y_{im_j jt}=k)}{\sum_{i=1}^{n} \theta_{i(c_{jt})}}$$

To include the covariate effects on the distribution of latent profiles, $\gamma_{u|d}$ should be rewritten as $\gamma_{u|d}(\boldsymbol{X}_i) = \exp(\boldsymbol{X}_i\boldsymbol{\beta}_{u|d}) \,/\, \sum_{s=1}^{S} \exp(\boldsymbol{X}_i\boldsymbol{\beta}_{s|d})$, and thus the estimator for $\gamma_{u|d}$ in (12) is no more available. Thus, we obtain $\boldsymbol{\beta}$ estimates by Newton-Raphson method for baseline multinomial logistic regression. Apart from the estimation problem in conventional baseline logistic regression, the first and second derivatives of log-likelihood functions were written in a function of posterior probabilities obtained from Eq.9. The first and second derivatives of observed-data log-lilkelihood in Eq. (6) are available in Appendix.

2.2 *Model diagnosis and selection*

Since the models with different number of latent classes are not in nested relationship, LRT test is not available for testing the goodness of model fit. Alternatively, we adopt AIC and BIC which are popular criteria to assess relative model fit among candidate models with different number of classes. The model with smaller AIC (or BIC) is preferred.

2.3 *Simulations*

The simulation study was designed to check whether GLCPA model properly provides the parameter estimates and their asymptotic standard errors. We generated datasets under GLCPA model, and calculated the ML estimates using the EM algorithm. 95% confidence interval for each parameters were constructed based on parameter estimates and standard errors, and the empirical coverage of the confidence intervals were calculated during 100 iterations. The standard errors of the estimates were calculated through asymptotic variance-covariace matrix, by taking the negative inverse of hessian matrix. Data was simulated to have three time stages with two latent variable and one group latent variable with two classes respectivly. Each latent variable was measured by 4 binary item response variables, and the rho-parameters were designed to be equal over time. The latent profile variable was designed to have 2-profiles structure. The number of sample size was 500 (see *Web Table 1*) and 250 (see *Web Table 2*), respectively. The simulation results can be found in Appendix A.

3. Application to NLSY 97 Data

3.1 *Data description*

The National Longitudinal Survey on Youth 97 (NLSY 97) Cohort is a longitudinal project that tracks the lives of a sample of American youth born between $1980 - 84$, and 8,984 respondents were first interviewed in 1997, ages from $14 \sim 17$. Five items were adopted for measuring substance use behaviors, alcohol consumption behaviors, and depression symptoms respectively. Response variables related with *Depression* were collected in 2000 when respondents are $17 \sim 20$, and the responses for substance use and depression were collected on 2000, 2002, and 2004.

To measure *Depression* latent class variable, we select the following five survey questions: (a) How often respondent has been a nervous person in past month? (b) How often respondent felt calm and peaceful in past month? (c) How often respondent felt down and blue in past month? (d) How often respondent has been a happy person in past month? and (e) How often respondent depressed in last month? Response variable (b) and (d) were re-coded so they can be consistent in the manner that the higher response values implies more exposure to depression symptoms. In this way, we define the each binary manifest item indicating whether the respondent had suffered that feeling at least one time or not, as *Nervous*, *Not calm*, *Down*, *Not happy*, and *Depressed*, respectively.

For *Alcohol Use* latent class variable, the following three survey items are selected and re-coded: (a) Number of days respondent drink alcohol last 30 days? (b) Number of days respondent had 5 or more drinks per day last 30 days? (c) Number of days drink at schools or work per day last 30 days? The quantitative question (a) was used for creating two binary manifest items whether one had ever drunken alcohol in last 30 days (*CurrentDRK*), whether had ever drunken 5 and more days (*FrequentDRK*), and whether had ever drunken 20 and more days (*HeavyDRK*). Questionaire (b) and (c) were transformed into binary variable, having 'Yes' if its value is higher than 0, 'No' otherwise.

Similarly, the quantitative question for smoking and marijuana use behaviors were transformed into two binary items whether one had ever smoked in last 30 days (*CurrentSMK*), whether had ever smoked in daily manner for last 30 days (*FrequentSMK*), and whether had ever tried marijuana 20 or more cigarettes per day in last 30 days (*HeavySMK*). Finally, the variable '*CurrentMari*' was 'Yes' if one had ever smoked in last 30 days, and '*FrequentMari*' was assigned to be 'Yes' if one used marijuana more than 5 times in last 30 days. Table 4 shows the percentages of respondents who responded 'yes' to the 15 binary response variables, and the proportion of the non-responses.

[Table 1 about here.]

By introducing the GLCPA approach to the substance use and depression measurement items, we expect to study the following properties of the population: (a) what kinds of latent classes may be found for alcohol use, substance use behavior, and depression symptom?: (b) what kinds of common sequential patterns of alcohol and substance use behavior can be identified?: (c) how does the prevalence of latent profiles of alcohol and substance use behavior change as the latent group membership of depression symptom is varied?

3.2 *Model selection*

Web Table 3 shows the goodness-of-fit statistics with the different number of classes for each latent variable. Both AIC and BIC selected the 4-class model for *Substance Use*, 3-class model for *Depression*, and 3-class model for *Alcohol Use*.

Web Table 4 shows the list of AIC and BIC values from GLCPA models whose number of latent profile are varied from 2 to 6. BIC showed the lowest value in 5-class model. Since the class interpretations for fourth and fifth profile was obscure, we adopted the 4-latent profile structure as our final model.

Given the selected latent structure, we tested whether the primary measurement parameters can be equal across the time stages. This homogeneity assumption for ρ-parameter is critical in longitudinal latent class model, because the interpretation of each identified latent classes are solely determined based on the ρ-parameter estimates, and the meaning of each latent class should be kept equal across the stages for the identification of sequential patterns. We adopted a likelihood ratio test because the model with equal ρ-parameters over time is nested in the one with no constraints.

Web Table 5 shows the LR test result for equal ρ-parameters. The null hypothesis (H_0 : ρ-parameters for each latent variables are equal across the time) was not rejected under $\alpha = 0.05$ (p-value $= 0.067$, $\chi^2 = -2(L_0 - L_{sat}) = 88.42$), and thus we set ρ-parameters to be equal across time. Such constraints on primary measurement parameters not only reduces the number of ρ-parameters from $3 \times (5 \times 3 + 5 \times 4) = 105$ to $5 \times 3 + 5 \times 4 = 35$, but also allows the each latent classes keep same interpretation for all time stage. Finally, we fitted GLCPA with covariate using Gender (Male / Female) and Race (White / Black / Others) as covariate, and obtained the estimated odds ratios to investigate their effect on identified latent profiles.

