Bayesian Discrete Time Survival Analysis of Multivariate Reoccurrable Events: Surviving Early Depressive Moods

Chih-Chien Yang *

Graduate School of Educational Measurement & Statistics
National Taichung Teacher College

Abstract

Multivariate survival surveys in which the expected events were occurred repeatedly in a discrete time data collecting system are analyzed by the Bayesian latent transition analysis model. The extended latent transition analysis model is implemented by Markov Chain Monte Carlo (MCMC) techniques, specifically, the Gibbs sampling estimation procedure. The extended model is newly designed; therefore, it can incorporate time-invariant or -variant covariates in predicting latent survival or hazard statuses. Theoretic derivations and computational procedures, including convergence diagnosis and sensitivity evaluations are outlined and discussed. Substantively, the Bayesian latent transition analysis model is applied in the survival analysis of early depression moods among Taiwanese adolescents. Parameter estimations, model assessments, and substantive interpretations are also discussed for the empirical example.

Keywords: Multivariate Survival Analysis, Discrete Time, Latent Transition Analysis, Bayesian Statistics, Gibbs Sampling, Early Depression.

*Part of this paper was presented at an invited special session of the Fifth International Conference of the International Chinese Statistical Association (ICSA) on latent variable models that was organized by Dr. Michelle Liou from Academia Sinica, Taipei. The author thanks Dr. Wu, Chyi-In of the Institute of Sociology, Academia Sinica, Taiwan for his kindly allowing the usage of the empirical datasets. All the errors of this paper remain author's solo responsibility.

1 Introduction

Survival analysis was typically designed for repeatedly measured univariates in which the expected events are non-reoccurrable for each individual under a continuous time recording system. In contrast, latent transition analysis (LTA) proposed by this paper is particularly suitable for repeatedly recorded multivariate survival data in which the expected events are reoccurable under a discrete time data collection system. For example, in longitudinal studies of early psychological or behavioral development, the latent traits measured by multiple diagnostic items often showed great variations and had changing trends. LTA allows the assumptions of dynamic latent statuses among time-spots and reveals latent survival and hazard probabilities behind multivariate observations. Therefore, LTA has advantages over linear regression or latent growth curve like methods that often assume a stationary changing rate for the whole study period. This limited assumption can only be beneficial if a researcher would like to have an assessment on the overall trend during the whole study time; however, it also imposed an equal constraint on all unstationary changing rates. This constraint can be a potential deficiency if the rates were actually changing over time.

The LTA models will be estimated under a Bayesian framework by employing a Markov Chain Monte Carlo (MCMC) technique, i.e., Gibbs sampling estimation procedure. In addition to the novelty of LTA model applied in multivariate survival analysis, the estimation method also outlines some new modeling possibilities. Traditionally, LTA models were often fitted by using maximum likelihood estimators (MLE) in literature. Because of computational difficulties, LTA modeling method is seldom conducted by using Bayesian inference that can have advantages over the traditional estimation method. General advantages or reasons for doing a Bayesian inference can be seen in Jackman (2000), while technical reasons for Bayesian inference in latent variable or latent class models can be read in Yang, Muthen, and Yang (1999). Thanks the advancement of modern statistical algorithms and computations (see e.g., Spiegelhalter, Thomas, Best, & Gilks, 1995a & 1995b) that make the Bayesian inference of latent transition analysis more realistic. Theoretic derivations and computational implementations of Bayesian LTA using Gibbs sampling estimations are discussed in the later sections.

Adolescent surviving from the symptoms of depressive moods during a longitudinal investigation is studied by utilizing latent transition analysis within Bayesian framework. Using a similar constructing method to the depressive moods outlined in DSM-IV (American Psychiatric Association, 1994), this study abridged the original questionnaires of Wu's (1999) study and established a set of items for recording depressive mood symptoms. In particular, survival and hazard rates of experiencing early depressions were analyzed by a dynamic statistical system which explicitly modeled the changing latent hazard rates of depression statuses.

The organization of the rest of this paper is described as follows. In Section II, methodological reviews of discrete time survival analysis and latent transition analysis are presented. Section III contains mathematical derivations and statistical properties of Bayesian LTA models. Substantive Bayesian LTA applications in analyzing multivariate survival reoccurrable data are included in Section IV. Model assessments are presented in Section V. In the last section, conclusion and discussion conclude the paper.

2 Reviews

Typical discrete time survival analysis was intended for univariates in which the expected events are non-reoccurrable within each individual (e.g., Singer, & Willett, 1993). In order to analyze multivariate survival series in which events are reoccurable, latent transition analysis (LTA) is proposed to solve the problems. Substantively, statuses or traits in behavioral and psychological sciences are often not directly measurable and have to be estimated through a set of multiple diagnostic items (e.g., Yang, Muthen, & Yang, 1999). LTA solves these problems by providing predictions on a subject's consequences by analyzing its antecedents (see e.g., Collins & Wugalter, 1992; Reboussin, Liang, & Reboussin, 1999; Graham, Collins, Wugalter, Chung, & Hansen, 1991). A path diagram showed in Figure 1 illustrates the LTA model. It shows that multiple indicators (U's) define a discrete latent status (C) at a certain observation spot; furthermore, an earlier latent status (C_{e-1}) can influence the probability of its succeeding latent status (C_e).

