



Published in final edited form as:

Lifetime Data Anal. 2018 January ; 24(1): 94–109. doi:10.1007/s10985-017-9397-0.

Joint analysis of interval-censored failure time data and panel count data

Da Xu¹, Hui Zhao², Jianguo Sun^{1,3}

¹Center for Applied Statistical Research, School of Mathematics, Jilin University, Changchun 130012, China

²Department of Statistics, Central China Normal University, Wuhan 430079, China

³Department of Statistics, University of Missouri, Columbia, MO 65211, USA

Abstract

Interval-censored failure time data and panel count data are two types of incomplete data that commonly occur in event history studies and many methods have been developed for their analysis separately (Sun in *The statistical analysis of interval-censored failure time data*. Springer, New York, 2006; Sun and Zhao in *The statistical analysis of panel count data*. Springer, New York, 2013). Sometimes one may be interested in or need to conduct their joint analysis such as in the clinical trials with composite endpoints, for which it does not seem to exist an established approach in the literature. In this paper, a sieve maximum likelihood approach is developed for the joint analysis and in the proposed method, Bernstein polynomials are used to approximate unknown functions. The asymptotic properties of the resulting estimators are established and in particular, the proposed estimators of regression parameters are shown to be semiparametrically efficient. In addition, an extensive simulation study was conducted and the proposed method is applied to a set of real data arising from a skin cancer study.

Keywords

Bernstein polynomial; Event history study; Frailty model; Sieve maximum likelihood estimation

1 Introduction

Interval-censored failure time data and panel count data are two types of incomplete data that commonly occur in event history studies (Sun 2006; Sun and Zhao 2013). The former concerns the occurrence rate of a failure event or the event that occurs only once or only whose first occurrence is of interest, while the latter provides information on the occurrence rate of a recurrent event. Furthermore, the former means that the occurrence time of the event is observed or known only to belong to an interval, while the latter means that one only observes the numbers of the occurrences of the event between some discrete observation times. A common feature behind them is that they typically involve or occur with a periodic follow-up observation scheme. In the following, we will focus on the

situations where the two types of the incomplete data occur together and discuss their joint analysis.

There exist many fields that often produce both interval-censored failure time data and panel count data together such as economic studies, medical studies, reliability experiments and social sciences. In an economic study, for example, one may have or need to analyze the data both on the length of holding a certain credit card and on employment changes. It is easy to see that such data are usually collected from periodic follow-up studies and thus given in the forms of interval-censored data and panel count data, respectively. An example from medical studies is clinical trials with two composite prime endpoints defined by a failure time event and a recurrent event. It is well-known that clinical trials usually employ periodic follow-up schemes and thus only interval-censored data and panel count data are available.

Many authors have discussed the analysis of either interval-censored failure time data or panel count data. For example, some early references on interval-censored data are given by Turnbull (1976) and Finkelstein (1986). The former discussed nonparametric estimation of a survival function based on interval-censored data and the latter investigated regression analysis of interval-censored data under the proportional hazards model. Some more recent references on interval-censored data include Huang and Rossini (1997), Banerjee and Sen (2007) and especially, Chen et al. (2012) and Sun (2006) give relatively complete reviews of the literature on interval-censored data. The authors who have investigated the analysis of panel count data include Groeneboom and Wellner (1992), Sun and Wei (2000), Wellner and Zhang (2000), and Zhang et al. (2005). Especially, Sun and Zhao (2013) provides a relatively complete review of the literature on them, and several authors have considered sieve estimation methods for either interval-censored or panel count data separately (Hua et al. 2014; Lu et al. 2009; Ma et al. 2015; Zhang et al. 2010). However, it does not seem to exist an established approach for joint analysis of the two types of data together. In the following, we will present a sieve maximum likelihood approach for the problem based on Bernstein polynomials.

To describe the approach, we will first introduce some notation and the models as well as the resulting likelihood function in Sect. 2. Section 3 presents the proposed sieve maximum likelihood estimation procedure and in the approach, Bernstein polynomials are employed to approximate the unknown baseline mean and baseline cumulative hazard functions. In addition, the asymptotic properties of the resulting estimators are established. In Sect. 4, we present some results obtained from an extensive simulation study for the assessment of the proposed method and they suggest that it works well for practical situations. Section 5 provides an illustrative example and some discussion and concluding remarks are given in Sect. 6.

2 Notation, models and likelihood function

Consider an event history study that involves n independent subjects and two events of interest, a failure event and a recurrent event. For subject i , suppose that there exists a p -dimensional vector of covariates denoted by Z_i , and let T_i and $N_i(t)$ denote the occurrence time of the failure event and the number of the occurrences of the recurrent event up to time

t , respectively, $i = 1, 2, \dots, n$. Also suppose that each subject is observed only at a sequence of time points denoted by $s_{i1} < s_{i2} < \dots < s_{i, m_i}$, where m_i denotes the number of observations on subject i . Hence for each subject, one only observes

$$X_i = \{m_i, N_i(s_{ij}), \delta_{ik} = I(T_i \in (s_{i, k-1}, s_{i, k}]);$$

$$j = 1, 2, \dots, m_i, k = 1, 2, \dots, m_i + 1\},$$

where $s_{i0} = 0$ and $s_{i, m_i+1} = \infty$. That is, we only have interval-censored data on the T_i 's and panel count data on the $N_i(t)$'s.

To describe the effects of covariates on T_i and $N_i(t)$ and the possible correlation between T_i and $N_i(t)$, we will assume that there exists a latent variable η_i with mean 1 and unknown variance $\gamma > 0$. Suppose that given Z_i and η_i , the cumulative hazard function of T_i has the form

$$\Lambda(t|Z_i, \eta_i) = \eta_i \Lambda_1(t) e^{\alpha' Z_i}, \quad (1)$$

where Λ_1 denotes an unknown baseline cumulative hazard function and α is a vector of regression parameters. That is, T_i follows the proportional hazards frailty model. For $N_i(t)$, we will assume that it is a nonhomogeneous Poisson process with the proportional mean function

$$E\{N_i(t)|Z_i, \eta_i\} = \eta_i \Lambda_2(t) e^{\beta' Z_i}, \quad (2)$$

where Λ_2 is an unknown, nondecreasing baseline mean function and β a vector of regression parameters as α . Note that it is easy to see that η_i represents a measure of the association between T_i and $N_i(t)$ and $\eta_i = 1$ means that they are independent given covariates. In the following, we will assume that the frailty η_i is independent of $\{Z_i, T_i, N_i(t)\}$ and given Z_i and η_i , T_i and $N_i(t)$ are independent. Also it will be assumed that given Z_i , $\{m_i, s_{ij}; j = 1, 2, \dots, m_i\}$ and $\eta_i, T_i, N_i(t)$ are independent, and the conditional distribution of $\{m_i, s_{ij}; j = 1, 2, \dots, m_i\}$ given Z_i does not involve the parameters in models (1) and (2).

