האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

SIMPSON'S PARADOX FOR THE COX MODEL

by

CLELIA DI SERIO, YOSEF RINOTT
and MARCO SCARSINI

Discussion Paper # 441 January 2007

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Simpson's paradox for the Cox model

Clelia Di Serio*
Dipartimento di Psicologia
Università Vita-Salute San Raffaele
Via Olgettina 58
I–20132 Milano, Italy
diserio.clelia@hsr.it

Yosef Rinott†
Department of Statistics
and Center for the Study of Rationality
Hebrew University of Jerusalem
Mount Scopus
Jerusalem 91905, Israel
rinott@mscc.huji.ac.il

Marco Scarsini‡
Dipartimento di Scienze Economiche e Aziendali
LUISS
Viale Pola 12
I–00198 Roma, Italy
and HEC, Paris
mscarsini@luiss.it

January 2007

*Partially supported by MIUR-COFIN
†Partially supported by the Israel Science Foundation grant No. 473/04
‡Partially supported by MIUR-COFIN

Abstract

In the context of survival analysis, we define a covariate X as *protective* (*detrimental*) for the failure time T if the conditional distribution of $[T|X = x]$ is stochastically increasing (decreasing) as a function of x. In the presence of another covariate Y, there exist situations where $[T|X = x, Y = y]$ is stochastically decreasing in x for each fixed y, but $[T|X = x]$ is stochastically increasing. When studying causal effects and influence of covariates on a failure time, this state of affairs appears paradoxical and raises the question of whether X should be considered protective or detrimental. In a biomedical framework, for instance when X is a treatment dose, such a question has obvious practical importance. Situations of this kind may be seen as a version of Simpson's paradox.

In this paper we study this phenomenon in terms of the well-known Cox model. More specifically, we analyze conditions on the parameters of the model and the type of dependence between X and Y required for the paradox to hold. Among other things, we show that the paradox may hold for residual failure times conditioned on $T > t$ even when the covariates X and Y are independent. This is due to the fact that independent covariates may become dependent when conditioned on the failure time being larger than t.

AMS 2000 Subject Classification: Primary 62N99, Secondary 62P10.

Keywords: Detrimental covariate, protective covariate, proportional hazard, omitting covariates, positive dependence, total positivity.

1 Introduction

Consider a failure time T and a covariate X. We say that X is *detrimental* for T if $P(T > t|X = x)$ is decreasing in x, and *protective* if it is increasing, that is, T is stochastically decreasing, or increasing in X, respectively. In the framework of survival analysis framework, data sometimes show situations where a certain covariate X appears to be detrimental to life in each subgroup defined by the values of another covariate Y, but seems protective when there is no conditioning on Y. This may be puzzling, and one may then find it hard to understand the nature of the influence of X on the failure time T. In fact this is an important case of the well-known Simpson paradox. In this paper we analyze this phenomenon in terms the classical Cox regression model and obtain conditions under which the paradox is natural, rather than being a surprising pathology.

A huge body of literature exists on Simpson's paradox (Simpson (1951)) and related phenomena. An early example concerning survival appears in Cohen and Nagel (1934), who cited actual death rates from tuberculosis in 1910 in two cities (Richmond, Virginia, and New York, New York):

- The death rate for African-Americans was lower in Richmond than in New York.
- The death rate for Caucasians was lower in Richmond than in New York.
- The death rate for the total combined population of African-Americans and Caucasians was higher in Richmond than in New York.

In the above terms, living in New York is detrimental for each ethnic subgroup, but protective for the combined population.

The paradox, which in some form goes back to Yule (1903), was examined more recently by Blyth (1972a,b, 1973) in connection with some principles of decision theory. Necessary conditions for the paradox were studied by Lindley and Novick (1981), and Mittal (1991). Good and Mittal (1987) studied conditions for related paradoxes, called amalgamation paradoxes. Cohen (1986) showed some implications of the paradox in demography.

Blyth (1973) gave a simple description of Simpson's paradox in terms of conditional probabilities. Given three events E, F, H, the paradox is the simultaneous occurrence of the following three inequalities

$$\begin{aligned} P(E|F \cap H) &\geq P(E|F^c \cap H), \\ P(E|F \cap H^c) &\geq P(E|F^c \cap H^c), \\ P(E|F) &< P(E|F^c). \end{aligned} \tag{1.1}$$

Samuels (1993) extended the consideration of the paradox from events to random variables and explained it as a particular case of the *association reversal* or of the *association distortion* phenomena. This idea has been further extended by Scarsini and Spizzichino (1999) who determined necessary conditions for the paradox when different notions of positive dependence are considered. Rinott and Tam (2003) described conditions under which an instance of the

paradox is natural rather than surprising. The above list of papers on Simpson's paradox is by no means exhaustive.