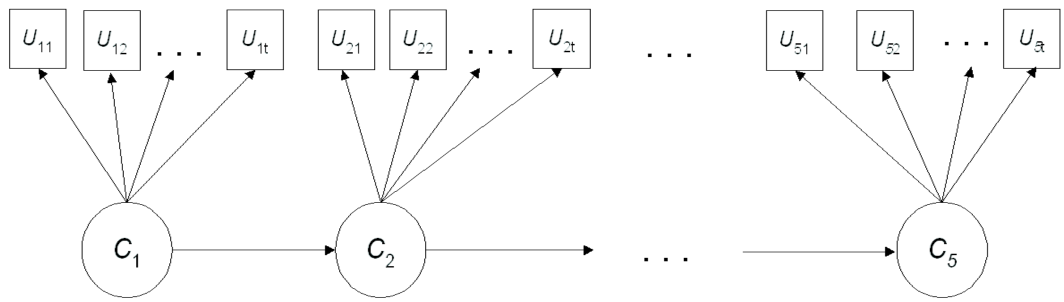

Figure 1 Path Diagram for Latent Transition Analysis

In literature, latent transition analysis was proposed as early as Wiggins (1973) in which a simple latent Markov model was created. Poulsen (1982), Van de Pol and De Leeuw (1986), Van de Pol and Langeheine (1990) and Langeheine (1994) followed Wiggins and applied LTA in sociological and behavioral sciences. Reboussin, Liang and Reboussin (1999) used an extended LTA model in biometrics while Collins and Wugalter (1992) and Graham, Collins, Wugalter, Chung and Hansen (1991) applied LTA models for psychiatric studies. Yang and Wu (2001) applied latent transition analysis for modeling deviant behaviors in a longitudinal study of Taiwanese adolescents. The popularity of LTA models in various research areas demonstrated that LTA models were substantively applicable in a wide range.

In comparing with existing literature, the current paper not only provides originality in multivariate survival models for reoccurrable latent events but also demonstrates new paradigm of Bayesian inference for LTA models. Traditionally, maximum likelihood estimator (MLE) was the most frequently selected estimation procedure for discrete time survival analysis (e.g., Singer, & Willett, 1993; Willett, & Singer, 1995) and LTA applications (e.g., Collins & Wugalter, 1992; Graham, Collins, Wugalter, Chung, & Hansen, 1991; Yang, & Wu, 2001). The usages of generalized estimating equations (GEE) were much less often seen in literature although GEE was used by Reboussin, Liang and Reboussin (1999) to handle a complicated LTA model and showed some promising results. Bayesian inferences by using various Markov chain Monte Carlo (MCMC) methods were applied in the hidden Markov model (HMM), an LTA closely related model. Example HMM applications within the framework of Bayesian inference can be seen in Robert, Celeux and Diebolt (1993) as well as in Robert, Ryden and Titterington (2000) and Chib (1996).

Performing latent transition analyses under the Bayesian framework can be difficult; nevertheless, the invention of Markov chain Monte Carlo (MCMC) methods makes the task more feasible and practical. Advantages of employing MCMC approaches as well as more intuitive explanations of MCMC can be seen in Jackman (2000). To summarize, MCMC employs a posterior sampling approach rather than a direct optimization over likelihood as often used in traditional MLE or pseudo-likelihood methods. Therefore, difficulties raised because of integrations on complex likelihood functions in Bayesian inferences can be avoided. Gibbs sampling has one further step of simplification over other MCMC methods by only working with conditional distributions instead of full likelihoods. Gibbs sampling is perhaps one of the most popular MCMC methods that can successfully solve various Bayesian estimation problems.

3 LTA & ELTA for survival analysis

In this section, the statistical properties and mathematical derivations of LTA and ELTA models for survival analysis are briefly introduced. Given a discrete time point in a survival study, let $\pi_{kc} = Pr(u_{ik} = 1|C = c)$ be the conditional probability of the kth binary response $(k = 1, \ldots, t)$ being 1 given that $C = c \quad (c = 0, 1)$ where C is the latent variable to indicate survival status. The probability that the ith subject will have diagnostic outcomes $U_i = (u_{il}, \ldots, u_{it})'$ is given by

$$g(U_i) = \sum_{c=0}^{1} \lambda_c g_c(U_i),$$

where $\lambda_c = Pr(C = c)$ and $g_c(U_i) = Pr(U_i|C = c)$. Given the latent variable, conditional independence of t binary responses can be assumed. Therefore, $g_c(U_i)$ has the form $g_c(U_i) = \Pi_{k=1}^{t} \pi_{kc}^{u_{ik}}(1 - \pi_{kc})^{1-u_{ik}}$. Thus, the likelihood function for each subject in this model based on the conditional independence assumption is given by

$$g(U_i) = \sum_{c=0}^{1} \lambda_c \prod_{k=1}^{t} \pi_{kc}^{u_{ik}}(1 - \pi_{kc})^{1-u_{ik}}.$$

To construct the model by using Bayesian framework, the conditional function $f(u_{ik}|c_i)$ is set to have a Bernoulli distribution for dichotomous responses and c_i

also have a Bernoulli distribution for the 2-latent class (survived or non-survived) model. While the number of latent class can be easily extended to more than two latent classes, the 2-latent-class case is focused in order to meet the assumptions of two survival statuses. In addition, parameters of both Bernoulli distributions have Dirichlet prior distributions. The reason for choosing Dirichlet distributions is that they conjugate the likelihoods and produce more trackable marginal posteriors, which can be modeled by using Gibbs sampling approach (Spiegelhalter, Thomas, Best, & Gilks, 1995a & 1995b).