Define $\theta = (\alpha', \beta', \gamma, \Lambda_1, \Lambda_2)$. Then the likelihood function of θ is given by $L_n(\theta) = \prod_{i=1}^n L(\theta|X_i)$, where

$$L(\theta|X_i) = E_{\eta} \left[\prod_{k=1}^{m_i+1} \left[e^{-\eta_i \Lambda_1(s_{i,k-1}) e^{\alpha' Z_i}} - e^{-\eta_i \Lambda_1(s_{i,k}) e^{\alpha' Z_i}} \right]^{\delta_{ik}} \right. \\ \left. \times \prod_{j=1}^{m_i} \left[\eta_i \Delta \Lambda_2(s_{ij}) e^{\beta' Z_i} \right]^{\Delta N_i(s_{ij})} e^{-\eta_i \Delta \Lambda_2(s_{ij}) e^{\beta' Z_i}} \right],$$

with

$$\Delta N_i(s_{ij}) = N_i(s_{ij}) - N_i(s_{i,j-1}), \Delta N_i(s_{i1}) = N_i(s_{i1}),$$

$$\Delta \Lambda_2(s_{ij}) = \Lambda_2(s_{ij}) - \Lambda_2(s_{i,j-1}), \Delta \Lambda_2(s_{i1}) = \Lambda_2(s_{i1}).$$

Note that if the η_j 's are assumed to follow the gamma distribution, the likelihood contribution $L(\theta|X_i)$ can be simplified to

$$L(\theta|X_i) = Q_i \prod_{k=1}^{m_i+1} (A_{i,k-1} - A_{i,k})^{\delta_{ik}},$$

where

$$Q_i = \frac{\Gamma(N_i + \gamma^{-1})}{\Gamma(\gamma^{-1})} \prod_{j=1}^{m_i} \Delta \Lambda_2(s_{ij})^{\Delta N_i(s_{ij})} \left(\gamma e^{\beta' Z_i} \right)^{N_i},$$

$$A_{i0} = \left[1 + \gamma e^{\beta' Z_i} \Lambda_2(s_{i,m_i}) \right]^{-N_i - \gamma^{-1}}, A_{i,m_i+1} = 0,$$

$$A_{ik} = \left[1 + \gamma e^{\beta' Z_i} \Lambda_2(s_{i,m_i}) + \gamma e^{\alpha' Z_i} \Lambda_1(s_{ik}) \right]^{-N_i - \gamma^{-1}}, k = 1, 2, \dots, m_i$$

with $N_i = \sum_{j=1}^{m_i} \Delta N_i(s_{ij}) = N_i(s_{i,m_i})$. In the following, for the simplicity, we will focus on the situation where the η_j 's follow the gamma distribution but the method developed below applies to general situations. In the next section, we will discuss the estimation of θ .

3 Estimation and inference procedures

Now we will discuss the estimation of the parameters θ and for this, it is apparent that a natural way would be to maximize the likelihood function $L_n(\theta)$. On the other hand, it is easy to see that this would be difficult or not straightforward. Thus instead, by following Huang and Rossini (1997) and Ma et al. (2015) as well as others, we propose to employ the sieve maximum likelihood estimation approach. More specifically, define

$$\theta = \{ \theta = (\alpha', \beta', \gamma, \Lambda_1, \Lambda_2) \in \mathcal{B} \otimes \mathcal{M}^1 \otimes \mathcal{M}^2 \},$$

the parameter space of θ , and

$$\theta_n = \{ \theta_n = (\alpha', \beta', \gamma, \Lambda_{1n}, \Lambda_{2n}) \in \mathcal{B} \otimes \mathcal{M}_n^1 \otimes \mathcal{M}_n^2 \},$$

the sieve space. In the above,

$$\mathcal{B} = \{ (\alpha', \beta', \gamma) \mid (\alpha', \beta', \gamma) \in R^{2p} \times R^+, \|\alpha\| + \|\beta\| + \|\gamma\| \leq M \}$$

with M being a positive constant, \mathcal{M}^j denotes the collection of all bounded and continuous non-decreasing, non-negative functions over the interval $[c_j, u_j]$, and

$$\mathcal{M}_n^j = \left\{ \Lambda_{jn}(t) = \sum_{k=0}^m \phi_{jk} B_k(t, m, c_j, u_j) : \sum_{0 \leq k \leq m} |\phi_{jk}| \leq M_n, \quad 0 \leq \phi_{j0} \leq \phi_{j1} \leq \dots \leq \phi_{jm} \right\}$$

with

$$B_k(t, m, c_j, u_j) = \binom{m}{k} \left(\frac{t - c_j}{u_j - c_j} \right)^k \left(1 - \frac{t - c_j}{u_j - c_j} \right)^{m-k}, \quad k = 0, \dots, m,$$

the Bernstein basis polynomials of degree $m = \alpha(n^\nu)$ for some $\nu \in (0, 1)$, where $0 \leq c_j < u_j < \infty$ with $[c_j, u_j]$ usually taken as the range of observed data, $j = 1, 2$.

It is easy to see that one can approximate $\Lambda_j(t)$ by $\Lambda_{jn}(t)$ with the coefficients $\phi_{jk} = \Lambda_j(c_j + (k/m)(u_j - c_j))$ or approximate the parameter space θ by the Bernstein polynomials-based sieve space θ_n . The use of Bernstein polynomials transfers an estimation problem about both finite-dimensional and infinite-dimensional parameters into a simpler estimation problem that involves only finite-dimensional parameters. Of course, one may use other approximations such as splines and piecewise linear functions (Huang and Rossini 1997; Ding and Nan 2011). One advantage of Bernstein polynomials is that they can naturally model the nonnegativity and monotonicity of Λ_1 and Λ_2 under the constraint $0 \leq \phi_{j0} \leq \phi_{j1} \leq \dots \leq \phi_{jm}$, which can be easily removed through reparameterization in implementation (Lorentz 1986; Osman and Ghosh 2012). Also it is known that the Bernstein polynomial has the optimal shape preserving property among all approximation polynomials (Carnicer and

Peña 1993). Furthermore, although the use of Bernstein polynomials may seem to be complex, it actually can be relatively easily implemented as seen below and they do not require the specification of interior knots as spline functions. One can show that the size of the sieve space defined above can be controlled by $M_n = O(n^a)$ with a being a positive constant (Lorentz 1986; Shen 1997).

