

Supplementary Material to the Paper: Estimation of Partly Linear Additive Hazards Model with Left Truncated and Right Censored Data

Arfan Raheen Afzal¹, Cheng Dong², and Xuewen Lu³

¹ Department of Mathematics & Statistics University of Calgary Calgary, Alberta,
T2N 1N4, Canada

² Cheng Dong Department of Statistics University of Missouri-Columbia Columbia,
MO 65211-6100, U.S.A

³ Department of Mathematics & Statistics University of Calgary Calgary, Alberta,
T2N 1N4, Canada

Address for correspondence: Xuewen Lu, Department of Mathematics & Statistics University of Calgary Calgary, Alberta, T2N 1N4, Canada.

E-mail: xlu@ucalgary.ca.

Phone: (+1) 403 220 6620.

Fax: (+1) 403 282 5150.

Abstract: This supplementary material provides the R code and dataset for the simulation studies and data analysis in the paper.

Key words: R code for simulation studies; nickel dataset

1 R code for simulation studies

All the R codes for the simulation studies are included in separated files in the Online Supplementary Material.

The R code for Example 1 is given in `Example1.R` for right censored data without left-truncation.

The R code for Example 2 is given in `Example2.R` for left truncated and right censored data.

The R code for Example 3 is given in `Example3.R` for left truncated and right censored data and the nonlinear functions are B-spline functions.

The R code for Example 4 is given in `Example4.R` for comparison with the kernel-based approach.

2 Nickel dataset used for data analysis

The data were obtained from the South Wales nickel refiners study ([Breslow and Day, 1987](#), Appendix D). The data contained complete records for 679 men workers employed before 1925 in a nickel refinery in South Wales. The purpose of the study is to determine the risk of developing carcinoma of the bronchi and nasal sinuses

associated with the refining of nickel. The follow-up through 1981 uncovered 137 lung cancer deaths among men aged 40 ~ 85 years and 56 deaths from cancer of the nasal sinus. Since the workers had been working in the company for various periods of time before the follow-up was initiated, their survival times were subject to left-truncation. A right-censored observation arose either because the worker died from a competing cause or because he was still alive on the date of data listings.

The data set can be found in the R library `Epi` by the following commands:

```
library(Epi)
data(nickel)
```

The data set contains seven variables: id =Case ID, icd =Primary ICD, $exposure$ =Exposure, dob =Date of Birth, $age1st$ =Age first employed, $agein$ =Age follow-up began, $ageout$ =Age at death or loss.

The censoring indicator is computed from $I(icd == 160)$, there were 56 deaths from cancer of nasal sinus. The left-truncation time= $agein - age1st$ and the survival time= $ageout - age1st$.

References

- Breslow, N. E. and Day, N. E. (1987). *Statistical methods in cancer research*, volume 2. International Agency for Research on Cancer Lyon.