Mathematical derivations for the Bayesian LTA model are described briefly as follows. In a time point of survival series, the conditional probabilities are given as,

$$u_{iky}|c_{iy} \sim \text{Bernoulli}(\pi_{kcy}),$$

where $\pi_{kcy} \sim \text{Dirichlet}(\alpha_1, \alpha_2)$. Specifically, in the initial time point ($y = 1$),

$$c_{iy} \sim \text{Bernoulli}(\lambda_y),$$

where $\lambda_{y=1} \sim \text{Dirichlet}(\alpha_1, \alpha_2)$. The latent survival statuses of two different time points are linked by the first order Markov process. When $y > 1$, the link is,

$$\text{logit} P(c_{iy} = 1|u_{iky}, c_{iy-1}) = \beta_0 + \beta_1 c_{iy-1} \tag{1}$$

where $\tilde{\beta} \sim \text{Normal}(\mu_\beta, \sigma_\beta^2)$. Without loss of generality, the linkage between two sequentially observed spots in the survival series is denoted by a transition probability matrix. For example,

$$\tilde{\lambda} = \begin{bmatrix} \lambda_{00} & \lambda_{01} \\ \lambda_{10} & \lambda_{11} \end{bmatrix}$$

where $\lambda_{ab} = P(c_{iy} = a|c_{iy-1} = b)$. The λ_{ab} can be interpreted as the probabilities of how subjects with latent status b at time point $y-1$ transfers to latent status a at time y. Mathematically, the transition probability matrix is obtained by an inverse logit transformation from Equation 1. Similar setups were utilized in Diggle, Liang, and Zeger (1994) but for observed Markov chains.

$$\lambda_{00} = \frac{1}{1+e^{\beta_0}},$$

$$\lambda_{01} = \frac{e^{\beta_0}}{1+e^{\beta_0}},$$

$$\lambda_{10} = \frac{1}{1+e^{\beta_0+\beta_1 C_{ij-1}}}, \text{and}$$

$$\lambda_{10} = \frac{e^{\beta_0+\beta_1 C_{ij-1}}}{1+e^{\beta_0+\beta_1 C_{ij-1}}},$$

To estimate parameters in the Bayesian LTA model, core Gibbs sampling steps are outlined as follows. The Gibbs sampling procedure is initiated by assigning reasonable starting values to all the fixed parameters of prior distributions where proper but non-informative prior were chosen. Sensitivity analyses on selections of different prior distributions and starting values are presented in later model assessment section. After the initiating stage, the Gibbs sampling procedure is proceeded by iterating the following steps.

Step 1: Sample $c_{iy}^{(j)}$ from $p(c_{iy}|u_{ij}, \pi_{jcd}^{(j-1)}, \lambda^{(j-1)})$ and obtain updated parameter values.

Step 2: Sample $\lambda_{iy}^{(j)}$ from $p(\lambda|u_{ij}, \pi_{jcd}^{(j-1)}, c_{iy}^{(j-1)})$ and obtain updated parameter values.

Step 3: Sample $\pi_{jcd}^{(j)}$ from $p(\pi_{jcd}|u_{ij}, \pi_{jcd}^{(j-1)}, \lambda^{(j-1)}, c_{iy}^{(j-1)})$ and obtain updated parameter values.

Step 4: Sample $\beta^{(j)}$ from $p(\beta|c_{iy}^{(j-1)}, \lambda^{(j-1)}, c_{i(y-1)}^{(j-1)})$ and obtain updated parameter values.

Step 5: If convergence of sampled chains is achieved, stop; otherwise, go to step 1.

An extension of LTA model, called ELTA model, is proposed in this paper. With a similar Gibbs sampling estimation procedure, the ELTA is able to include the analyses of covariate effects. In particular, both time-variant and -invariant covariates can be included in this newly proposed model for multivariate survival analysis. Figure 2 shows the ELTA setups in a path diagram where LS_i represent the latent survival statuses across the survey years.

