

MARGINAL MODEL FOR CLUSTERED CORRELATED FAILURE TIME DATA WITH AUXILIARY COVARIATES

GUO Li-sha^{1,2}, JIN Ling-hui³

(1. *School of Math. and Statistics, South-Central University for Nationalities, Wuhan 430074, China*)

(2. *School of Math. and Statistics, Wuhan University, Wuhan 430072, China*)

(3. *City College, Wuhan university of Science and Technology, Wuhan 430083, China*)

Abstract: In this article, we consider an inference procedure for clustered correlated failure time with auxiliary covariate information. We propose an estimated pseudo partial likelihood estimator under the marginal hazard model framework and develop the asymptotic properties for the proposed estimator.

Keywords: marginal hazard model; clustered failure data; auxiliary covariate; pseudopartial likelihood; validation sample

2010 MR Subject Classification: 62N01; 62N02

Document code: A

Article ID: 0255-7797(2014)06-1025-08

1 Introduction

With the continuing advancement in the use of biological markers in epidemiology and genetic studies, which often involve expensive assays, there is a growing incentive to further improve study efficiency and power by optimally incorporating into the statistical analysis the available auxiliary covariate. Some proposed methods have been developed for the univariate survival time data in the areas of mismeasured covariates, missing data, and auxiliary covariate problems. This includes, but is not limited to [1–3].

Models dealing with multivariate failure time data where the true covariates of interest are fully available for all subjects have been well studied. In particular, if the correlation among the observations is not of interest, the marginal proportional hazards model is widely used, e.g., [4–9]. There has been limited progress on the methods for dealing with covariate measurement error for multivariate failure time. Greene and Cai [10] proposed using the SIMEX approach for handling measurement errors in the marginal hazards model for multivariate failure time data, when a validation set is not available. Liu, Zhou and Cai [11] consider an inference procedure for multivariate failure time with auxiliary covariate information.

* **Received date:** 2013-11-26

Accepted date: 2014-04-02

Foundation item: Supported by the Fundamental Research Funds for the Central Universities, South-Central University for Nationalities (CZQ12015).

Biography: Guo Lisha (1980–), female, born at Wuhan, Hubei, lecturer, major in biostatistics.

Clustered failure time data arise when the study subjects are sampled in clusters so that the failure times within the same cluster tend to be correlated. In this article, assuming a validation set is available, we develop an estimated pseudopartial likelihood method for handling auxiliary covariates for clustered failure time data under the framework of the marginal hazards model with distinguishable baseline hazards.

The rest of the article is organized as follows. Section 2 outlines the marginal hazard model and present the estimated pseudopartial likelihood estimator. In Section 3, We characterize the asymptotic properties of the proposed estimator and propose a variance estimator. We conclude the article with some discussion in Section 4. Outline of the proof for theoretical results are given in the Appendix.

2 Model and Estimation

2.1 Notation and Data Structure

Suppose that there are n independent clusters. In cluster i , there are J subjects. For subject j in cluster i , K different types of failures may occur. Let (i, j, k) denote the k th type of failure on subject j in i th cluster, for $i = 1, \dots, n; j = 1, \dots, J; k = 1, \dots, K$. Let T_{ijk} and C_{ijk} denote the potential failure time and censoring time, respectively. With censoring, we observe $X_{ijk} = \min(T_{ijk}, C_{ijk})$. Let $\Delta_{ijk} = I(X_{ijk} \leq C_{ijk})$ be the failure indicator and $Y_{ijk}(t) = I(X_{ijk} \geq t)$ denote the at-risk indicator process. Let (E_{ijk}, Z_{ijk}) denote a set of covariate, where E_{ijk} is the primary exposure subjecting to missing and $Z_{ijk} = (Z_{ijk1}, \dots, Z_{ijkd})'$ is the remaining observed covariates vector that is always. We denote variable A as an auxiliary variable for the exposure variable E , assuming that conditional on E , A provides no additional information to the regression model, i.e.,

$$\lambda(t; E(t), Z(t), A(t)) = \lambda(t, E(t), Z(t)).$$

Suppose that there is a simple random validation sample with sample size n_V , denote by V , such that (i, j, k) belonging to V have their (E, A) measured. Similarly, let $\bar{V}$ denote the remaining subjects, the nonvalidation set, the subjects in $\bar{V}$ will only have their A measured. Hence, the observed data structure for (i, j, k) is

$$\begin{aligned} &\{X_{ijk}, \Delta_{ijk}, Z_{ijk}, A_{ijk}, E_{ijk}\}, \text{ if } (i, j, k) \in V, \\ &\{X_{ijk}, \Delta_{ijk}, Z_{ijk}, A_{ijk}\}, \text{ if } (i, j, k) \in \bar{V}. \end{aligned}$$