This problem is related to the issue of ignoring latent variables. In biostatistics latent heterogeneity is modelled by means of an individual frailty variable: frailer individuals tend to die earlier, thus introducing a crucial diversity in the population. For frailty models see e.g., Hougaard (1984, 1986), Andersen et al. (1993). Neglecting unobserved covariates could result in misleading and ambiguous interpretation of the the role of an observable covariate on the failure times distribution (see, e.g., Di Serio (1997, 2003), Everitt and Dunn (1998)).

We formalize the paradox for residual failure times as follows. Given a failure time T and two covariates X, Y, Simpson's paradox at (t, s) is the simultaneous occurrence of the following two conditions:

$$P(T > t + s | T > t, X = x, Y = y) \text{ is decreasing in } x \text{ for all } y \tag{1.2a}$$

$$P(T > t + s | T > t, X = x) \text{ is strictly increasing in } x. \tag{1.2b}$$

In words, (1.2a) means that conditionally on every value of the covariate Y, higher values of X stochastically reduce the survival time (i.e., X is *detrimental* for T given Y), whereas (1.2b) means that, unconditionally, the opposite is true (i.e., X is *protective* for T). In a similar way one may consider situations where all monotonicities are reversed.

To see the relation between (1.2) and (1.1), define $E = \{T > t + s\}$, $F = \{X = x\}$, $H = \{Y = y\}$, and

$$G = \{X \in \{x, x'\}\} \cap \{Y \in \{y, y'\}\} \cap \ \{T > t\}, \tag{1.3}$$

The inequalities (1.1) applied to the conditional probability $P(\cdot|G)$ for all $x' > x$, $y' > y$, become (1.2). Conditioning on G amounts to treating X and Y as dichotomous, and looking at T only after time t. Indeed the dichotomous case will receive special attention, since the problem essentially reduces to this case.

In this paper we show that under certain conditions the paradox occurs, and in fact it is quite natural. We study the range of values of (t, s) for which the paradox (1.2) holds, and show that under some circumstances it holds for all $s, t > 0$, and that in general the sets of s and t where it holds must be intervals.

It is clear that the classical Simpson's paradox (1.1) can arise only if the conditioning events exhibit some form of dependence, that is, if F and H are independent then the paradox of (1.1) is impossible. However, our analysis shows that we can have Simpson's paradox even if the covariates X and Y are independent. This surprising phenomenon is due to the fact that conditioning destroys independence, so, as time goes by, the covariates may become dependent, and, for suitable values of the parameters, the paradox may arise.

The paper is organized as follows. Section 2 describes the model. Section 3 deals with dichotomous covariates. We show the possibility of the paradox for independent covariates in Section 4. Section 5 considers the case of continuously distributed covariates. Section 6 deals with the issue of omitting relevant covariates in a Cox model. In Section 7 a biomedical example of the paradox is shown within a gene therapy context. Data from failure times of mice are examined conditional on two covariates which describe properties of the cancer

treatment they receive. The existence of Simpson's paradox in such data raises the question whether a particular outcome of a treatment is desirable or not, since its influence on failure times in the whole population and within subpopulations is inconsistent. All proofs are contained in Section 9.

2 The model

The conditional survival function of T will be modeled via the following well-known proportional hazard Cox model (see Cox (1972))

$$h(t|x,y) = h_0(t)\exp\{\beta_X x + \beta_Y y\}, \tag{2.1}$$

where h is the conditional hazard function defined by

$$h(t|x,y) := \lim_{\varepsilon \searrow 0} \frac{P(t+\varepsilon > T > t | T > t, X = x, Y = y)}{\varepsilon},$$

and h_0 is the underlying baseline hazard rate. We assume that h_0 is a positive function such that $\int_t^\tau h_0(u)\ \mathrm{d}u$ is finite if and only if $\tau < \infty$ for all $t > 0$, that is

$$\infty > \int_t^\tau h_0(u)\ \mathrm{d}u \to \infty \quad \text{as} \quad \tau \to \infty\,. \tag{2.2}$$

This condition on h_0 corresponds to assuming that the failure time T is finite, but cannot be bounded with probability 1 by any finite constant. The latter assumption is technically useful, since it simplifies the presentation, but can be avoided.

By standard calculations the conditional survival function for the Cox model can be written as

$$P(T > t+s | T > t, X = x, Y = y) = \exp\left\{\left(-\int_t^{t+s} h_0(u)\ \mathrm{d}u\right)\exp\{\beta_X x + \beta_Y y\}\right\}. \tag{2.3}$$

To avoid trivialities we will assume that both $\beta_X, \beta_Y \neq 0$. Note that when $\beta_X > 0$ the probability in (2.3) is decreasing in x for all $t, s > 0$. In fact, under the Cox model the monotonicity of (2.3) and the monotonicity condition (1.2a) are both equivalent to assuming $\beta_X > 0$. This assumption will be made throughout the paper.