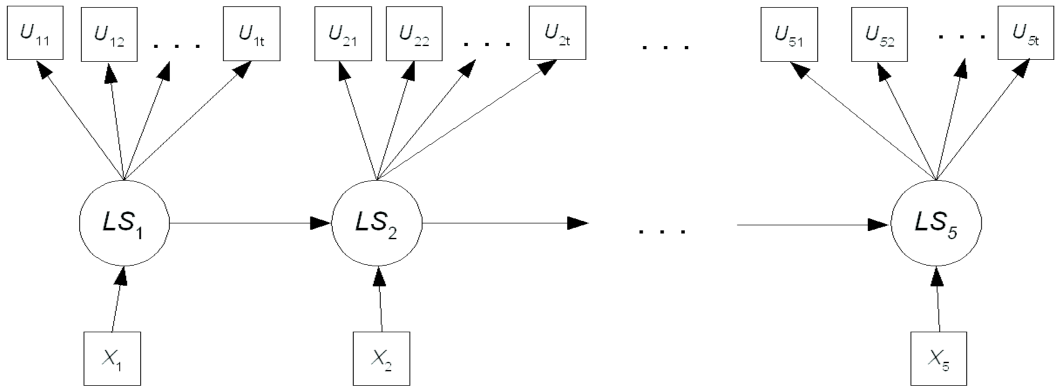

Figure 2 Path diagram for extended latent transition analysis

4 Surviving early depressions

A substantive example using Bayesian LTA method to analyze a longitudinal survey of Taiwanese adolescent depressive moods is presented. The longitudinal study is an ongoing large-scale investigation (Wu, 1999; Yang, & Wu, 2001), which has been conducted in the northern areas of Taiwan since 1997. Subjects involved in the study contain about 1,500 middle school students and their teachers and parents. In the study, several questionnaires were designed to survey important school and family related longitudinal information of these subjects. Specifically, items that surveyed adolescent depressive moods are selected to analyze probabilities of surviving the early depressive moods. Table 1 included descriptions of the diagnostic items for depressive moods. In the survey, an item is coded as 1 if a participant had a positive response on the item; otherwise, the item is coded as 0. Some descriptive statistics for these items are calculated and presented in Table 2.

Table 1 Depressive Moods Symptoms

Variable Name	Description: During the past week, did you ever feel
Lonely	lonely? (0,1)
Down	run-down? (0,1)
Worry	worried very much? (0,1)
Insomnia	sleepless or awaken all the time? (0,1)
Quarrel	easily involved a quarrel? (0,1)
Scream	desiring to scream and/or break things? (0,1)

Table 2 Descriptive Statistics for Early Depressions

	Positive Percentages (%)				
Variable	1st year	2nd year	3rd year	4th year	5th year
Lonely	18.1	22.9	27.0	14.1	40.2
Down	21.2	27.0	32.8	28.1	45.9
Worry	28.4	28.6	35.9	24.1	38.8
Insomnia	25.5	20.6	28.4	12.3	47.8
Quarrel	16.4	23.9	20.6	3.4	12.1
Scream	7.6	7.3	7.4	2.3	11.1
	Male : Female = 51.8 : 48.2 (%)				
	Total N = 873				

The early depression is fitted by the Bayesian ELTA model. The objectives of the analysis are to reveal the survival and hazard trends for depression in these adolescents. In particular, one of the strength of ELTA is to show cause-effects on the survival/hazard statuses from the covariates. In the current example, a time-invariant (gender) effect is included although time-variant effects can be easily incorporated given the proposed ELTA model setups. Figures 3 and 4 show the latent hazard and survival trends for depression moods of different genders in survey years.