2.2 Models and Estimated Pseudopartial Likelihood Function

Assume that, the marginal hazard function for the k th failure type of subject j in cluster i takes the form

$$\lambda_{ijk}(t; Z_{ijk}(t), E_{ijk}(t)) = Y_{ijk}(t) \lambda_{0jk}(t) \exp\{\beta'_2 Z_{ijk}(t) + \beta'_1 E_{ijk}^*(t)\}, \quad (2.1)$$

where E_{ijk}^* is an m -vector consisting of E_{ijk} and possibly interaction terms between E_{ijk} and some fully observed covariates, $\beta = (\beta'_1, \beta'_2)'$ is the parameter to be estimated, and $\lambda_{0jk}(t)$ is an unspecified marginal distinct baseline hazard function pertaining to the type k failure.

If (i, j, k) belongs to the validation set, then Z_{ijk} and E_{ijk} are observed and the marginal model takes the form as in equation (2.1). If (i, j, k) belongs to the nonvalidation set $\bar{V}$, we only observe $Z_{ijk}(t)$ and $A_{ijk}(t)$. Under this situation, we can show, using the argument of Liu [11], that the hazard function for $\lambda_{ijk}(t; Z_{ijk}(t), A_{ijk}(t))$ satisfied the induced model

$$\begin{aligned}\lambda_{ijk}(t; Z_{ijk}(t), A_{ijk}(t)) &= Y_{ijk} \lambda_{0jk}(t) e^{\beta'_2 Z_{ijk}(t)} E\{e^{\beta'_1 E_{ijk}^*(t)} | Y_{ijk}(t) = 1, A_{ijk}(t), Z_{ijk}(t)\} \\ &= Y_{ijk} \lambda_{0jk}(t) e^{\beta'_2 Z_{ijk}(t)} E\{e^{\beta'_1 E_{ijk}^*(t)} | Y_{ijk}(t) = 1, A_{ijk}^*(t)\},\end{aligned}\quad (2.2)$$

where A^* includes auxiliary variable A and the part of the information in covariate Z that, given A , are still related to E . That is, A^* satisfying the following conditional dependence $f(E_{ijk}(t) | X_{ijk}(t) \geq t, Z_{ijk}(t), A_{ijk}(t)) = f(E_{ijk}(t) | X_{ijk}(t) \geq t, A_{ijk}^*(t))$. Notice that under this formulation, A^* still satisfies the auxiliary assumption that given E and Z , A^* does not contribute to the regression model, i.e., $\lambda(t; Z(t), E(t), A^*(t)) = \lambda(t; Z(t), E(t))$.

Equation (2.2) implies that this induced hazard model is also a proportional hazard model with the relative risk function $\exp(\beta'_2 Z_{ijk}(t)) \phi_{ijk}(\beta_1; t)$, where

$$\phi_{ijk}(\beta_1; t) = E\{e^{\beta'_1 E_{ijk}^*(t)} | Y_{ijk}(t) = 1, A_{ijk}^*(t)\}.$$

Based on equations (2.1) and (2.2), the relative risk function can be written as

$$r_{ijk}(\beta, t) = R_{ijk}(\beta_1, t) \exp(\beta'_2 Z_{ijk}(t)),$$

where $R_{ijk}(\beta_1, t) = \exp(\beta'_1 E_{ijk}^*(t)) \rho_{ijk} + \phi_{ijk}(\beta_1, t)(1 - \rho_{ijk})$ and the binary variable $\rho_{ijk} = 1$ or 0 denote whether (i, j, k) is in validation set V or not. If $f(E_{ijk}(t) | X_{ijk}(t) \geq t, A_{ijk}^*(t))$ is a known function up to a parameter θ , then the inference about β and θ can be drawn from a pseudopartial likelihood [4, 6]. However, misspecification of such parameterization may lead to biased estimates. We develop an estimated pseudopartial likelihood approach for clustered correlated failure time data that avoids making undesirable parametric assumptions on the conditional distribution.