3.3 *Parameter estimates for multiple latent group variables*

Web Table 6 shows the primary measurement parameter estimates for *Alcohol Use* variable, which is the latent subgroup of population in GLCPA model. The ρ-estimates in first class are all close or equal to 0, implying that individuals in the first class are not likely exposed to alcohol use behavior so named as *'Not Drinker'*. The second class shows the high probabilitis for current drinking behavior, so named as *'Current Drinker'*. The individuals in third class are labeled as *'Heavy Drinker'*, because they have large probability of current, frequent, and binge drinking behaviors.

Web Table 7 shows the five classes of *Substance use* latent variable and their estimated ρ-parameter estimates. The first latent class has low probabilities for all items, meaning *'Not User'*. The second class can be named as *'Marijuana User'*, because it shows high probabilities for *Current Mari*. Third class has high probabilities for *'Current SMK'*, *Frequent SMK*, and *Heavy SMK* items, thus named *Heavy Smoker*. The fourth class was *'Heavy User'*, showing the high probability for all response variables.

For *Depression*, the estimated ρ-parameters for the three identified latent classes are given in *Web Table 8*. The first class has probabilities that are lower than 0.5 for all binary responses thus named as *'Not Depressed'*. The second class has high probabilities for *Nervous*, *Down*, and *Depressed* variables compared to the first sub-group, thus named as *'Middle level Depressed'*. The third class has the high probabilities for all items except *Not Happy* items, meaning *'Seriously Depressed'*.

Web Table 9 shows the estimated secondary measurement parameters (i.e., η-parameters) for each latent class membership given latent profile membership. In Profile 1, all η-parameter estimates for *Alcohol Use* and *Substance Use* show the highest probabilities for *'Not User'* for all time waves, and thus implies *'Not involved in any substance disorder'*. In Profile 2, parameter estimates for *Alcohol Use* are mainly concentrated on *'Heavy Drinker'*, and while the prevalence on *Substance Use* were mainly distributed on *'Not User'*. As a result, the observations in Profile 2 can be named as *'Heavy Alcohol Drinker'*. In Profile 3, the prevalence on *Alcohol Use* in 2000 is the highest for *'Not User'* and monotonely moved to *'Heavy User'* across 2002 and 2004. Likewise, the probabilities for *'Heavy Smoker'* in *Subtance Use* behaviors showed consistent increase from 0.532 up to 0.818. Consequently, Profile 3 can be named as *'Developing Heavy Substance User'*. On the other hand, Profile 4 identified a subgroup whose conditional probabilities for both *Alcohol Use* and *Substance Use* are distributed on *'Heavy Drinker'* and *Serious User*. Clealy, Profile 4 represents the observations who are seriously exposed to the *Alcohol Use* and *Substance Use* behavior throught the all time waves and thus can be labeled as *'Serious Substance User'*.

We fitted multinomial logistic regression model to examine the effect of individual characteristics on latent profile memberships. *Web Table 10* shows the estimated odds ratios and their 95% confidence intervals that are obtained from the coefficients of multinomial logistic regression, given the identified *Depression* levels. Profile 1 was set as the baseline category, thus the estimated parameters represents the odds ratios of belonging to the certain latent profile compared to the Profile 1. We considered gender (female was set to be baseline) and race (White was set to be baseline) as the individual covariates, and the estimated coefficients were transformed into odds ratios for interpretation. No covariate effect had significant effect on prevalence of profiles given the *Depression* membership is *'Not Depressed'*. When the *Depression* level is middle, boys were 2.41 times more likely to belong to Profile 4 compared to baseline than girls, Black and Other students were 0.178, 0.409 times less likely to be in

Profile 4 versus baseline than White students. In *'Seriously Depressed'* latent groups, male students were 4.212 times more likely to belong to Profile 4 compared to baseline than girls.

Finally, *Web Table 11* shows the γ-estimates which represent the prevalence of four latent profiles given the *Depression* class memberships discovered in *Web Table 8* and *Web Table 9*. Profile 1 was the most prevalent class (0.512) among four profiles when the *Depression* class was *'Not Depressed'*, but decreases to 0.325 as *Depression* class becomes severe level to *'Seriously Depressed'*. Profile 2 showed relatively consistent proportion throughout the all depression levels, ranging from 0.219 to 0.261. On the other hand, the Profile 3 and 4 showed the increasing trend as the level of *Depression* becomes severe, from *Not Depressed* to *Seriously Depressed*. This is a noticeable result from GLCPA model compared to other previous categorical latent models, in that the $\hat{\gamma}$ estimates provide the quantitative measures for the associations between two categorical latent variables. *Web Table 11* evidently shows that as indicivuals exposed to more severe *Depression* levels, they are likely to experience the more serious *Alcohol Use* and *Substance Use* behaviors.

4. Discussion

This article suggested a new type of latent variable model to examine the complex structure of categorical latent variables, especially in the cases that the we study for longitudinal trends of latent variables that are identified through repeated measured item variables. GLCPA can systemically specify the effect of a latent group memberhip on the probability of having a certain sequential patterns.

Through the analysis of NLSY 97 data, we found four representative sequential patterns of young adolescents who had experienced the Authoritarian parental style. The proportions of these four latent profiles were varied by the levels of depression symptoms that the individuals were exposed to. GLCPA model discovered that as the levels of depression symptoms increase, the probability of not being exposed to the any types of substance use behavior decreases, and the prevalence of the adolescents with severe levels of substance use behaviors increases.