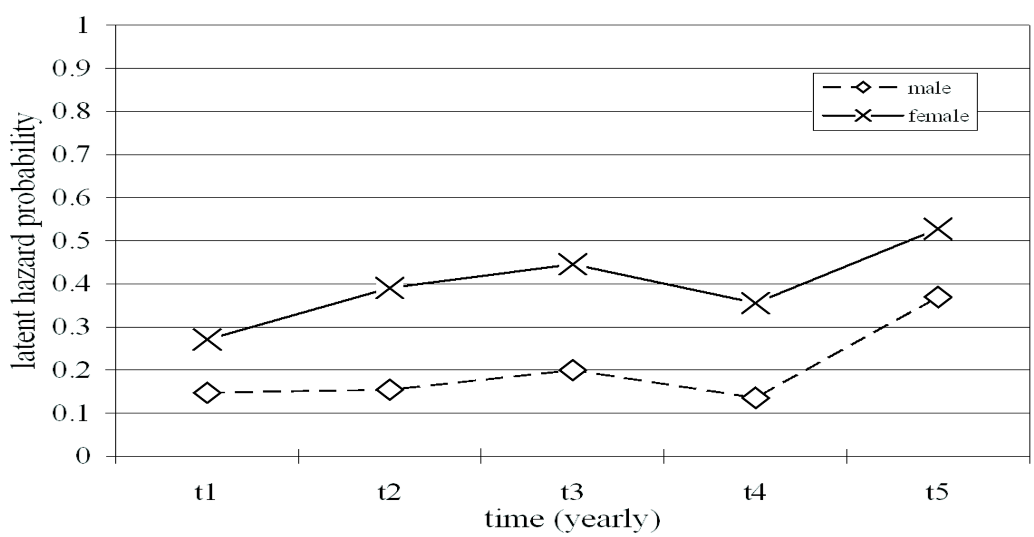

Figure 3 latent hazard probability of depression

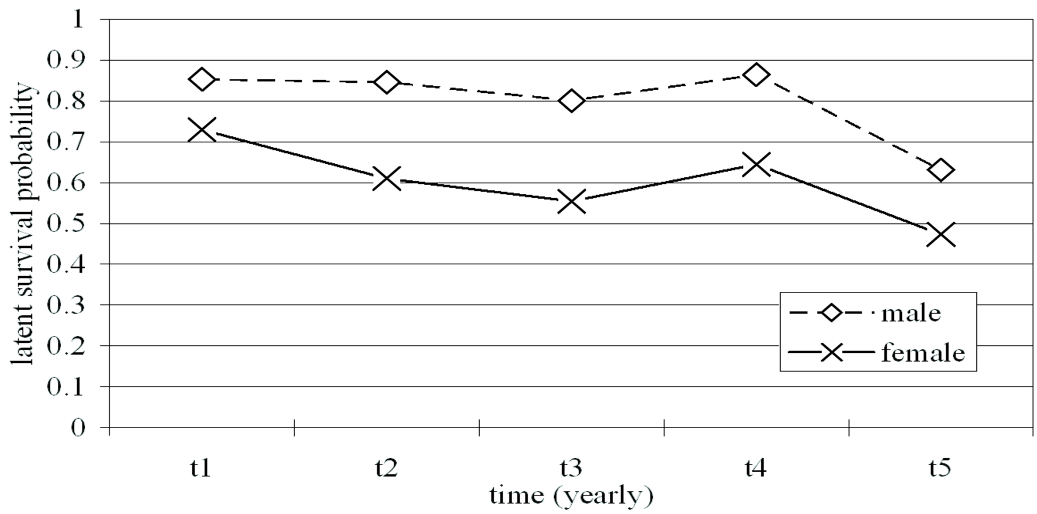

Figure 4 latent survival probability of depression

In Figure 3, the hazard rates and trends of exposures under depressions have interesting and substantive meaningful interpretations. For example, in the first three years during the survey period, hazard rates of depressions were constantly increased for both genders. The largest common event for these adolescents in Taiwan is that they have to prepare for the high school entrance exams after the end of these three years. Similarly and interestingly, there is a substantial drop of hazard rates after they entered high schools (the forth survey year), it seems a relief from the entrance exam. The depression risks, however, increased considerably again when these Taiwanese adolescents moved closer to the college entrance examinations that will take place in the end of the sixth survey year. It would not be surprising if additional studies could prove the events caused the increasing hazard rates of depression for these adolescents. The increasing hazard rates of depression, however, deserve great attentions for educators and parents of these adolescents. In addition, the increasing hazard trends of depressions for female adolescents were more obvious than their male peers in Figure 3. On other hand, the male adolescents survived more over the hits of depression moods in Figure 4. Particular cares are needed for these female adolescents because of their lower survival rates.

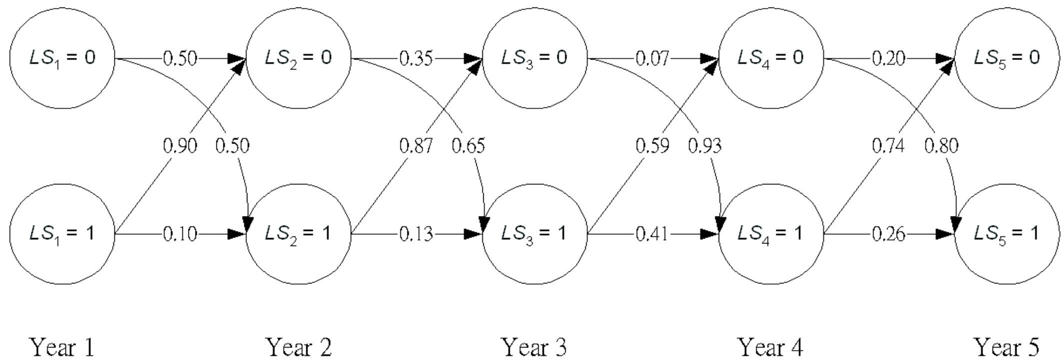

Figure 5 Latent transitions of survival probabilities (Male)

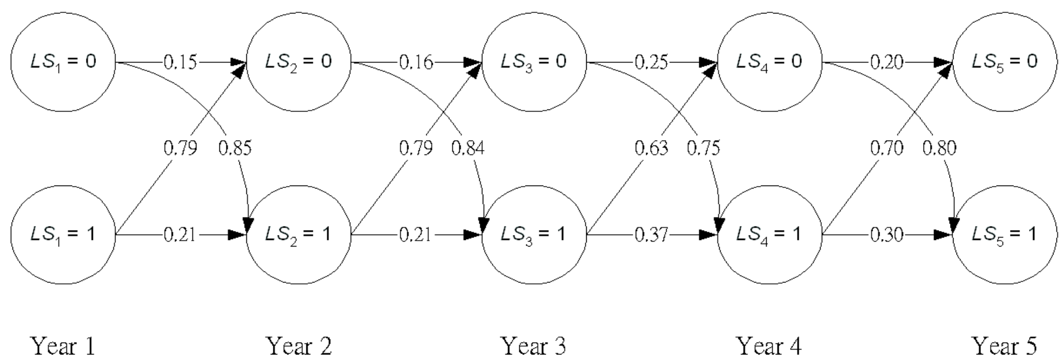

Figure 6 Latent transitions of survival probabilities (Female)