If all the observations were independent, we could write the partial likelihood as

$$PPL(\beta) = \prod_{k=1}^K \prod_{j=1}^J \prod_{i=1}^n \left[\frac{r_{ijk}(\beta, X_{ijk})}{\sum_{l=1}^n Y_{ljk}(X_{ljk}) r_{ljk}(\beta, X_{ljk})} \right]^{\Delta_{ijk}}. \quad (2.3)$$

When the failure times within a subject are not independent, the above function is referred to as the pseudopartial likelihood [4, 6]. Without loss of generality, we assume that $\{A_{ijk}^*\}$ are identically distributed categorical variables with the distribution $\Pr(A^* = a_m) = p_m, m = 1, \dots, L, \sum_{m=1}^L p_m = 1$. Hence, if (i, j, k) is in the nonvalidation set $\bar{V}$, we will estimate the

induced hazard function, $\phi_{ijk}(\beta_1, t)$, as

$$\hat{\phi}_{ijk}(\beta_1, t) = \frac{\sum_{(p,q,s) \in V} Y_{pqs}(t) I(A_{pqs}^*(t) = A_{ijk}^*(t)) \exp(\beta_1' E_{pqs}^*(t))}{\sum_{(p,q,s) \in V} Y_{pqs}(t) I(A_{pqs}^*(t) = A_{ijk}^*(t))}. \quad (2.4)$$

It follows that the estimated relative risk function is

$$\hat{r}_{ijk}(\beta, t) = \hat{R}_{ijk}(\beta_1, t) \exp[\beta_2' Z_{ijk}(t)],$$

where

$$\hat{R}_{ijk}(\beta_1, t) = \exp(\beta_1' E_{ijk}^*(t)) \rho_{ijk} + \hat{\phi}_{ijk}(\beta_1, t) (1 - \rho_{ijk}).$$

Replacing $r_{ijk}(\beta, t)$ by $\hat{r}_{ijk}(\beta, t)$ in equation (2.3), we obtain an estimated pseudopartial likelihood function

$$EPPL(\beta) = \prod_{k=1}^K \prod_{j=1}^J \prod_{i=1}^n \left[\frac{\hat{r}_{ijk}(\beta, X_{ijk})}{\sum_{l=1}^n Y_{ljk}(X_{ijk}) \hat{r}_{ljk}(\beta, X_{ijk})} \right]^{\Delta_{ijk}}. \quad (2.5)$$

We define our proposed estimator $\hat{\beta}_E$ as the maximizer of equation (2.5). $\hat{\beta}_E$ can be obtained by solving the estimated pseudo partial likelihood score equation, $\hat{U}(\beta) = 0$, where

$$\hat{U}(\beta) = \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^n \int_0^\tau \frac{\hat{r}_{ijk}^{(1)}(\beta, u)}{\hat{r}_{ijk}(\beta, u)} dN_{ijk}(u) - \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^n \int_0^\tau \frac{\sum_l Y_{ljk}(u) \hat{r}_{ljk}^{(1)}(\beta, u)}{\sum_l Y_{ljk}(u) \hat{r}_{ljk}(\beta, u)} dN_{ijk}(u) \quad (2.6)$$

and $N_{ijk}(t) = I(X_{ijk} \leq t, \Delta_{ijk} = 1)$ is the counting process corresponding to failure time T_{ijk} . For a function $g(\beta, u)$, $g^{(j)}(\beta, u)$ denotes the j th derivative of $g(\beta, u)$ with respect to β . A Newton-Raphson iterative procedure can be invoked to obtain $\hat{\beta}_E$.