EM algorithm is widely adopted for the parameter estimation of the finite mixture model due to the difficulties with unobservable structures. Even though it provides the stable ML estimation, the computational cost is relatively huge compared to the other estimation strategies, and the burden of computational complexity becomes even worse if the number of time stage increases. The Reculsive method discussed in Section 2.1 significantly reduced the computational complexity by skipping the calculation of redundant posterior terms from (8). For the actual simulation result, see Chang and Chung (2013) which showed the superiority of reculsive EM estimation for univariate LCPA model in time efficiency. EM algorithm also requires the appropriate initial values to guarantee the converged solution to be global maximum. To achieve global maximum, we used 100 different sets of starting values and chose the one with the highest likelihood as a final solution, which requires another huge cost of calculation and time. To avoid the difficulty of choosing appropriate initial value, the deterministic annealing EM algorithm which ensures the global maximum. See Chang and Chung (2013),Lee and Chung (2017) for more details. To this end, we have made a program for GLCPA model written in R language (version 3.3.4) which is available on request.

Supplementary Materials

All Web Appendices, Figures, and Tables referenced in this paper are available under the submitted Wep Material, *Supporting Information for (A Multivariate Latent Class Profile Analysis with Latent Group).*

References

Chang, H.-C. and Chung, H. (2013). Dealing with multiple local modalities in latent class profile analysis. *Computational Statistics & Data Analysis* **68,** 296–310.

Chung, H., Anthony, J. C., and Schafer, J. L. (2011). Latent class profile analysis: an application to stage sequential processes in early onset drinking behaviours. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **174,** 689–712.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **39,** 1–22.

Jeon, S., Lee, J., Anthony, J. C., and Chung, H. (2017). Latent class analysis for multiple discrete latent variables: A study on the association between violent behavior and drug-using behaviors. *Structural Equation Modeling: A Multidisciplinary Journal* **24,** 911–925.

Lee, J. W. and Chung, H. (2017). Latent class analysis with multiple latent group variables. *Communications for Statistical Applications and Methods* **24,** 173–191.

Received October 2007. *Revised February* 2008. *Accepted March* 2008.

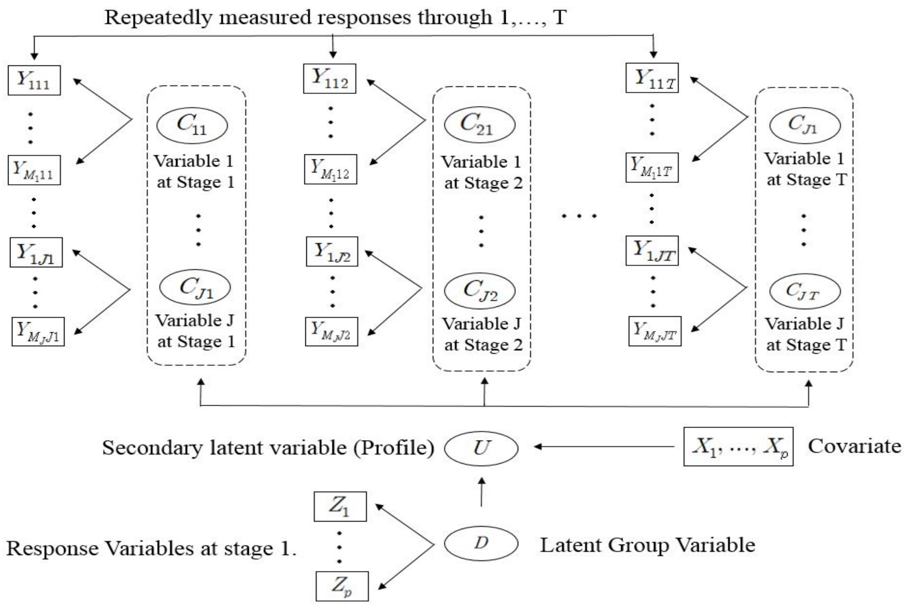

Figure 1. A diagram of LCPA with latent group variables.

Table 1

Response variables for Alcohol Use, Substance Use, *and* Depression

	Response variable	2000		2002		2004	
		Yes(%)	Mis(%)	Yes(%)	Mis(%)	Yes(%)	Mis(%)
Alcohol Use	*Current DRK*	42.51	9.48	51.04	11.11	52.21	17.04
	Frequent DRK	18.11	9.48	24.91	11.11	26.64	42.61
	Heavy DRK	4.41	9.69	4.72	11.11	2.54	17.85
	Binge DRK	24.96	9.53	28.51	11.26	30.44	18.46
	Work DRK	8.77	9.43	6.85	11.06	5.94	16.74
Substance Use	*Current SMK*	35.92	10.95	35.92	10.95	34.40	17.40
	Frequent SMK	17.19	9.43	20.59	10.95	20.64	17.40
	Heavy SMK	13.03	9.58	15.12	10.95	15.17	17.40
	Current Mari	26.13	9.94	24.10	11.36	19.08	17.19
	Frequent Mari	11.52	9.44	12.43	11.01	9.69	16.59

Depression (2000)	*Nervous*	*NotCalm*	*Feel Down*	*NotHappy*	*Depressed*
Yes(%)	62.86	6.34	69.05	2.64	37.95
Mis(%)	9.69	9.63	9.74	9.69	9.69

BIOMETRICS 64, 1–6 DOI: 10.1111/j.1541-0420.2005.00454.x
December 2008

Supporting Information for (A Multivariate Latent Class Profile Analysis with Latent Group).

This paper has been submitted for consideration for publication in *Biometrics*

1. Web Appendix

1.1 *Web Appendix A : Elements of the score function*

Let $\mathbf{\Theta}$ be a vector of all free parameters for the GLCPA. The score function $S(\mathbf{\Theta})$ is obtained by the first-ordered derivatives of the log-likelihood of the GLCPA given in observed data likelihood with respect to the model parameters $\mathbf{\Theta}$. Let $\boldsymbol{\beta}$ be the vectorized β-parameters in the GLCPA model. The elements of the first-derivative vector with respect to $\boldsymbol{\beta}$ (i.e., $\sum_{i=1}^{n} \partial \log L(\mathbf{x}_i)/\partial\boldsymbol{\beta}$) are given by

$$\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \beta_{qu|d}} = \sum_{i=1}^{n} x_{iq} \left[\theta_{i(u,d)} - \gamma_{u|d}(x_i)\theta_{i(d)}\right]$$