For estimation of θ , define the sieve maximum likelihood estimator $\hat{\theta}_n = (\hat{\alpha}'_n, \hat{\beta}'_n, \hat{\gamma}'_n, \hat{\Lambda}_{1n}, \hat{\Lambda}_{2n})'$ to be the value of θ that maximizes the log likelihood function $l_n(\theta) = \log \{L_n(\theta)\}$ over the sieve space Θ_n . Let $\theta_0 = (\alpha'_0, \beta'_0, \gamma'_0, \Lambda_{10}, \Lambda_{20})'$ denote the true value of θ . For the asymptotic properties of $\hat{\theta}_n$, we need the following regularity conditions.

Condition 1 The covariate Z has a bounded support in R^p .

Condition 2 If there exist constant ξ_0 and a_0 such that $\xi'_0 Z = a_0$ almost surely, then $a_0 = \xi_0 = 0$.

Condition 3 For $j = 1, 2$, the κ th derivatives of $\Lambda_{j0}(\cdot)$, denoted by $\Lambda_{j0}^{(\kappa)}(\cdot)$, is Holder continuous such that $|\Lambda_{j0}^{(\kappa)}(t_1) - \Lambda_{j0}^{(\kappa)}(t_2)| \leq \mathcal{M}|t_1 - t_2|^\eta$ for some $\eta \in (0, 1]$ and $t_1, t_2 \in [c, u]$, where $0 < c < u < \infty$ and \mathcal{M} are some constants. Define $r = \kappa + \eta$.

Condition 4 For any $\varepsilon > 0$, $\sup_{d(\theta, \theta_0) \geq \varepsilon} Pl(\theta, X) < Pl(\theta_0, X)$.

Condition 5 The matrix $E(S_\vartheta S'_\vartheta)$ is finite and positive definite with $\vartheta = (\alpha', \beta', \gamma')$, where S_ϑ is defined below.

Note that the conditions above are generally mild and satisfied in practical situations (Huang and Rossini 1997). In particular, Condition 2 is needed for the identifiability of the parameters, which is equivalently to the linear independence of the components of Z . The following theorems give the consistency, the rate of convergence and the asymptotic normality of the estimator.

Theorem 1 Suppose that the Conditions 1 – 4 described above hold. Then $\hat{\alpha}_n, \hat{\beta}_n$ and $\hat{\gamma}_n$ are strongly consistent, and as $n \rightarrow \infty$,

$$\|\hat{\Lambda}_{1n} - \Lambda_{10}\|_2 \rightarrow 0, \|\hat{\Lambda}_{2n} - \Lambda_{20}\|_2 \rightarrow 0$$

almost surely, where $\|f(X)\|_2 = (\int |f|^2 dP)^{1/2}$ is defined as norm for a function f with P being the probability measure for X .

Theorem 2 Suppose that the Conditions 1 – 4 described above hold and $r > 2$ with r defined in Condition 3. Then as $n \rightarrow \infty$, we have that

$$\|\hat{\Lambda}_{1n} - \Lambda_{10}\|_2 + \|\hat{\Lambda}_{2n} - \Lambda_{20}\|_2 = O_p\left(n^{-(1-\nu)/2} + n^{-r\nu/2}\right).$$

Theorem 3 Suppose that the Conditions 1 – 5 described above hold and $r > 2$. Then as $n \rightarrow \infty$, we have

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \rightarrow N(0, \Sigma)$$

in distribution and furthermore $\hat{\vartheta}_n$ is semiparametrically efficient, where Σ is defined in the Appendix, $\hat{\vartheta}_n = (\hat{\alpha}'_n, \hat{\beta}'_n, \hat{\gamma}_n)'$ and $\vartheta_0 = (\alpha'_0, \beta'_0, \gamma_0)'$.

The proof of the theorems above are sketched in the Appendix. To make use of the results above, it is apparent that one needs to estimate the covariance matrix of $\hat{\vartheta}_n$. One natural way would be to employ the inverse of the information matrix of the log likelihood function $l_n(\theta)$. On the other hand, this is quite difficult because of the complicated form of the information matrix. To deal with, suggested by a referee, we adopt the following simple bootstrap procedure. Let B denote a prespecified positive integer. For each b , where $1 \leq b \leq B$, draw a simple random sample of size n ,

$$D^{(b)} = \left\{ m_i^{(b)}, N_i^{(b)}(s_{ij}^{(b)}), \delta_{ik}^{(b)} = I(T_i \in (s_{i,k-1}^{(b)}, s_{i,k}^{(b)}]); j = 1, 2, \dots, m_i^{(b)}, k = 1, 2, \dots, m_i^{(b)} + 1, i = 1, 2, \dots, n \right\},$$

with replacement from the observed data

$$D = \left\{ m_i, N_i(s_{ij}), \delta_{ik} = I(T_i \in (s_{i,k-1}, s_{i,k}]); j = 1, 2, \dots, m_i, k = 1, 2, \dots, m_i + 1, i = 1, 2, \dots, n \right\}.$$

Let $\hat{\vartheta}_n^{(b)}$ denote the proposed estimate of ϑ based on the data set $D^{(b)}$ defined above. Then a natural estimate of the covariance matrix of $\hat{\vartheta}_n$ is given by

$$\hat{\Sigma} = \frac{1}{B-1} \sum_{b=1}^B \left\{ \hat{\vartheta}_n^{(b)} - \frac{1}{B} \sum_{b=1}^B \hat{\vartheta}_n^{(b)} \right\}^{\otimes 2}.$$

For the choice of positive integer B , for a practical problem, one may start with some reasonable values and then increase them until the results are stable. For example, it is common to choose $B = 100$. For a simulation study, if using enough replications, one may actually only need to use smaller values to save the computational effort.