3 Asymptotic Properties

To investigate the asymptotic properties of the estimated pseudopartial likelihood estimator $\hat{\beta}_E$, we define the following notations. For a vector a , define $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, $a^{\otimes 2} = aa'$, $\|a\| = \sup_i |a_i|$. For a matrix A , define $\|A\| = \sup_{i,j} |a_{ij}|$. We also define

$$\begin{aligned} s_{jk}^{(0)}(\beta, t) &= E(Y_{ijk}(t) r_{ijk}(\beta, t)), s_{jk}^{(d)}(\beta, t) = E(Y_{ijk}(t) r_{ijk}^{(d)}(\beta, t)), d = 1, 2, \\ e_{1jk}(\beta, t) &= E\left(Y_{ijk}(t) \left(\frac{r_{ijk}^{(1)}(\beta, t)}{r_{ijk}(\beta, t)}\right)^{\otimes 2} r_{ijk}(\beta_0, t)\right), \\ e_{2jk}(\beta, t) &= E\left(Y_{ijk}(t) \left(\frac{r_{ijk}^{(2)}(\beta, t)}{r_{ijk}(\beta, t)}\right)^{\otimes 2} r_{ijk}(\beta_0, t)\right). \end{aligned}$$

Assume that the study duration is from 0 to τ . Suppose that $\beta_0 = (\beta_{10}', \beta_{20}')$ is the true hazards parameter. Our asymptotic results rely on the following assumptions:

- [A1] $\int_0^\tau \lambda_{0jk}(t) < \infty, j = 1, \dots, J; k = 1, \dots, K$.
- [A2] $P_r(Y_{ijk}(t) = 1 | A_{ijk}^*(t) = a_m) > 0, m = 1, \dots, L$.

[A3] For any $j = 1, \dots, J; k = 1, \dots, K$, there exists a neighborhood B_2 of β_{20} such that

$$E\left(\sup_{B_2 \times [0, \tau]} \|Z_{ijk}(t)\|^2 e^{\beta_2' Z_{ijk}(t)}\right) < \infty.$$

[A4] There exists an open set B_1 , containing β_{10} , such that $\phi_{ijk}(\beta_1, t)$ is bounded away from 0 on $B_1 \times [0, \tau]$. $\sum(\beta_0)$, as defined in Theorem 3.2, is positive definite.

[A5] For any $j = 1, \dots, J; k = 1, \dots, K$,

$$E\left(\sup_{B_1 \times [0, \tau]} [Y_{ijk}(t) R_{ijk}^{(d)}(\beta, t)]\right) < \infty, d = 0, 1, 2,$$

$$E\left(\sup_{B_1 \times [0, \tau]} \left[Y_{ijk}(t) \left\| \left(\frac{R_{ijk}^{(1)}(\beta, t)}{R_{ijk}(\beta_0, t)} \right)^{\otimes 2} \right\|^d R_{ijk}(\beta_0, t) \right]\right) < \infty, d = 1, 2,$$

$$E\left(\sup_{B_1 \times [0, \tau]} \left[Y_{ijk}(t) \left\| \frac{R_{ijk}^{(2)}(\beta, t)}{R_{ijk}(\beta, t)} \right\|^d R_{ijk}(\beta_0, t) \right]\right) < \infty, d = 1, 2.$$

[A6] $\sup_{t \in [0, \tau]} |L_k^{(d)}(t)| = O_p(1)$, $d = 0, 1$, where

$$L^{(d)}(t) = \sqrt{n_v} \left[\frac{1}{n_v} \sum_{(i,j,k) \in V} I_{(Y_{ijk}(t)=1, A_{ijk}^*=a)} \gamma_{ijk}^{(d)}(\beta_1, t) - E\left(I_{(Y_{ijk}(t)=1, A_{ijk}^*=a)} \gamma_{ijk}^{(d)}(\beta_1, t)\right) \right]$$

and $\gamma_{ijk}(\beta_1, t) = \exp(\beta_1' E_{ijk})$.

Following closely the argument of [11, 12], we can show the asymptotic properties of $\hat{\beta}_E$. We summarize the results in the following theorems and give the outline of the proofs in the Appendix.

Theorem 3.1 (Consistency) $\hat{\beta}_E$ is a consistent estimator of β_0 under assumptions (A1)–(A6).