for $q = 1, \ldots, p$, $u = 1, \ldots, S-1$, $d = 1, \ldots, D$. Also, let $\boldsymbol{\rho}^{(j,t)}_{m_j|c_{jt}} = [\rho^{(j,t)}_{m_j1|c_{jt}}, \ldots, \rho^{(j,t)}_{m_jr_{m_j}|c_{jt}}]^T$, and $\boldsymbol{\eta}^{(j,t)}_{t|s} = [\eta^{(j,t)}_{1t|s}, \ldots, \eta^{(j,t)}_{K_jt|s}]^T$ be the vectorized ρ-and η-parameters, respectively for $m_j = 1, \ldots, M_j$, $c_{jt} = 1, \ldots, K_j$, $j = 1, \ldots, J$, $t = 1, \ldots, T$, and $s = 1, \ldots, S$. The elements of the first-derivative vector with respect to $\boldsymbol{\rho}^{(j,t)}_{m_j|c_{jt}}$ and $\boldsymbol{\eta}^{(j,t)}_{t|s}$ are obtained by

$$\begin{aligned}
\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \rho^{(j,t)}_{m_jk|c_{jt}}} &= \sum_{i=1}^{n} \frac{\theta_{i(c_{jt})}\zeta_{y_{im_jjt}k}}{\rho^{(j,t)}_{m_jk|c_{jt}}}, \\
\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \eta^{(j,t)}_{c_{jt}|u}} &= \sum_{i=1}^{n} \frac{\theta_{i(u,c_{jt})}}{\eta^{(j,t)}_{c_{jt}|u}}.
\end{aligned}$$

Here, $\zeta_{y_{im_jjt}k}$ is the indicator function which has the value of 1 if $y_{im_jjt} = k$, otherwise 0. Note that there are $r_{m_j} - 1$ and $K_j - 1$ free parameters in $\boldsymbol{\rho}^{(j,t)}_{m_j|c_{jt}}$ and $\boldsymbol{\eta}^{(j,t)}_{t|s}$, respectively. Therefore, the score function of the free parameters for $\boldsymbol{\rho}^{(j,t)}_{m_j|c_{jt}}$ and $\boldsymbol{\eta}^{(j,t)}_{t|s}$ can be obtained as follows:

$$\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \boldsymbol{\rho}^{(j,t)}_{m_j|c_{jt}}} \mathbf{A}^T_{r_{m_j}t} \text{ and } \sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \boldsymbol{\eta}^{(j,t)}_{t|s}} \mathbf{A}^T_{K_j},$$

where $\mathbf{A}_k$ is a $(k-1)\times k$ matrix, composed of an identity matrix in the first $k-1$ columns and a column vector of -1 in the last column for $m_j = 1, \ldots, M_j$, $c_{jt} = 1, \ldots, K_j$, $j = 1, \ldots, J$, $t = 1, \ldots, T$, and $s = 1, \ldots, S$. Likewise, let $\boldsymbol{\phi}_{p|d} = [\phi_{p1|d}, \ldots, \phi_{pr_p|d}]^T$, and $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_G]^T$ be the vectorized ϕ- and δ-parameters, respectively for $p = 1, \ldots, P$, and $d = 1, \ldots, G$. The elements of the first-derivative vector with respect to $\boldsymbol{\phi}_{p|d}$ and $\boldsymbol{\delta}$ are obtained as follows:

$$\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \phi_{ph|d}} = \sum_{i=1}^{n} \frac{\theta_{i(d)}\zeta_{z_{ip}h}}{\phi_{ph|d}},$$

$$\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \delta_d} = \sum_{i=1}^{n} \frac{\theta_{i(d)}}{\delta_d}.$$

Here, $\zeta_{z_{ip}h}$ is the indicator function which has the value of 1 if $z_{ip} = h$, otherwise 0. Note that there are $r_p - 1$ and $G - 1$ free parameters in $\boldsymbol{\phi}_{p|d}$ and $\boldsymbol{\delta}$, respectively. Therefore, the score function of the free parameters for $\boldsymbol{\phi}_{p|d}$ and $\boldsymbol{\delta}$ can be obtained by

$$\sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \boldsymbol{\phi}_{p|d}} \mathbf{A}_{r_p}^T \quad \text{and} \quad \sum_{i=1}^{n} \frac{\partial \log L(\mathbf{x}_i)}{\partial \boldsymbol{\delta}} \mathbf{A}_G^T,$$

where $\mathbf{A}_k$ is a $(k-1) \times k$ matrix, composed of an identity matrix in the first $k-1$ columns and a column vector of -1 in the last column for $p = 1, \ldots, P$, $d = 1, \ldots, G$.

1.2 *Web Appendix B : Elements of the Hessian Matrix*

The Hessian matrix is the second derivatives of the log-likelihood with respect to all model parameters $\boldsymbol{\Theta}$. The second derivatives of log-observed data likelihood with respect to β and $\boldsymbol{\rho}^{(j,t)}_{m_j jt|c_{jt}}$, $\boldsymbol{\eta}^{(j,t)}_u$, $\boldsymbol{\gamma}_d$ and $\boldsymbol{\phi}_{p|d}$ are obtained as follows:

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(x_i)}{\partial \beta_{qu|d} \partial \beta_{q'u'|d'}} = \sum_{i=1}^{n} x_{iq} x_{iq'} \left\{ \zeta_{dd'} [\omega_{i(u,d)} (\zeta_{uu'} - \gamma_{u|d}(x_i)) - \omega_{i(u,d')} \gamma_{u|d}(x_i)] - \omega_{i(u',d')} \omega_{i(u,d)} \right\}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(x_i)}{\partial \delta_d \partial \beta_{qu|d'}} = \sum_{i=1}^{n} \frac{x_{iq} \left\{ (\zeta_{dd'} - \theta_{i(d)}) \theta_{(u,d')} - \gamma_{u|d'}(x_i) \theta_{i(d)} \theta_{i(d')} \right\}}{\partial \delta_d}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(x_i)}{\partial \eta_{c_{jt}|u} \partial \beta_{qu'|d}} = \sum_{i=1}^{n} \frac{x_{iq} \left\{ \zeta_{uu'} \theta_{i(c_{jt},u',d)} - \theta_{i(u',d)} \theta_{i(u,c_{jt})} - \gamma_{u'|d}(x_i) (\theta_{i(c_{jt},u,d)} - \theta_{i(d)} \theta_{i(u,c_{jt})}) \right\}}{\partial \eta^{(j,t)}_{c_{jt}|u}}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(x_i)}{\partial \phi_{ph|d} \partial \beta_{qu|d'}} = \sum_{i=1}^{n} \frac{x_{iq} \left\{ \theta_{i(u,d')} (\zeta_{dd'} - \theta_{i(d)}) - \gamma_{u|d'}(x_i) \theta_{i(d)} \theta_{i(d')} \right\} \zeta_{z_{ip}h}}{\phi_{ph|d}}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(x_i)}{\partial \rho^{(j,t)}_{m_j j|c_{jt}} \partial \beta_{qu|d}} = \sum_{i=1}^{n} \frac{x_{iq} \left\{ \theta_{i(c_{jt},u,d)} - \theta_{i(u,d)} \theta_{i(c_{jt})} - \gamma_{u|d}(x_i) [\theta_{i(c_{jt},d)} - \theta_{i(c_{jt})} \theta_{i(d)}] \right\} \zeta_{y_{im_j jt}k}}{\rho_{m_j kjt|c_{jt}}}$$