For the implementation of the estimation procedure proposed above, two issues need to be addressed. One is that there exist some restrictions on the parameters due to the nonnegativity and monotonicity of the functions Λ_1 and Λ_2 , and for this, one can easily remove them by using some reparameterization. A natural way is to reparameterize the frailty variance parameter γ as $\exp(\gamma^*)$ and the parameters $\{\phi_0, \dots, \phi_{jm}\}$ as the cumulative

sums of $\{\exp(\phi_{j0}^*), \dots, \exp(\phi_{jm}^*)\}$, $j = 1, 2$, giving the total number parameters to be estimated being $2(p + m) + 3$. Another issue is the selection of the degree m of the Bernstein polynomials for the parameter space Θ_p , which controls the roughness or smoothness of the approximation. It is apparent that a simple approach is to use several different values that are in the order $\alpha(n^V)$ and compare the results. As an alternative, by following the BIC criterion commonly used for model selection (Burnham and Anderson 2003), one can choose the value of m that minimizes

$$\text{BIC} = -2l_n(\hat{\theta}_n) + (2(p + m) + 3)\log n.$$

4 A simulation study

An extensive simulation study was conducted to assess the performance of the estimation procedure proposed in the previous sections, including the general performance and the robustness to some assumptions. In the study, we used the design similar to that used in Hua et al. (2014) and others and generated the covariate Z_i 's and the latent variables η_i 's from the Bernoulli distribution with the probability of success 0.5 and the gamma distribution with mean 1 and variance γ , respectively. Furthermore, the total number of observation times m_i was assumed to follow the uniform distribution over $\{1, 2, 3, 4, 5\}$, and given m_i the observation times s_{ij} 's were taken to be the order statistics of the m_i random variables from the uniform distribution over $(0.02, 5)$. Then given the Z_i 's, η_i 's, m_i 's and s_{ij} 's, the failure times T_i 's were generated under model (1) with $\Lambda_1(t) = t$ and the panel count data $N_i(s_{ij})$'s under model (2) with $\Lambda_2(t) = t$ such that

$$N_i(s_{i,1}) \sim \text{Poisson}\{\eta_i \Lambda_2(s_{i,1}) \exp(\beta Z_i)\},$$

$$N_i(s_{ij}) - N_i(s_{i,j-1}) \sim \text{Poisson}\{\eta_i [\Lambda_2(s_{ij}) - \Lambda_2(s_{i,j-1})] \exp(\beta Z_i)\},$$

for $j = 2, \dots, m_i$, $i = 1, \dots, n$. The results given below are based on $n = 200$, $B = 100$ and 1000 replications.

Table 1 presents the results obtained on estimation of the parameters α , β and γ with $\alpha_0 = -0.5, 0$ or 0.5 , $\beta_0 = -0.5, 0$ or 0.5 , and $\gamma_0 = 1.2$. Here we took $[c_j, u_j] = [0.02, 5]$ and used $m = \lceil n^{1/2} \rceil = 3$, largest the integer smaller than $n^{1/4}$. The results include the estimated bias (Bias) calculated as the average of the point estimates minus the true value, the sample standard errors (SSE) of the point estimates, the average of the bootstrap standard error estimates (BSE), and the 95% empirical coverage probability (CP). They suggest that the proposed estimator seems to be unbiased and the bootstrap variance estimation also seems to be appropriate. In addition, the normal approximation to the distribution of the proposed estimators appears to be reasonable. We also considered several other values for m and obtained similar results. In other words, the proposed estimator seems to be robust to the choice of m .

Note that in the study above, we generated the latent variables η_i 's from the gamma distribution and a question of practical interest is the performance of the proposed method in the case of misspecified distributions for the latent variables. To investigate this, we repeated the simulation study above in which we generated the η_i 's from the log-normal distribution with the location parameter $\mu = -0.168$ and the scale parameter $\sigma^2 = 0.336$ but pretended they were from the gamma distribution. Table 2 gives the results obtained on estimation of the regression parameters α and β and they seem to yield similar conclusions as those given in Table 1. In other words, this suggests that the proposed estimation procedure seems to be robust to the misspecification of the distribution of the latent variables.

In the next scenario, we repeated the study that gave the results in Table 2 but generated the panel count data $N_i(s_{ij})$'s from the mixed-Poisson processes. Specifically, we first generated a random sample $\{\nu_1, \nu_2, \dots, \nu_n\}$ from $\{-0.25, 0, 0.25\}$ with $P(\nu_i = -0.25) = P(\nu_i = 0.25) = 1/4$ and $P(\nu_i = 0) = 1/2$. For each i , given ν_i , η_i and Z_i , the $\{N_i(s_{ij})\}_{j=1}^{m_i}$ were then generated from the Poisson process with the mean function $\eta_i(1 - \nu_i) \Lambda_2(t) \exp(\beta Z_i)$. The results obtained on estimation of the regression parameters are presented in Table 3 and indicate that the proposed estimation procedure seems to work well for the situations considered. In other words, the estimation procedure appears to be robust against the Poisson assumption.

5 An application

In this section, we apply the estimation procedure proposed in the previous sections to a set of real data arising from a skin cancer chemoprevention trial conducted by the University of Wisconsin Comprehensive Cancer Center in Madison, Wisconsin (Sun and Zhao 2013). It is a five-year double-blinded and placebo-controlled randomized Phase III clinical trial. The primary objective of this trial is to evaluate the effectiveness of 0.5g/m²/day PO difluoromethylornithine (DFMO) in reducing new skin cancers in a population of the patients with a history of non-melanoma skin cancers: basal cell carcinoma and squamous cell carcinoma. During the study, the patients were scheduled to be assessed or observed every 6 months for the development of new skin cancers. As expected, the real observation times differ from patient to patient and so as the follow-up times. The study consists of 291 patients randomized to either the placebo group (147) or the DFMO group (144), and the data include the numbers of occurrences of both basal cell carcinoma and squamous cell carcinoma between observation times. On the time to the first recurrence of the squamous cell carcinoma and the overall recurrence process of the basal cell carcinoma, only interval-censored data and panel count data are available, respectively. Note that both variables are important for characterizing the carcinoma recurrence process, and thus for assessing the treatment effect on the carcinoma recurrence process, one may want to consider both together.