Distinguishing advantages of Bayesian LTA model for analyzing reoccurrable survival datasets include the revealing of dynamic latent hazard and survival rates. In Figures 5 and 6, the transitional probabilities between latent hazard and survival statuses are summarized for the two genders. Detailed transitional probabilities to depressed or non-depressed are provided for each year. For example, a probability of 0.84 would be expected for non-depressed ($LS_2 = 0$) female adolescents to become depressed ($LS_3 = 1$) between Year 2 and Year 3. In the same time period, however, a probability of 0.65 is expected for non-depressed ($LS_2 = 0$) male adolescents to become depressed ($LS_3 = 1$). The information is particularly useful for depression prevention and intervention programs designed for these adolescents because it provides the most efficient and valid time points for these programs to be activated.

5 Model assessment

In this section, various issues in Bayesian LTA model assessments, including convergence of Gibbs samplers and choices of different prior distributions and starting values, are discussed. Almost all MCMC methods, including Gibbs sampling, had some spaces for further research in assessing convergence of parameter estimations. In other words, there is no "the" best diagnostic method for detecting convergence of Gibbs sampling that could satisfy all experts in the field (see e.g., Kass, Carlin, Gelman, & Neal, 1998; Gelman, & Rubin, 1992). Therefore, two diagnostic methods are selected and provide cross references and comparisons. First, a non-parametric graphical method provides pioneer information for assessing convergences. In this stage, the Markov chains produced by Gibbs sampling are recorded sequentially; therefore, sampling traces and kernel densities can be plotted. Second, a parametric method, Geweke's Z-statistic (Geweke, 1992) is also used to test the convergences of each Gibbs sampling chain.

Recording Gibbs sampling traces can show how the Markov chains were sequentially generated. Thus, if a sampling procedure produced non-focused traces, i.e., a sign of non-convergence, it can be detected immediately. A trace plot with centrally condensed traces showed that the sampling procedure was likely converged. In contrast, a plot with many scattered spikes and/or wondered traces without a centriod showed no convergences. In addition, the kernel density plot of each chain can reveal its distribution shape. Convergence can be decided by whether the distribution curve meets the original parameter assumptions. All these plots for each parameter were carefully examined and they indicated that the parameter sampling procedures very likely reached their convergences. Example plots were shown in Figure 7. Notice that only part of the examined plots was shown because the number of parameters was big. Including all of them can result too many plots to fit in this presentation.

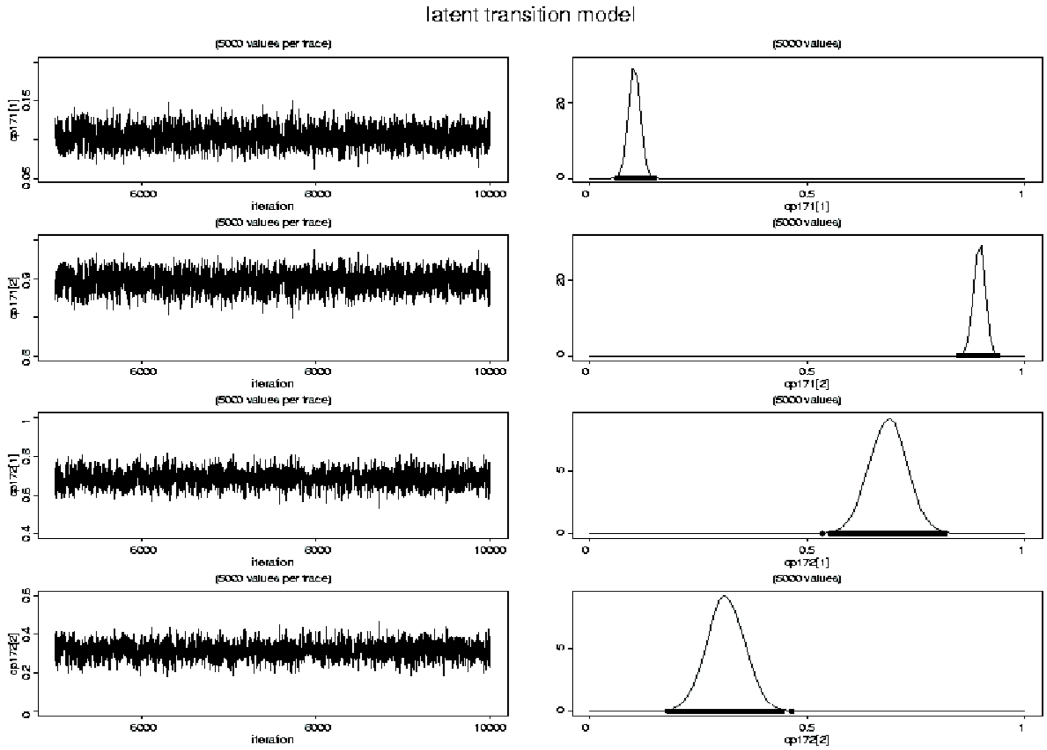

Figure 7 Latent transition model

In the parametric tests of convergence, Geweke's Z-statistic was used as the main parametric criterion for diagnosing convergence. The Geweke's Z-statistics is basically a stationary test to detect whether the comparisons of pre-half and post-half of Gibbs sampling chains reached stationary. Geweke's Z-statistics (Spiegelhalter, et al., 1995a) is described in equations as follows.