where $\omega_{i(u,d)} = \theta_{i(u,d)} - \gamma_{u|d}(x_i)\theta_{i(d)}$ for $q, q' = 1, \ldots, L$, $d = 1, \ldots, G$, $\beta_{S|d} = [\beta_{1S|d}, \cdots, \beta_{LS|d}] = 0$, $u, u' = 1, \ldots, S-1$, and $\zeta_{dd'} = 1$ if $d = d'$, 0 otherwise. for $m_j = 1, \ldots, M_j$, $k = 1, \ldots, r_{m_j}$,

$c_j = 1, \ldots, K_j$, $w = 1, \ldots, D-1$, $p = 1, \ldots, r_p$, $j = 1, \ldots, J$, and $s = 1, \ldots, S$. Here, $\zeta_{y_{im_j j}k}$ is an indicator function which has the value of 1 if $y_{im_j j} = k$, 0 otherwise. Note that there are $S-1$, $r_p - 1$, $r_{m_j} - 1$, and $K_j - 1$ free parameters in $\boldsymbol{\gamma}$, $\boldsymbol{\phi}_{p|w}$, $\boldsymbol{\rho}_{m_j j|c_j}$, and $\boldsymbol{\eta}_u^{(j)}$, respectively.

The elements of the Hessian matrix with respect to γ are as follows:

$$
\begin{aligned}
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \gamma_{u|d} \partial \gamma_{u'|d'}} &= -\sum_{i=1}^{n} \frac{\theta_{i(u,d)} \theta_{i(u',d')}}{\gamma_{u|d} \gamma_{u'|d'}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \delta_d \partial \gamma_{u|d'}} &= \sum_{i=1}^{n} \frac{\left(\zeta_{dd'} - \theta_{i(d)}\right) \theta_{i(u,d')}}{\delta_d \gamma_{u|d'}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \eta^{(j,t)}_{c_{jt}|u} \partial \gamma_{u'|d}} &= \sum_{i=1}^{n} \frac{\zeta_{uu'} \theta_{i(c_{jt},u,d)} - \theta_{i(c_{jt},u)} \theta_{i(u',d)}}{\eta^{(j,t)}_{c_{jt}|u} \, \gamma_{u'|d}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \phi_{ph|d} \partial \gamma_{u|d'}} &= \sum_{i=1}^{n} \frac{\left(\zeta_{dd'} - \theta_{i(d)}\right) \theta_{i(u,d')} \zeta_{z_{ip}h}}{\phi_{ph|d} \gamma_{u|d'}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \rho^{(j,t)}_{m_j k|c_{jt}} \partial \gamma_{u|d}} &= \sum_{i=1}^{n} \frac{\left(\theta_{i(c_{jt},u,d)} - \theta_{i(c_{jt})} \theta_{i(u,d)}\right) \zeta_{y_{im_j j}k}}{\rho^{(j,t)}_{m_j k|c_{jt}} \gamma_{u|d}}
\end{aligned}
$$

for $d = 1, \ldots, G$, $p = 1, \ldots, P$, $h = 1, \ldots, r_p$, $m_j = 1, \ldots, M_j$, k, $k' = 1, \ldots r_{m_j}$, $c_{jt} = 1, \ldots, K_j$, u, $u' = 1, \ldots S$, and j, $j' = 1, \ldots, J$. Here, $\zeta_{y_{im_j j}k}$ is an indicator function which has the value of 1 if $y_{im_j j} = k$, 0 otherwise.

The elements of the Hessian matrix with respect to δ are as follows:

$$
\begin{aligned}
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \delta_d \partial \delta_{d'}} &= -\sum_{i=1}^{n} \frac{\theta_{i(d)} \theta_{i(d')}}{\delta_d \delta_{d'}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \eta^{(j,t)}_{c_{jt}|u} \partial \delta_d} &= \sum_{i=1}^{n} \frac{\theta_{i(c_{jt},u,d)} - \theta_{i(u,c_{jt})} \theta_{i(d)}}{\eta^{(j,t)}_{c_{jt}|u} \delta_d} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \phi_{ph|d} \partial \delta_{d'}} &= \sum_{i=1}^{n} \frac{(\zeta_{dd'} - \theta_{i(d)}) \theta_{i(d')} \zeta_{z_{ip}h}}{\phi_{ph|d} \delta_{d'}} \\
\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \rho^{(j,t)}_{m_j k|c_{jt}} \partial \delta_d} &= \sum_{i=1}^{n} \frac{\left(\theta_{i(d,c_{jt})} - \theta_{i(c_{jt})} \theta_{i(d)}\right) \zeta_{y_{im_j jt}k}}{\rho^{(j,t)}_{m_j k|c_{jt}} \delta_d}.
\end{aligned}
$$

for d, $d' = 1, \ldots, G$.