For the study subject, in addition to the treatment indicator, we also have information on three baseline covariates, gender, age at the diagnosis and the number of prior skin cancers from the first diagnosis to randomization. For the analysis below, we will focus on the 290 patients (147 in the placebo group and 143 in the DFMO group) with at least one observation. To apply the estimation approach proposed in the previous sections, for patient

i , let T_i denote the time to the first recurrence of squamous cell carcinoma and $N_i(t)$ be the number of the basal cell carcinoma that have occurred up to time t , $i = 1, \dots, 290$. Also define the covariate Z_{i1} to be 1 if the i th patient was in the DFMO group and 0 otherwise, Z_{i2} and Z_{i3} to denote the number of prior skin cancers and the age of the patient, respectively, and $Z_{i4} = 1$ if patient i is female and 0 otherwise. Table 4 presents the analysis results given by the application of the proposed estimation procedure with $m = 3, 4, 5, 6$ and $B = 1000$, and they include the estimated covariate effects, the estimated standard error based on the bootstrap procedure (BSE) and the p -values for testing no covariate effects.

One can see from Table 4 that it seems that the DFMO treatment did not have any significant effect in reducing the risk of the first recurrence of squamous cell carcinoma and the recurrence process of basal cell carcinoma. Also both the first recurrence of squamous cell carcinoma and the recurrence process of basal cell carcinoma did not seem to be significantly related to the gender of the patient. However, both variables seem to be positively related to the number of prior skin cancers. The age of the patient seems to be positively related to the the first recurrence of squamous cell carcinoma but negatively related to the recurrence process of basal cell carcinoma. Note that the results also indicate that the time to the first squamous cell carcinoma recurrence and the overall recurrence of basal cell carcinoma were not significantly correlated. In addition, one can see that the results are similar for different values of m , which suggests that the proposed approach seems to be robust to the choice of m as expected.

6 Discussion and concluding remarks

This paper discussed joint regression analysis of interval-censored failure time data and panel count data. As mentioned above, a large literature exists for the analysis of either type of the data separately, but there does not seem to exist an established approach for their joint analysis. The proposed procedure allows one to combine two different sources together for the analysis or perform treatment comparison based on composite endpoints. The proposed method made use of the sieve approach and Bernstein polynomials and the resulting estimators of regression parameters are consistent, asymptotically normal and semiparametrically efficient. Also the simulation study suggested that it works well for practical situations.

Note that in the method described above, for the simplicity, we have assumed that the observation processes on the failure time and recurrent event process of interest are the same, but sometimes they could be different. In other words, the time observation times, say $s_{ik}^{(1)}$'s, on the T_i 's and the time observation times, say $s_{ij}^{(2)}$'s, on the $N_i(t)$'s are different. In this case, the observed data would have the form

$$\{N_i(s_{ij}^{(2)}), \delta_{ik} = I(T_i \in (s_{i,k-1}^{(1)}, s_{i,k}^{(1)}))\}$$

and more complicated structures. However, one can still follow the idea used above to develop a similar estimation procedure as above.

The proposed method relies on several assumptions or has some limitations, which also provide several directions for extensions. One is the Poisson process assumption on the underlying recurrent event process, which can be treated as a working assumption (Wellner and Zhang 2000). For this, note that the simulation study suggested that the estimation procedure is robust to it and also this has already been demonstrated theoretically by others under similar contexts. Also on the underlying recurrent event process, we only considered the situation where it follows the proportional mean model and it is apparent that it is useful to generalize the method to more general models. In the method, it has been assumed that the failure time of interest follows the proportional hazards model. Although this model is quite common, it would be helpful to develop similar methods for other commonly used models. Another assumption used above is that the observation process is noninformative and sometimes this may not be true (Sun and Zhao 2013). Note that for the latter, one faces three related processes, failure time process, the underlying recurrent event process and the observation processes, which is also a counting process. In other words, for the development of the corresponding inference method, one needs to model all three processes together.

Acknowledgements

The authors wish to thank the Editor-in-Chief, Dr. Mei-Ling Lee, an Associate Editor and two reviewers for their many helpful comments and suggestions. The work was partly supported by the National Nature Science Foundation of China grants 11471135 and 11571133 and the Central China Normal University grant MOE 15ZD011 to the second author and the NIH grant R21CA198641 to the third author.

Appendix:

Proofs of the asymptotic properties of $\hat{\theta}_n$

In this Appendix, we will sketch the proofs for the results given in Theorems 1–3 by using the empirical process theory (van der Vaart and Wellner 1996) some techniques commonly used in nonparametric literature.

First let $X = \{X_1, \dots, X_n\}$ denote the observed data and define $\mathcal{L}_n = \{l(\theta, X) : \theta \in \Theta_n\}$.

Proof of Theorem 1 For any **Proof of Theorem 1**. For any $\theta^1 = (\alpha^1, \beta^1, \gamma^1, \Lambda_1^1, \Lambda_2^1)'$ and $\theta^2 = (\alpha^2, \beta^2, \gamma^2, \Lambda_1^2, \Lambda_2^2)'$ $\in \Theta_n$, it is easy to show that

$$|l(\theta^1, X) - l(\theta^2, X)| \leq K(\|\alpha^1 - \alpha^2\| + \|\beta^1 - \beta^2\| + \|\gamma^1 - \gamma^2\| + \|\Lambda_1^1 - \Lambda_1^2\|_\infty + \|\Lambda_2^1 - \Lambda_2^2\|_\infty)$$

$$(1)$$

by the Taylor's series expansion. According to the conclusion in page 94 of van der Vaart and Wellner (1996), we can show that the covering number of \mathcal{L}_n satisfies

$$N(\epsilon, \mathcal{L}_n, L_1(P_n)) \leq N\left(\frac{\epsilon}{3M_n}, \mathcal{B}, \|\cdot\|\right) \cdot N\left(\frac{\epsilon}{3M_n}, \mathcal{M}_n^1, L_\infty\right) \cdot N\left(\frac{\epsilon}{3M_n}, \mathcal{M}_n^2, L_\infty\right) \leq \left(\frac{9M_n^2}{\epsilon}\right)^{2p+1} \\ \cdot \left(\frac{9M_n^2}{\epsilon}\right)^{m+1} \left(\frac{9M_n^2}{\epsilon}\right)^{m+1} \leq K M_n^{4p+2} M_n^{4m+4} \epsilon^{-p_m},$$

where $p_m = 2p + 2m + 3$. Then it follows from the inequality (31) of Pollard (1984)(page 31) that

$$\sup_{\theta \in \Theta_n} |P_n I(\theta, X) - P I(\theta, X)| \rightarrow 0 \quad (2)$$

in probability. Define $\Theta_\epsilon = \{\theta : d(\theta, \theta_0) < \epsilon, \theta \in \Theta_n\}$ and let $J(\theta, X) = -I(\theta, X)$, $\zeta_{1n} = \sup_{\theta \in \Theta_n} |P_n J(\theta, X) - P J(\theta, X)|$, and $\zeta_{2n} = P_n J(\theta_0, X) - P J(\theta_0, X)$. Then

$$\inf_{\Theta_\epsilon} P J(\theta, X) = \inf_{\Theta_\epsilon} \{P J(\theta, X) - P_n J(\theta, X) + P_n J(\theta, X)\} \leq \zeta_{1n} + \inf_{\Theta_\epsilon} P_n J(\theta, X).$$