$$Z = \frac{\mu^{early} - \mu^{late}}{\sqrt{(\sum^{early} + \sum^{late})/25}}$$

where μ^{early} and μ^{late} are the means of early and late segments in a Markov chain, respectively. In addition, $\sum^{early}$ and $\sum^{late}$ represent the corresponding variances of early and late parts in the same Gibbs sampling sequence. The number of 25 is arbitrarily set for the number of segmented sequence values in each bin. The usage of Geweke's Z-statistics follows rules of regular statistical inference; i.e., it rejects the convergence of a sampling chain when its Z-statistic is larger than 1.96 given a 95Figure 8 shows the distributions of Geweke's Z-statistics of some parameters in the Bayesian ELTA model.

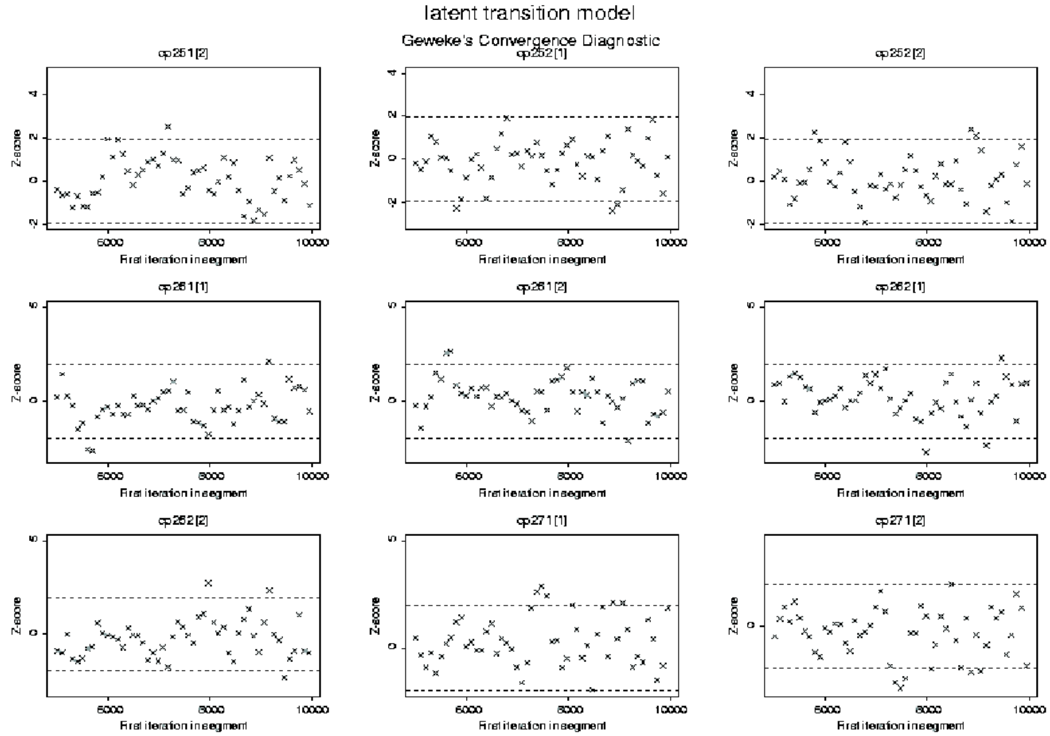

Figure 8

Although sequences in Markov chain Monte Carlo may converge eventually, it is not guaranteed that convergence will happen rapidly (Raftery, & Lewis, 1992; Yang, & Yang, 1999). To accelerate the convergence, Yang, Muthen, and Yang (1999) suggested setting a burn-in time for which sample values could be actually taken after such a period. In the current paper, we started to record samples after the 5,000th burn-in sample. To avoid autocorrelation between samples, we only saved every 5th samples after the burn-in period until we reached the 50,000th sample; therefore, there were 10,000 Gibbs sampling samples in the final estimations.

Proper and non-informative priors were selected for the parameters in Bayesian LTA models, for example, α_1 and α_2 in the Dirichlet distributions were set equal to 1. These priors were changed to various reasonable values to test the sensitivities by using these priors. In addition, different starting values were tried to test the stability of estimations. These examinations ensure the reliability and stability of parameters presented in this paper.

6 Conclusion and discussion

In this paper, a new modeling paradigm is successfully proposed for multivariate survival surveys in which survival and hazard trends of reoccurrable latent traits can be estimated. The modeling procedure is fully presented and special cares are given to the potential problematic areas that may cause instable estimations. It was shown that the Bayesian latent transition analysis is methodologically feasible and substantively applicable in the research field.