The elements of the Hessian matrix with respect to $\boldsymbol{\eta}$ are as follows:

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \eta^{(j,t)}_{c_{jt}|u} \partial \eta^{(j',t')}_{c'_{j't'}|u'}} = \sum_{i=1}^{n} \frac{\theta_{i(u,c_{jt},c'_{j't'})}\zeta_{uu'}(1-\zeta_{j'j}\zeta_{t't}) - \theta_{i(u,c_{jt})}\theta_{i(u',c'_{j't'})}}{\eta^{(j,t)}_{c_{jt}|u}\ \eta^{(j',t')}_{c'_{j't'}|u'}}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \phi_{ph|d} \partial \eta^{(j,t)}_{c_{jt}|u}} = \sum_{i=1}^{n} \frac{\left(\theta_{i(c_{jt},u,d)} - \theta_{i(u,c_{jt})}\theta_{i(d)}\right)\zeta_{z_{ip}h}}{\phi_{ph|d}\eta^{(j,t)}_{c_{jt}|u}}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L_i}{\partial \rho^{(j,t)}_{m_jk|c_{jt}} \partial \eta^{(j',t')}_{c'_{j't'}|u}} = \sum_{i=1}^{n} \frac{(1-\zeta_{j'j}\zeta_{t't})\,\theta_{i(u,c'_{j't'},c_{jt})} + \theta_{i(u,c'_{j't'})}\left(\zeta_{c'_{j't'}c_{jt}} - \theta_{i(c_{jt})}\right)\zeta_{y_{m_jjt}k}}{\eta^{(j',t')}_{c'_{j't'}|u}\ \rho^{(j,t)}_{m_jk|c_{jt}}}$$

for $c_{jt} = 1, \ldots, K_j$, $j = 1, \ldots, J$, $t = 1, \ldots, T$, $d = 1, \ldots, G$ and u, $u' = 1, \ldots, S$. Here, $\zeta_{j'j}$ is the indicator function whose value is 1 if $j = j'$ and 0 otherwise.

The elements of the Hessian matrix with respect to ϕ are as follows:

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \phi_{ph|d'} \partial \phi_{p'h'|d'}} = \sum_{i=1}^{n} \frac{\theta_{i(d)}\left\{(1-\zeta_{dd'}) + (1-\zeta_{pp'}) - \theta_{i(d')}\right\}\zeta_{z_{ip}h}\zeta_{z_{ip'}h'}}{\phi_{ph|d}\rho_{p'h'|d'}}$$

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \rho^{(j,t)}_{m_jk|c_{jt}} \partial \phi_{ph|d}} = \sum_{i=1}^{n} \frac{\left(\theta_{i(c_{jt},d)} - \theta_{i(c_{jt})}\theta_{i(d)}\right)\zeta_{z_{ip}h}\zeta_{y_{im_jjt}k}}{\phi_{ph|d}\rho_{m_jkjt|c_{jt}}}$$

for $p = 1, \ldots, P$, $h = 1, \ldots r_p$, $d = 1, \ldots, G$. Here, $\zeta_{z_{ip}h}$ is an indicator function which has the value of 1 if $z_{ip} = h$, otherwise 0.

The second derivatives of log-observed data likelihood with respect to $\boldsymbol{\rho}$ are as follows:

$$\sum_{i=1}^{n} \frac{\partial^2 \log L(\mathbf{x}_i)}{\partial \rho^{(j,t)}_{m_jk|c_{jt}} \partial \rho^{(j',t')}_{m'_{j'}k'|c_{j't'}}} = \sum_{i=1}^{n} \frac{\zeta_{y_{im_jjt}k}\zeta_{y_{im'_{j'}j't'}k'}}{\rho^{(j,t)}_{m_jk|c_{jt}}\ \rho^{(j',t')}_{m'_{j'}k'|c'_{j't'}}}$$
$$\times\left(\theta_{i(c_{jt},c'_{j't'})}(1-\zeta_{j'j}\zeta_{t't}) + \theta_{i(c_{jt})}\left\{\zeta_{j'j}\zeta_{t't}\left[(1-\zeta_{c_{jt}c'_{j't'}}) + (1-\zeta_{m_jm'_{j'}})\right] - \theta_{i(c'_{j't'})}\right\}\right)$$

for $k = 1, \ldots, r_{m_j}$, $m_j = 1, \ldots, M_j$, k, $k' = 1, \ldots r_{m_j}$, $c_{jt} = 1, \ldots, K_j$, t, $t' = 1, \ldots T$, and j, $j' = 1, \ldots, J$. Here, $\zeta_{y_{im_jj}k}$ is an indicator function which has the value of 1 if $y_{im_jj} = k$, 0 otherwise.

2. Web Table

2.1 *Web Table 1 : Simulation Results 1*

[Table 1 about here.]

2.2 *Web Table 2 : Simulation Results 2*

[Table 2 about here.]

Table 2.1 and 2.2 shows that the average of parameter estimates, mean square errors, and 95% coverage probabilities. The average estimates from the EM algorithm were considerably similar with the true values, and the coverage probabilities of the 95% confidence intervals are fairly close to 0.95 in both simulation. This implies that the parameter estimation and model identification are working properly.

2.3 *Web Table 3 : LCA model fit measures*

[Table 3 about here.]

2.4 *Web Table 4 : The list of AIC and BIC values from GLCPA models*

[Table 4 about here.]

2.5 *Web Table 5 : Likelihood ratio test for time constraints*

[Table 5 about here.]

2.6 *Web Table 6 : The estimated ρ-parameters for* Substance Use *classes.*

[Table 6 about here.]

2.7 *Web Table 7 : The estimated ρ-parameters for* Alcohol Use.

[Table 7 about here.]

2.8 *Web Table 8 : The estimated ρ-parameters for* Depression.

[Table 8 about here.]

2.9 *Web Table 9 :The estimated conditional probabilities of the latent class membership for a given latent profile membership (i.e., the η-parameters).*

[Table 9 about here.]

2.10 *Web Table 10 : The estimated odds ratio for a latent profile memberships given a* Depression *membership and 95% confidence intervals.*

[Table 10 about here.]

2.11 *Web Table 11 : The estimated odds ratio for a latent profile memberships given a* Depression *membership and 95% confidence intervals.*

[Table 11 about here.]

Received October 2007. *Revised February* 2008. *Accepted March* 2008.

Table 1: Average estimates (EST), mean square error (MSE), and coverage probability (CP) of 95% confidence intervals for parameter estimates (N=500).