Theorem 3.2 (Asymptotic Normality) Under the assumptions (A1)–(A6) in Appendix, we have that $n^{1/2}(\hat{\beta}_E - \beta_0)$ is asymptotically normally distributed with mean zero and variance matrix $\sum_{EPLL}(\beta_0) = \Sigma^{-1}(\beta_0) \Sigma_1(\beta_0) \Sigma^{-1}(\beta_0)$, where

$$\begin{aligned} \sum(\beta_0) &= - \int_0^\tau \sum_{k=1}^K \sum_{j=1}^J \left[\left(\frac{s_{jk}^{(1)}(\beta_0, t)}{s_{jk}^{(0)}(\beta_0, t)} \right)^{\otimes 2} s_{jk}^{(0)}(\beta_0, t) - e_{1jk}(\beta_0, t) \right] \lambda_{0jk}(t) dt, \\ \sum_1(\beta_0) &= (1 - q) E(g_{ijk}(\beta_0) g'_{ijk}(\beta_0)) + q E(h_{ijk}(\beta_0) h'_{ijk}(\beta_0)), \\ g_{ijk}(\beta_0) &= \int_0^\tau \sum_{k=1}^K \left[\left(\frac{\phi_{ijk}^{(1)}(\beta_{10}, t)}{\phi_{ijk}(\beta_{10}, t)} \right) - \frac{s_{jk}^{(1)}(\beta_0, t)}{s_{jk}^{(0)}(\beta_0, t)} \right] dM_{ijk}(t), \\ h_{ijk}(\beta_0) &= \int_0^\tau \sum_{k=1}^K \left[\left(\frac{\phi_{ijk}^{(1)}(\beta_{10}, t)}{\phi_{ijk}(\beta_{10}, t)} \right) - \frac{s_{jk}^{(1)}(\beta_0, t)}{s_{jk}^{(0)}(\beta_0, t)} \right] dM_{ijk}(t) - \frac{1 - q}{q} \begin{pmatrix} Q_{ijk}(\beta_0) \\ H_{ijk}(\beta_0) \end{pmatrix}, \\ Q_{ijk}(\beta_0) &= \int_0^\tau \sum_{k=1}^K \left(\frac{\phi_{ijk}^{(1)}(\beta_{10}, t)}{\phi_{ijk}(\beta_{10}, t)} - \frac{s_{jk}^{(11)}(\beta_0, t)}{s_{jk}^{(0)}(\beta_0, t)} \right) Y_{ijk}(t) \left(e^{\beta_{10}' E_{ijk}^*} - \phi_{ijk}(\beta_0, t) \right) \delta_{jk}^*(\beta_0, t) \lambda_{0jk}(t) dt, \\ H_{ijk}(\beta_0) &= \int_0^\tau Y_{ijk}(t) \left(e^{\beta_{10}' E_{ijk}^*} - \phi_{ijk}(\beta_0, t) \right) \delta_{jk}^{**}(\beta_0, t) \lambda_{0jk}(t) dt. \end{aligned}$$

Here $s_{jk}^{(11)}(\beta_0, t)$ is the first m elements of $s_{jk}^{(1)}(\beta_0, t)$ and $s_{jk}^{(12)}(\beta_0, t)$ is the remaining p elements,

$$\delta_{jk}^*(\beta_0, t) = E\left(e^{\beta'_{20} Z_{ijk}(t)} | Y_{ijk}(t) = 1, A_{ijk}^*(t)\right),$$

$$\delta_{jk}^{**}(\beta_0, t) = E\left(\left[Z_{ijk}(t) - \frac{s_{jk}^{(12)}(\beta_0, t)}{s_{jk}^{(0)}(\beta_0, t)}\right] e^{\beta'_{20} Z_{ijk}(t)} | Y_{ijk}(t) = 1, A_{ijk}^*(t)\right),$$

$q = \lim_{n \rightarrow \infty} (n_v/n)$, $M_{ijk}(t) = N_{ijk}(t) - \int_0^t \lambda_{ijk}(u) du$ is the marginal martingale.