If $\hat{\theta}_n \in \Theta_\epsilon$, we have

$$\inf_{\Theta_\epsilon} P_n J(\theta, X) = P_n J(\hat{\theta}_n, X) \leq P_n J(\theta_0, X) = \zeta_{2n} + P J(\theta_0, X).$$

By Condition 4, we have that $\inf_{\Theta_\epsilon} P J(\theta, X) - P J(\theta_0, X) = \delta_\epsilon > 0$. Thus,

$$\inf_{\Theta_\epsilon} P J(\theta, X) \leq \zeta_{1n} + \zeta_{2n} + P J(\theta_0, X) = \zeta_n + P J(\theta_0, X)$$

with $\zeta_n = \zeta_{1n} + \zeta_{2n}$. Hence we obtain that $\zeta_n < \delta_\epsilon$ and furthermore $\{\hat{\theta}_n \in \Theta_\epsilon\} \subseteq \{\zeta_n \geq \delta_\epsilon\}$. By (2) and Strong Law of Large Numbers, we have $\zeta_{1n} = o_p(1)$, $\zeta_{2n} = o_p(1)$ and then $\zeta_n = o_p(1)$ almost surely. Therefore, $\cup_{k=1}^\infty \cap_{n=k}^\infty \{\hat{\theta}_n \in \Theta_\epsilon\} \subseteq \cup_{k=1}^\infty \cap_{n=k}^\infty \{\zeta_n \geq \delta_\epsilon\}$, which shows the strong consistency of $\hat{\theta}_n$.

Proof of Theorem 2 To establish the convergence rate, note that by the Theorem 1.6.2 of Lorentz (1986) or the proof of Theorem 2 in Osman and Ghosh (2012), if $m = o(n^\nu)$, there

exist Bernstein polynomials A_{1n0} and A_{2n0} such that $\|A_{1n0} - A_{10}\|_\infty = O\left(n^{-\frac{N\nu}{2}}\right)$ and

$\|A_{2n0} - A_{20}\|_\infty = O\left(n^{-\frac{r\nu}{2}}\right)$, respectively. For any η , define the class

$\mathcal{F}_\eta = \{l(\theta_{n0}, X) - l(\theta, X) : \theta \in \Theta_n, d(\theta - \theta_{n0}) < \eta\}$ with $\theta_{n0} = (\beta'_0, \alpha'_0, \gamma'_0, \Lambda_{1n0}, \Lambda_{2n0})'$. Following the calculation of Shen and Wong (1994, p. 597), we can establish that $\log N_{[]}(\varepsilon, \mathcal{F}_\eta, \|\cdot\|_2) \leq CN \log(\eta/\varepsilon)$ with $N = 2(m + 1)$. Moreover, some algebraic calculations lead to $\|l(\theta_{n0}, X) - l(\theta, X)\|_2^2 \leq C\eta^2$ for any $l(\theta_{n0}, X) - l(\theta, X) \in \mathcal{F}_\eta$. Therefore it follows from Lemma 3.4.2 of van der Vaart and Wellner (1996) that

$$E_P \|n^{1/2}(P_n - P)\|_{\mathcal{F}_\eta} \leq CJ_\eta(\varepsilon, \mathcal{F}_\eta, \|\cdot\|_2) \left\{ 1 + \frac{J_\eta(\varepsilon, \mathcal{F}_\eta, \|\cdot\|_2)}{\eta^2 n^{1/2}} \right\}, \quad (3)$$

where $J_\eta(\varepsilon, \mathcal{F}_\eta, \|\cdot\|_2) = \int_0^\eta \{1 + \log N_{[]}(\varepsilon, \mathcal{F}_\eta, \|\cdot\|_2)\}^{1/2} d\varepsilon \leq CN^{1/2}\eta$.

Note that the right-hand side of (3) gives $\phi_n(\eta) = C(N^{1/2}\eta + N/n^{1/2})$. Also it is easy to see that $\phi_n(\eta)/\eta$ decreases in η and $r_n^2 \phi_n(1/r_n) = r_n N^{1/2} + r_n^2 N/n^{1/2} < 2n^{1/2}$, where $r = N^{-1/2}n^{1/2} = n^{(1-\nu)/2}$ with $0 < \nu < 0.5$. Hence $n^{(1-\nu)/2}d(\hat{\theta}_n, \theta_{n0}) = O_p(1)$ by Theorem 3.4.1 of van der Vaart and Wellner (1996). This, together with $d(\theta_{n0}, \theta_0) = O_p(n^{-r\nu/2})$, yields that $d(\hat{\theta}_n, \theta_0) = O_p(n^{-(1-\nu)/2} + n^{-r\nu/2})$. The choice of $\nu = 1/(1+r)$ yields the rate of convergence $d(\hat{\theta}_n, \theta_0) = O_p(n^{-r/(2+2r)})$.

Proof of Theorem 3 As above, let θ_0 denote the true value of parameter θ and define V to be the linear span of $\Theta - \theta_0$. Also let $l(\theta, X)$ be the log-likelihood for a sample of size one and $\delta_n = n^{-(1-\nu)/2} + n^{-r\nu/2}$. For any $\theta \in \{\theta \in \Theta : \|\theta - \theta_0\| = O(\delta_n)\}$, define the first order directional derivative of $l(\theta, X)$ at the direction $v \in V$ as

$$\dot{l}(\theta, X)[v] = \frac{dl(\theta + sv, X)}{ds} \Big|_{s=0},$$

and the second order directional derivative as

$$\ddot{l}(\theta, X)[v, \tilde{v}] = \frac{d\dot{l}(\theta + s\tilde{v}, X)}{ds} \Big|_{s=0} = \frac{d^2 l(\theta + sv + s\tilde{v}, X)}{ds ds} \Big|_{s=s=0}.$$

Also define the Fisher inner product on the space V as

$$\langle v, \tilde{v} \rangle = P\{\dot{l}(\theta, X)[v]\dot{l}(\theta, X)[\tilde{v}]\}$$

and the Fisher norm for $v \in V$ as $\|v\|^{1/2} = \langle v, v \rangle$. Let \bar{V} be the closed linear span of V under the Fisher norm. Then $(\bar{V}, \|\cdot\|)$ is a Hilbert space.