Parameter	True	EST	MSE	CP	Parameter	True	EST	MSE	CP
$\rho^{(1,t)}_{11\|1}$	0.90	0.900	0.0001	0.94	$\eta^{(1,1)}_{1\|1}$	0.80	0.799	0.0012	0.93
$\rho^{(1,t)}_{21\|1}$	0.90	0.899	0.0001	0.95	$\eta^{(1,2)}_{1\|1}$	0.80	0.799	0.0017	0.96
$\rho^{(1,t)}_{31\|1}$	0.90	0.901	0.0001	0.95	$\eta^{(1,3)}_{1\|1}$	0.80	0.794	0.0009	0.94
$\rho^{(1,t)}_{41\|1}$	0.90	0.901	0.0001	0.98	$\eta^{(2,1)}_{1\|1}$	0.80	0.802	0.0011	0.93
$\rho^{(1,t)}_{11\|2}$	0.10	0.100	0.0002	0.94	$\eta^{(2,2)}_{1\|1}$	0.20	0.202	0.0013	0.93
$\rho^{(1,t)}_{21\|2}$	0.10	0.101	0.0001	0.94	$\eta^{(2,3)}_{1\|1}$	0.80	0.794	0.0010	0.93
$\rho^{(1,t)}_{31\|2}$	0.10	0.102	0.0001	0.97	$\eta^{(1,1)}_{1\|1}$	0.20	0.194	0.0007	0.93
$\rho^{(1,t)}_{41\|2}$	0.10	0.104	0.0001	0.93	$\eta^{(1,2)}_{1\|2}$	0.80	0.802	0.0007	0.96
$\rho^{(2,t)}_{11\|1}$	0.10	0.101	0.0002	0.93	$\eta^{(1,3)}_{1\|2}$	0.20	0.202	0.0008	0.93
$\rho^{(2,t)}_{21\|1}$	0.10	0.101	0.0002	0.94	$\eta^{(2,1)}_{1\|2}$	0.20	0.197	0.0007	0.98
$\rho^{(2,t)}_{31\|1}$	0.10	0.099	0.0002	0.93	$\eta^{(2,2)}_{1\|2}$	0.20	0.203	0.0014	0.96
$\rho^{(2,t)}_{41\|1}$	0.10	0.100	0.0002	0.98	$\eta^{(2,3)}_{1\|2}$	0.20	0.193	0.0010	0.98
$\rho^{(2,t)}_{11\|2}$	0.90	0.901	0.0002	0.99	$\phi_{11\|1}$	0.90	0.902	0.0003	0.96
$\rho^{(2,t)}_{21\|2}$	0.90	0.898	0.0002	0.97	$\phi_{21\|1}$	0.90	0.902	0.0005	0.94
$\rho^{(2,t)}_{31\|2}$	0.90	0.899	0.0001	0.95	$\phi_{31\|1}$	0.90	0.899	0.0004	0.95
$\rho^{(2,t)}_{41\|2}$	0.90	0.899	0.0001	0.96	$\phi_{41\|1}$	0.90	0.907	0.0004	0.93
$\gamma_{1\|1}$	0.20	0.202	0.0009	0.95	$\phi_{11\|2}$	0.10	0.099	0.0003	0.98
$\gamma_{1\|2}$	0.80	0.797	0.0007	0.96	$\phi_{21\|2}$	0.10	0.098	0.0003	0.93
δ_1	0.50	0.501	0.0004	0.97	$\phi_{31\|2}$	0.10	0.097	0.0004	0.96
					$\phi_{41\|2}$	0.10	0.101	0.0004	0.97

Table 2: Average estimates (EST), mean square error (MSE), and coverage probability (CP) of 95% confidence intervals for parameter estimates (N=250).

Parameter	True	EST	MSE	CP	Parameter	True	EST	MSE	CP
$\rho^{(1,t)}_{11\|1}$	0.90	0.897	0.0002	0.97	$\eta^{(1,1)}_{1\|1}$	0.80	0.812	0.0027	0.97
$\rho^{(1,t)}_{21\|1}$	0.90	0.899	0.0001	0.97	$\eta^{(1,2)}_{1\|1}$	0.80	0.799	0.0024	0.97
$\rho^{(1,t)}_{31\|1}$	0.90	0.899	0.0001	0.98	$\eta^{(1,3)}_{1\|1}$	0.80	0.797	0.0021	0.99
$\rho^{(1,t)}_{41\|1}$	0.90	0.901	0.0002	0.95	$\eta^{(2,1)}_{1\|1}$	0.80	0.802	0.0021	0.97
$\rho^{(1,t)}_{11\|2}$	0.10	0.100	0.0004	0.98	$\eta^{(2,2)}_{1\|1}$	0.20	0.201	0.0019	0.99
$\rho^{(1,t)}_{21\|2}$	0.10	0.097	0.0004	0.95	$\eta^{(2,3)}_{1\|1}$	0.80	0.793	0.0017	0.95
$\rho^{(1,t)}_{31\|2}$	0.10	0.098	0.0004	0.95	$\eta^{(1,1)}_{1\|1}$	0.20	0.201	0.0024	0.94
$\rho^{(1,t)}_{41\|2}$	0.10	0.101	0.0003	0.95	$\eta^{(1,2)}_{1\|2}$	0.80	0.808	0.0025	0.97
$\rho^{(2,t)}_{11\|1}$	0.10	0.101	0.0003	0.98	$\eta^{(1,3)}_{1\|2}$	0.20	0.202	0.0016	0.92
$\rho^{(2,t)}_{21\|1}$	0.10	0.100	0.0004	0.96	$\eta^{(2,1)}_{1\|2}$	0.20	0.199	0.0021	0.96
$\rho^{(2,t)}_{31\|1}$	0.10	0.101	0.0004	0.97	$\eta^{(2,2)}_{1\|2}$	0.20	0.202	0.0016	0.98
$\rho^{(2,t)}_{41\|1}$	0.10	0.099	0.0004	0.95	$\eta^{(2,3)}_{1\|2}$	0.20	0.198	0.0015	0.98
$\rho^{(2,t)}_{11\|2}$	0.90	0.899	0.0001	0.95	$\phi_{11\|1}$	0.90	0.898	0.0007	0.94
$\rho^{(2,t)}_{21\|2}$	0.90	0.899	0.0002	0.99	$\phi_{21\|1}$	0.90	0.899	0.0009	0.92
$\rho^{(2,t)}_{31\|2}$	0.90	0.901	0.0002	0.95	$\phi_{31\|1}$	0.90	0.898	0.0009	0.97
$\rho^{(2,t)}_{41\|2}$	0.90	0.900	0.0003	0.96	$\phi_{41\|1}$	0.90	0.899	0.0007	0.96
$\gamma_{1\|1}$	0.20	0.202	0.0019	0.95	$\phi_{11\|2}$	0.10	0.098	0.0007	0.98
$\gamma_{1\|2}$	0.80	0.797	0.0021	0.96	$\phi_{21\|2}$	0.10	0.098	0.0009	0.93
δ_1	0.50	0.498	0.0004	0.93	$\phi_{31\|2}$	0.10	0.099	0.0009	0.96
					$\phi_{41\|2}$	0.10	0.097	0.0012	0.96