The particular trends of surviving over early depression moods were discovered in this study and they provided important information for the substantive researchers. The trends show that the gender difference on probabilities of surviving from the early depressions was considerable. In addition, the survival and hazard rates varied considerably across the survey years. This demonstrates that the early depression was not a stable or permanent psychological status for these adolescents. Further research is needed to examine other potential causes for these trends; in particular, some time variant covariates (e.g., parenting, academic performance, friendship, etc.) can also be influential to the depression statuses of these adolescents.

An interesting topic for further methodological research can be studies on evaluation methods of the model selections. Indices on the overall model fitness provide information of how appropriate the proposed models fit into the survey datasets. Yang (2004) provided studies for model selections of latent class analysis; yet, further research is needed for selection of Bayesian latent transition analysis models.

References

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, Washington, DC.

Chib, S. (1996). Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, 75, 79-97.

Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Journal of Multivariate Behavior Research*, 27, 131-157.

Diggle, P. J., Liang, K-Y, & Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457-472.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smiths (Eds.). *Bayesian Statistics 4*. Oxford: Oxford University Press.

Graham, J. W., Collins, L. M., Wugalter, S. E. Chung, N. K., & Hansen, W. B. (1991). Modelling transitions in latent stage-sequential processes: a substance use prevention example. *Journal of Consulting and Clinical Psychology*, 59, 48-57.

Jackman, S. (2000). Estimation and Inference via Bayesian Simulation: an introduction to Markov Chain Monte Carlo. *American Journal of Political Science*, 44, 375-404.

Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. M. (1998). Markov chain Monte Carlo in practice: a roundtable discussion. *The American Statistician*, 52, 93-100.

Langeheine, R. (1994). Latent variables Markov models. In A. von Eye and C. C. Clogg (eds.), *Latent Variables Analysis*, (pp. 373-395), Thousand Oaks, CA: Sage,.

Poulsen, C. A. (1982). *Latent structure analysis with choice modelling.* Arahus: Arahus School of Business Administration and Economics.

Raftery, A. E., & Lewis, S. M. (1992). How many iterations in the Gibbs sampler? In J. M. Bernardo, J.O. Berger, A. P. Dawid and A. F. M. Smiths (Eds.). *Bayesian Statistics 4.* Oxford: Oxford University Press.

Reboussin, B. A., Liang, K-Y, & Reboussin, D. M. (1999). Estimating equations for a latent transition model with multiple discrete indicators. *Biometrics*, 55, 839-845.

Robert, C. P., Celeux, G., & Diebolt, J. (1993). Bayesian estimation of hidden Markov chains: a stochastic implementation. *Statistics and Probability Letters*, 16, 77-83.

Robert, C. P., Ryden, T., & Titterington, D. M. (2000). Bayesian inference in hidden Markov models through the reversible jump Markov chain Monte Carlo method. *Journal of Royal Statistical Society*, Series B., 62, 57-75.

Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 18(2), 155-195.

Spiegelhalter, D., Thomas, A., Best, N., & Gilks, W. (1995a). *Bayesian Inference Using Gibbs Sampling (BUGS)*, Version 0.5, computer software manual, MRC Biostatistics Unit, IPH, Cambridge.

Spiegelhalter, D., Thomas, A., Best, N., & Gilks, W. (1995b). *BUGS Examples*, Volume 1-2., computer software manual, MRC Biostatistics Unit, IPH, Cambridge.

Van de Pol, F., & De Leeuw, J. (1986). A latent Markov model to correct for measurement error. *Sociological Models and Research*, 15, 118-141.

Van de Pol, F., & Langeheine, R. (1990). Mixed Markov latent class models. In C. C. Clogg (ed.), *Sociological Methodology 1990.* Oxford: Basil Blackwell.

Wiggins, L. M. (1973). *Panel Analysis.* Amsterdam: Elsevier.

Willett, J. B., & Singer, J. D. (1995). It's Deja Vu All Over Again: Using

Multiple-Spell Discrete-Time Survival Analysis. *Journal of Educational and Behavioral Statistics*, 20(1), 41-67.

Wu, C. I. (1999). The Intergenerational Transmission of Harsh Parenting in Modern Taiwan. In R. Richter and S. Supper (eds.), *New Qualities in the Life Course: Intercultural Aspects*. Germany: Ergon-Verlag.

Yang, C. C. (2004). Accuracy of Likelihood Ratio Statistics and Information Criteria in Selecting Numbers of Latent Classes, submitted.

Yang, C. C, Muthen, B., & Yang, C. C. (1999). Finite Mixture Multivariate Generalized Linear Models Using Gibbs Sampling and E-M Algorithm. Proceedings of the National Science Council Part A: *Physical Science and Engineering*, 23, 695-702.

Yang, C. C. & Wu, C. I. (2001). Latent transition models for longitudinal sociology study: A paradigm of adolescences deviant behavior development (in Chinese). *Survey Research*, 9, 5-33.

Yang, C. C., & Yang, C. C. (1999). The Application of Bayesian Extended Latent Class Analyses in Psychometrics (in Chinese), *Psychological Testing*, 46, pp. 73-84.