Furthermore, define the smooth functional of θ as

$$\psi(\theta) = b'_1\alpha + b'_2\beta + b'_3\gamma,$$

where $b = (b'_1, b'_2, b'_3)'$ is any vector of $2p + 1$ dimension with $\|b\| = 1$. For any $v \in V$, we denote

$$\dot{\psi}(\theta_0)[v] = \left. \frac{d\psi(\theta_0 + sv)}{ds} \right|_{s=0}.$$

Note that $\psi(\theta) - \psi(\theta_0) = \dot{\psi}(\theta_0)[\theta - \theta_0]$. It follows from the Riesz representation theorem that there exists $v^* \in \bar{V}$ such that $\dot{\psi}(\theta_0)[v] = \langle v^*, v \rangle$ for all $v \in \bar{V}$ and $\|v^*\|^2 = \|\dot{\psi}(\theta_0)\|^2$. Thus it follows from the Cramér-Wold device that to prove Theorem 3, it suffices to show that

$$n^{1/2} \langle \hat{\theta}_n - \theta_0, v^* \rangle \rightarrow N(0, b' \Sigma b) \text{ in distribution} \quad (4)$$

since $b' \{(\hat{\alpha}_n - \alpha_0)', (\hat{\beta}_n - \beta_0)', (\hat{\gamma}_n - \gamma_0)'\}' = \psi(\hat{\theta}_n) - \psi(\theta_0) = \dot{\psi}(\theta_0)[\hat{\theta}_n - \theta_0] = \langle \hat{\theta}_n - \theta_0, v^* \rangle$. In fact, (4) can be proved using the similar arguments of Theorem 1 of Shen (1997). For each component ϑ_q of ϑ , $q = 1, 2, \dots, 2p + 1$, we denote by $\zeta_q^* = (b_{1q}^*, b_{2q}^*)$ the solution to

$$\inf_{\zeta_q} \left\{ l_g \cdot e_q - l_{b_1}[b_{1q}] - l_{b_2}[b_{2q}] \right\}^2,$$

where $l_g = (l'_\alpha, l'_\beta, l'_\gamma)'$, and e_q is a $(2p + 1)$ -dimensional vector of zeros except the q -th element equal to 1. $l_{b_1}[b_1]$ and $l_{b_2}[b_2]$ are the directional derivatives with respect to A_1 and A_2 and can be calculated as directional derivatives defined at the beginning of the proof of Theorem 3. Now let $\zeta^* = (\zeta_1^*, \dots, \zeta_{2p+1}^*)$. By the calculations of Chen et al. (2012), we have

$$\|v^*\|^2 = \|\dot{\psi}(\theta_0)\|^2 = \sup_{v \in \bar{V}: \|v\| > 0} \frac{|\dot{\psi}(\theta_0)[v]|}{\|v\|} = b' \Sigma b, \text{ where } \Sigma = [E(S_g S_g')]^{-1},$$

$S_g = \{l_g - l_{b_1^*}[b_1^*] - l_{b_2^*}[b_2^*]\}'$. Hence the semiparametric efficiency can be established by applying the result of Theorem 4 in Shen (1997), which completes the proof.

References

- Banerjee M, Sen B (2007) A pseudolikelihood method for analyzing interval censored data. *Biometrika* 94:71–86
- Burnham KP, Anderson DR (2003) Model selection and multimodel inference: a practical information-theoretic approach. Springer, Berlin
- Carnicer JM, Peña JM (1993) Shape preserving representations and optimality of the bernstein basis. *Adv Comput Math* 1:173–196
- Chen D, Sun J, Peace K (2012) Interval-censored time-to event data: methods and applications. Chapman & Hall/CRC
- Ding Y, Nan B (2011) A sieve M-theorem for bundled parameters in semiparametric models, with application to the efficient estimation in a linear model for censored data. *Ann Stat* 39:3032–3061
- Finkelstein DM (1986) A proportional hazards model for interval-censored failure time data. *Biometrics* 42:845–854 [PubMed: 3814726]

- Groeneboom P, Wellner JA (1992) Information bounds and nonparametric maximum likelihood estimation. Springer, Berlin
- Hua L, Zhang Y, Tu W (2014) A spline-based semiparametric sieve likelihood method for over-dispersed panel count data. *Can J Stat* 42(2):217–245
- Huang J, Rossini AJ (1997) Sieve estimation for the proportional-odds failure-time regression model with interval censoring. *J Am Stat Assoc* 92(439):960–967
- Lorentz GG (1986) Bernstein polynomials. Chelsea Publishing Co, New York
- Lu M, Zhang Y, Huang J (2009) Semiparametric estimation methods for panel count data using monotone B-splines. *J Am Stat Assoc* 104(487):1060–1070
- Ma L, Hu T, Sun J (2015) Sieve maximum likelihood regression analysis of dependent current status data. *Biometrika* 102(3):731–738
- Osman M, Ghosh SK (2012) Nonparametric regression models for right-censored data using Bernstein polynomials. *Comput Stat Data Anal* 56(3):559–573
- Pollard D (1984) Convergence of stochastic processes. Springer, New York
- Shen X (1997) On methods of sieves and penalization. *Ann Stat* 25(6):2555–2591
- Shen X, Wong WH (1994) Convergence rate of sieve estimates. *Ann Stat* 22:580–615
- Sun J (2006) The statistical analysis of interval-censored failure time data. Springer, New York
- Sun J, Wei LJ (2000) Regression analysis of panel count data with covariate-dependent observation and censoring times. *J R Stat Soc Ser B (Stat Methodol)* 62(2):293–302
- Sun J, Zhao X (2013) The statistical analysis of panel count data. Springer, New York
- Turnbull BW (1976) The empirical distribution function with arbitrarily grouped censored and truncated data. *J R Stat Soc Ser B* 38:290–295
- van der Vaart AW, Wellner JA (1996) Weak convergence and empirical processes: with application to statistics. Springer
- Wellner JA, Zhang Y (2000) Two estimators of the mean of a counting process with panel count data. *Ann Stat* 28:779–814
- Zhang Z, Sun J, Sun L (2005) Statistical analysis of current status data with informative observation times. *Stat Med* 24(9):1399–1407 [PubMed: 15565735]
- Zhang Y, Hua L, Huang J (2010) A spline-based semiparametric maximum likelihood estimation method for the Cox model with interval-censored data. *Scand J Stat* 37(2):338–354