Table 3: Goodness-of-fit measures for a series of LCA models with the different number of classes for each latent variables

Latent variable	Number of classes	AIC	BIC	Bootstrap p-value
Alcohol	2	18246.5	18320.1	0.00
	3	18092.4	18206.0	0.06
	4	18093.2	18247.0	0.42
	5	18105.2	18299.1	0.52
Substance Use	2	20802.1	20875.6	0.00
	3	19669.7	19783.3	0.04
	4	18931.7	19085.4	0.52
	5	18942.4	19136.3	0.54
Depression	2	7444.7	7506.1	0.00
	3	7365.1	7460.1	0.08
	4	7364.9	7493.5	0.18
	5	7367.6	7529.6	0.77

Table 4: Goodness-of-fit measures for a series of GLCPA models with the different number of latent profiles.

Number of Profiles	2	3	4	5	6
AIC	42430.9	41847.7	41505.6	41368.8	41297.0
BIC	42799.6	42317.0	42075.4	42039.2	42067.9

Table 5: LR test table for time constraints

Constraints for ρ	log-likelihood	d.f	χ^2-statistics	p-value
Equal by time	-20639.8	103	88.42	0.067
Unequal	-20595.6	173		

Table 6: The estimated ρ-parameters for *Substance Use* classes.

Response Variable	*Substance Use*			
	Not User	Marijuana User	Heavy smoker	Serious User
Current SMK	0.079	0.380	1.000†	1.000†
Frequent SMK	0.000†	0.000†	0.867	0.947
Heavy SMK	0.000†	0.000†	0.524	0.581
Current Mari	0.039	1.000†	0.112	1.000†
Frequent Mari	0.000†	0.443	0.000†	0.737

† The estimated probabilities are constrained to be zero or one.

Table 7: The estimated ρ-parameters for *Alcohol Use*.

Response Variable	*Alcohol Use*		
	Not Drinker	Current Drinker	Heavy Drinker
Current DRK	0.064	1.000†	1.000†
Frequent DRK	0.000†	0.307	0.842
Heavy DRK	0.000†	0.000†	0.158
Binge DRK	0.000†	0.222	0.954
Work DRK	0.000†	0.117	0.189

† The estimated probabilities are constrained to be zero or one.

Table 8: The estimated ρ-parameters for *Depression*.

Response Variable	*Depression*		
	Not Depressed	Middle level Depressed	Seriously Depressed
Nervous	0.413	0.846	0.925
NotCalm	0.021	0.000†	0.772
Down	0.423	0.961	0.949
NotHappy	0.000†	0.012	0.280
Depressed	0.062	0.603	0.764

† The estimated probabilities are constrained to be zero or one.

Table 9: The estimated conditional probabilities of the latent class membership for a given latent profile membership (i.e., the η-parameters).

Profile	Year	*Alcohol Use*			*Substance use*			
		Not Drinker	Current Drinker	Heavy Drinker	Not User	Marijuana User	Heavy Smoker	Serious User
1	00	0.891	0.109	0.000†	0.975	0.000†	0.025	0.000†
	02	0.795	0.205	0.000†	0.981	0.009	0.010	0.000†
	04	0.690	0.272	0.038	0.950	0.011	0.039	0.000†
2	00	0.319	0.347	0.334	0.594	0.381	0.013	0.012
	02	0.195	0.352	0.453	0.602	0.375	0.000†	0.023
	04	0.167	0.352	0.481	0.636	0.302	0.023	0.039
3	00	0.483	0.233	0.284	0.304	0.057	0.532	0.107
	02	0.335	0.279	0.386	0.098	0.000†	0.824	0.078
	04	0.339	0.223	0.438	0.146	0.000†	0.818	0.036
4	00	0.161	0.173	0.666	0.059	0.170	0.148	0.623
	02	0.125	0.201	0.674	0.035	0.137	0.158	0.670
	04	0.053	0.182	0.765	0.030	0.084	0.270	0.616

† The estimated probabilities are constrained to be zero or one.

Table 10: The estimated odds ratio for a latent profile memberships given a *Depression* membership and 95% confidence intervals.

Depression	Profile	Intercept	Male	Black	Others
Not Depressed	2	0.435 [0.263, 0.718]	1.074 [0.615, 1.875]	1.237 [0.673, 2.275]	0.881 0.377, 2.059]
	3	0.394 [0.228, 0.680]	1.279 [0.747, 2.189]	1.197 [0.657, 2.179]	0.682 [0.341, 1.363]
	4	0.186 [0.089, 0.387]	1.449 [0.609, 3.444]	0.494 [0.177, 1.374]	0.229 [0.040, 1.312]
Middle level Depressed	2	0.822 [0.583, 1.161]	1.370 [0.865, 2.169]	0.559 [0.328, 0.953]	0.580 [0.312, 1.076]
	3	0.683 [0.471, 0.989]	1.392 [0.902, 2.144]	0.738 [0.447, 1.219]	0.757 [0.452, 1.266]
	4	0.543 [0.378, 0.781]	2.410 [1.512, 3.838]	0.178 [0.082, 0.384]	0.409 [0.229, 0.729]
Seriously Depressed	2	0.686 [0.258, 1.820]	0.983 [0.310, 3.111]	1.288 [0.344, 4.815]	0.834 [0.205, 3.392]
	3	0.766 [0.305, 1.926]	0.950 [0.314, 2.880]	1.045 [0.276, 3.955]	0.902 [0.259, 3.146]
	4	0.384 [0.113, 1.291]	4.212 [1.003, 17.797]	0.154 [0.014, 1.650]	0.217 [0.036, 1.3010]

Table 11: The estimated prevalence of latent profile for a given latent group (*Depression*).

	Not Depressed	Middle level Depressed	Seriously Depressed
Profile 1	0.512	0.337	0.325
Profile 2	0.219	0.261	0.248
Profile 3	0.200	0.248	0.297
Profile 4	0.069	0.154	0.130