The variance estimator for $\hat{\beta}_E$ can be consistently estimated by replacing the population quantities in the covariance matrix $\sum_{EPLL}(\beta_0)$ with their corresponding sample quantities. The cumulative hazard $\Lambda_{0jk}(t)$ can be estimated by Aalen-Breslow type of estimator:

$$\hat{\Lambda}_{ijk}(t) = \int_0^t \frac{\sum_{i=1}^n dN_{ijk}(s)}{\sum_{i=1}^n Y_{ijk}(s) \hat{r}_{ijk}(\hat{\beta}_E, s)} = \int_0^t \frac{1}{\hat{S}_{jk}^{(0)}(\hat{\beta}_E, s)} \frac{1}{n} \sum_{i=1}^n dN_{ijk}(s),$$

where $\hat{S}_{jk}^{(0)}(\beta, t) = n^{-1} \sum_{i=1}^n Y_{ijk} \hat{r}_{ijk}(\beta, t)$.

4 Concluding Remarks

In this article, we studied an estimated pseudopartial likelihood method for clustered failure time data with an auxiliary covariate. A key feature of this method is that it is nonparametric with respect to the association between the missing covariate and the observed auxiliary covariate. The auxiliary variable is assumed to be discrete with the number of categories fixed. One way to deal with a continuous auxiliary variable is to discretize it into categories and then apply the proposed method. Future work about common baseline hazard models and mixed baseline hazard models for clustered correlated failure time data with auxiliary covariates will be considered.

References

- [1] Lin D Y, Ying Z. Cox regression with incomplete covariate measurements [J]. J. American Statistical Association, 1993, 88: 1341–1349.
- [2] Zhou H, Pepe M S. Auxiliary covariate data in failure time regression analysis [J]. Biometrika, 1995, 82: 139–149.
- [3] Wang X, Zhou H. A semiparametric empirical likelihood method for biased sampling schemes with auxiliary covariates [J]. Biometrics, 2006, 62: 1149–1160.
- [4] Wei L J, Lin D Y, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions [J]. J. American Stat. Assoc., 1989, 84: 106–1073.
- [5] Liang K Y, Self S G, Chang Y. Modeling marginal hazards in multivariate failure time data [J]. J. Royal Statistical Society, Series B, 1993, 55: 441–453.

- [6] Cai J, Prentice R L. Estimating equations for hazard ratio parameters based on correlated failure time data [J]. Biometrika, 1995, 82: 151–164.
- [7] Cai J, Prentice R L. Regression analysis for correlated failure time data [J]. Lifetime Data Analysis, 1997, 3: 197–213.
- [8] Spiekerman C F, Lin D Y. Marginal regression models for multivariate failure time data [J]. J. American Statistical Association, 1998, 93: 1164–1175.
- [9] Clegg L X, Cai J, Sen P K. Modeling multivariate failure time data [J]. Handbook of Statistics, 2000, 18: 804–838.
- [10] Greene W F, Cai J. Measurement error in covariate in the marginal hazards model for multivariate failure time Data [J]. Biometrics, 2004, 60: 987–996.
- [11] Liu Y Y, Zhou H, Cai J. Estimated pseudo-partial-likelihood method for correlated failure time data with auxiliary covariates [J]. Biometrics, 2009, 65: 1184–1193.
- [12] Foutz R V. On the Unique consistent solution to the likelihood equations [J]. J. American Statistical Association, 1977, 72: 147–148.

带辅助协变量的分组相关失效时间数据的边际模型

郭丽莎^{1,2}, 金凌辉³

(1. 中南民族大学数学与统计学院, 湖北 武汉 430074)

(2. 武汉大学数学与统计学院, 湖北 武汉 430072)

(3. 武汉科技大学城市学院, 湖北 武汉 430083)

摘要: 本文研究了带有辅助协变量的分组相关失效时间数据的边际风险模型, 获得了未知参数的伪偏似然估计, 证明了所得估计的相合性和渐进正态性.

关键词: 边际风险模型; 分组失效数据; 辅助协变量; 伪偏似然; 有效样本

MR(2010) 主题分类号: 62N01; 62N02 中图分类号: O212.2

Appendix

In this appendix, we outline the proofs of the theorems.