Table 1

Estimation of regression parameters α and β and the variance parameter γ

(α_0, β_0)	Parameters	Bias	SSE	BSE	CP
(0, 0)	α	-0.0115	0.2818	0.2800	94.3
	β	-0.0076	0.1793	0.1876	94.6
	γ	-0.0204	0.1247	0.1283	95.4
(0, 0.5)	α	-0.0179	0.2693	0.2625	94.4
	β	-0.0110	0.1680	0.1733	95.3
	γ	-0.0126	0.1240	0.1221	94.6
(0, -0.5)	α	0.0130	0.2740	0.2746	94.4
	β	-0.0062	0.1749	0.1778	94.6
	γ	-0.0051	0.1375	0.1337	94.0
(0.5, 0)	α	-0.0119	0.2885	0.2992	94.7
	β	-0.0046	0.1792	0.1715	94.2
	γ	-0.0209	0.1229	0.1279	94.6
(0.5, 0.5)	α	0.0099	0.2734	0.2790	94.8
	β	-0.0084	0.1637	0.1740	95.6
	γ	-0.0125	0.1252	0.1194	93.8
(0.5, -0.5)	α	0.0161	0.2796	0.2865	94.4
	β	-0.0067	0.1746	0.1779	95.2
	γ	-0.0058	0.1361	0.1331	94.1
(-0.5, 0)	α	-0.0165	0.2855	0.2913	95.7
	β	-0.0111	0.1789	0.1815	95.3
	γ	-0.0205	0.1244	0.1289	95.6
(-0.5, 0.5)	α	-0.0009	0.2824	0.2875	94.9
	β	0.0062	0.1738	0.1794	95.5
	γ	-0.0153	0.1224	0.1272	95.5
(-0.5, -0.5)	α	0.0040	0.2750	0.2722	94.8
	β	-0.0062	0.1748	0.1779	94.8
	γ	-0.0038	0.1375	0.1346	94.0

Table 2

Estimation of regression parameters α and β with misspecified frailty distribution

(α_0, β_0)	Parameters	Bias	SSE	BSE	CP
(0, 0)	α	-0.0059	0.2002	0.2104	94.8
	β	-0.0020	0.1125	0.1162	94.4
(0, 0.5)	α	0.0360	0.2324	0.2208	94.3
	β	-0.0022	0.1207	0.1133	93.3
(0, -0.5)	α	-0.0154	0.2104	0.2174	95.5
	β	0.0009	0.1143	0.1137	94.6
(0.5, 0)	α	-0.0050	0.2096	0.2065	94.4
	β	0.0001	0.1147	0.1076	93.2
(0.5, 0.5)	α	0.0052	0.2318	0.2368	94.3
	β	0.0031	0.1186	0.1186	94.3
(0.5, -0.5)	α	0.0198	0.2282	0.2225	94.8
	β	-0.0163	0.1251	0.1214	95.0
(-0.5, 0)	α	-0.0054	0.2282	0.2403	95.7
	β	-0.0018	0.1249	0.1277	94.6
(-0.5, 0.5)	α	0.0198	0.2282	0.2225	94.8
	β	-0.0163	0.1252	0.1264	95.1
(-0.5, -0.5)	α	-0.0168	0.2145	0.2078	93.8
	β	-0.0024	0.1221	0.1217	94.4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Estimation of regression parameters α and β with misspecified frailty distribution and based on the data from mixed Poisson processes

(α_0, β_0)	Parameters	Bias	SSE	BSE	CP
(0, 0)	α	-0.0001	0.2167	0.2105	94.3
	β	0.0014	0.1209	0.1154	94.2
(0, 0.5)	α	-0.0181	0.2099	0.2082	94.8
	β	-0.0030	0.1208	0.1293	95.2
(0, -0.5)	α	0.0074	0.2146	0.2219	95.2
	β	0.0043	0.1260	0.1239	95.2
(0.5, 0)	α	0.0474	0.2290	0.2230	93.8
	β	0.0087	0.1222	0.1156	92.8
(0.5, 0.5)	α	0.0214	0.2094	0.2284	95.6
	β	-0.0054	0.1205	0.1146	93.9
(0.5, -0.5)	α	0.0341	0.2285	0.2255	94.8
	β	0.0026	0.1381	0.1435	95.1
(-0.5, 0)	α	-0.0400	0.2067	0.2093	94.6
	β	0.0038	0.1220	0.1195	94.4
(-0.5, 0.5)	α	-0.0106	0.2075	0.2184	95.6
	β	-0.0004	0.1227	0.1204	94.3
(-0.5, -0.5)	α	-0.0232	0.2263	0.2172	93.4
	β	-0.0061	0.1216	0.1283	95.2

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Joint analysis results of the skin cancer data

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\gamma}$
<i>m</i> = 3									
Estimator	-0.086	0.094	0.039	-0.335	-0.133	0.106	-0.015	-0.096	-0.294
BSE	0.227	0.025	0.008	0.279	0.171	0.020	0.007	0.179	0.178
<i>p</i> -value	0.707	0.000	0.000	0.230	0.438	0.000	0.038	0.529	0.098
<i>m</i> = 4									
Estimator	-0.097	0.093	0.035	-0.347	-0.145	0.106	-0.017	-0.105	-0.299
BSE	0.226	0.025	0.010	0.278	0.171	0.020	0.009	0.178	0.178
<i>p</i> -value	0.666	0.000	0.000	0.212	0.400	0.000	0.042	0.554	0.092
<i>m</i> = 5									
Estimator	-0.098	0.093	0.035	-0.347	-0.147	0.106	-0.018	-0.106	-0.299
BSE	0.225	0.025	0.012	0.279	0.171	0.020	0.009	0.178	0.178
<i>p</i> -value	0.664	0.000	0.002	0.213	0.390	0.000	0.049	0.550	0.091
<i>m</i> = 6									
Estimator	-0.100	0.093	0.034	-0.350	-0.148	0.106	-0.018	-0.106	-0.299
BSE	0.225	0.025	0.012	0.278	0.171	0.020	0.009	0.178	0.177
<i>p</i> -value	0.654	0.000	0.004	0.208	0.389	0.000	0.051	0.551	0.091