Proof of Theorem 3.1 Note that $\hat{\beta}_E$ solves $n^{-1}\hat{U}(\beta) = 0$. Follow closely the argument of [12], one can show that $\hat{\beta}_E$ is consistent for β_0 , provided:

[R1] $n^{-1}\partial\hat{U}(\beta)/\partial\beta$ exists and is continuous in an open neighborhood B of β_0 .

[R2] $n^{-1}\partial\hat{U}(\beta_0)/\partial\beta_0$ is negative definite with probability going to 1.

[R3] $n^{-1}\partial\hat{U}(\beta)/\partial\beta$ converges in probability to a fixed function, $\sum(\beta)$, uniformly in an open neighborhood of β_0 .

[R4] $n^{-1}\hat{U}(\beta_0) \rightarrow 0$ in probability.

Let

$$\hat{S}_{jk}^{(d)}(\beta, t) = n^{-1} \sum_{i=1}^n Y_{ijk} \hat{r}_{ijk}^{(d)}(\beta, t), d = 0, 1, 2,$$

$$\begin{aligned}\widehat{e}_{1jk}(\beta, t) &= n^{-1} \sum_{i=1}^n Y_{ijk} \left(\frac{\widehat{r}_{ijk}^{(1)}(\beta, t)}{\widehat{r}_{ijk}(\beta, t)} \right)^{\otimes 2} r_{ijk}(\beta_0, t), \\ \widehat{e}_{2jk}(\beta, t) &= n^{-1} \sum_{i=1}^n Y_{ijk} \frac{\widehat{r}_{ijk}^{(2)}(\beta, t)}{\widehat{r}_{ijk}(\beta, t)} r_{ijk}(\beta_0, t),\end{aligned}$$

similar to [11], we can show that the four conditions are satisfied. Therefore, $\widehat{\beta}_E$ converges in probability to β_0 .

Proof of Theorem 3.2 It can be shown that the score function $n^{-1} \partial \log EPPL(\beta) / \partial \beta$ can be expressed as

$$\begin{aligned}n^{-1/2} \widehat{U}(\beta) &= n^{-1/2} \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^n \int_0^\tau \left[\frac{\widehat{r}_{ijk}^{(1)}(\beta, u)}{\widehat{r}_{ijk}(\beta, u)} - \frac{\sum_l Y_{ljk}(u) \widehat{r}_{ijk}^{(1)}(\beta, u)}{\sum_l Y_{ljk}(u) \widehat{r}_{ijk}(\beta, u)} \right] dM_{ijk}(u) \\ &\quad + n^{-1/2} \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^n \int_0^\tau \left[\frac{\widehat{r}_{ijk}^{(1)}(\beta, u)}{\widehat{r}_{ijk}(\beta, u)} - \frac{\sum_l Y_{ljk}(u) \widehat{r}_{ijk}^{(1)}(\beta, u)}{\sum_l Y_{ljk}(u) \widehat{r}_{ijk}(\beta, u)} \right] \\ &\quad r_{ijk}(\beta_0, u) Y_{ijk}(u) \lambda_{0jk}(u) du.\end{aligned}$$

By Taylor expansion of $\widehat{U}(\beta_0)$, we have

$$n^{-1/2} \widehat{U}(\beta_0) = -n^{-1} \partial \widehat{U}(\beta_*) / \partial \beta_* \cdot n^{-1/2} (\widehat{\beta}_E - \beta_0),$$

where β_* is between $\widehat{\beta}_E$ and β_0 . To prove the asymptotic normality, it suffices to prove that $n^{-1/2} \widehat{U}(\beta_0)$ converges to a normal random variable in distribution and that $n^{-1} \partial \widehat{U}(\beta_*) / \partial \beta_*$ converges to an invertible matrix. By consistency of $\widehat{\beta}_E$ and the convergence proof of $n^{-1} \partial \widehat{U}(\beta) / \partial \beta$ for (R3), it can be shown that $n^{-1} \partial \widehat{U}(\beta_*) / \partial \beta_*$ converges to the invertible matrix $\Sigma(\beta_0)$. These results together with the Slutsky Lemma give the desired normally result for $\widehat{\beta}_E$ in Theorem 3.2.