Notice that for all values x_1, y_1, x_2, y_2 the difference

$$P(T > t+s | T > t, X = x_1, Y = y_1) - P(T > t+s | T > t, X = x_2, Y = y_2)$$

cannot change sign as a function of s. This is due to proportionality of the conditional hazard function (2.1), which implies that changing the value of some covariates moves the whole conditional hazard function up or down, but does not allow any crossing.

On the other hand the difference

$$P(T > t+s | T > t, X = x_1) - P(T > t+s | T > t, X = x_2)$$

can change sign, since the expression

$$P(T > t+s|T > t, X = x) = \int P(T > t+s|T > t, X = x, Y = y)\ \mathrm{d}F_Y(y|T > t, X = x) \tag{2.4}$$

is not a Cox model, but rather a mixture of Cox models, and mixing destroys proportionality of the hazards. This feature will be fundamental in proving results about Simpson's paradox for this model. Mixture of Cox models arise in Bayesian statistics (see, e.g., Gouget and Raoult (1999), Ibrahim et al. (2001)).

3 Dichotomous covariates

In this section we assume that the covariates X and Y are dichotomous. Although simple, this case is important in applications and it is rich enough to show the salient features of the paradox. Other cases will be discussed later.

In Theorem 3.1 below, first we fix t and show that the set of values of s for which the paradox (1.2) holds, is a single interval. Conditions for this interval to be empty, contain the origin and/or be unbounded, depend on the joint distribution of the covariates and on their relative influence. The last part of the theorem shows that the range of t for which the paradox holds for all $s > 0$ is an upper interval.

Define

$$p^t_{y|x} = P(Y = y|T > t, X = x), \quad A = \exp\{\beta_X\}, \quad \text{and} \quad B = \exp\{\beta_Y\}. \tag{3.1}$$

Theorem 3.1. *Consider the Cox model* (2.3) *where the covariates X and Y take only two values,* $0, 1$ *and are not degenerate. Let*

$$\beta_Y < 0 < \beta_X. \tag{3.2}$$

Then

(a) *The probability*

$$P(T > t+s|T > t, X = x, Y = y) \text{ is decreasing in } x \text{ for each } y \in \{0,1\} \text{ and } s, t > 0, \tag{3.3}$$

and there exist some $0 \le s_1 \le s_2 \le \infty$ depending on β_X, β_Y, t and h_0 such that inequality

$$P(T > t+s|T > t, X = 0) < P(T > t+s|T > t, X = 1) \tag{3.4}$$

holds if and only if $s \in (s_1, s_2)$. This interval may be empty. A sufficient condition for (s_1, s_2) to be empty is $p^0_{0|0} = 0$.

Moreover, for s_1, s_2 of (a) *we have:*

(b) $s_1 < s_2 = \infty$ *if and only if*

$$\beta_X < -\beta_Y \tag{3.5}$$

and

$$p^0_{1|0} = 0 \tag{3.6}$$

both hold.

(c) $0 = s_1 < s_2$ *if and only if*

$$p^t_{1|1} \geq \frac{A-1}{A-AB} + \frac{p^t_{1|0}}{A}. \tag{3.7}$$

(d) $s_1 = 0$ *and* $s_2 = \infty$ *if and only if* (3.6) *and*

$$p^t_{1|1} \geq \frac{A-1}{A-AB} \tag{3.8}$$

both hold. In this case there exists some $0 \leq t_0 \leq \infty$ *such that* (3.4) *holds for all* $t \in (t_0, \infty)$ *and all* $s > 0$.

Note again that the combination of (3.3) and (3.4) is exactly Simpson's paradox for those s and t such that (3.4) holds.

Under condition (3.6), it can be shown that X and Y exhibit very strong positive dependence: they are comonotone, that is, they are both increasing functions of some latent variable. It is easy to see that under (3.2) condition (3.7) implies $p^t_{1|1} \geq p^t_{1|0}$, which can be proved equivalent to X and Y being positively correlated, which for dichotomous variables implies most concepts of positive dependence.

From Theorem 3.1 we see that the paradox holds for large values of s only if a very strong form of dependence exists for the covariates. We will show in the proof of Theorem 3.1 that $p^0_{1|0} = 0$ if and only if $p^t_{1|0} = 0$ for any $t > 0$, so that condition (3.6) means that given $X = 0$ then $Y = 0$ at any given time.

Notice that the choice of $0, 1$ for the values of X, Y is arbitrary. Theorem 3.1 can be adapted to any other pair of values. It is essential that Y take only two values, whereas such a restriction is not needed for X; for each pair of values $x_1 < x_2$ we would have an interval of s-values where the inequality $P(T > t + s|T > t, X = x_1) < P(T > t + s|T > t, X = x_2)$ holds. Since the intersection of interval is itself an interval, we see that the conclusion of the theorem holds without restricting X to be dichotomous.

In our formulation, the paradox requires that the covariates exhibit positive dependence. However, a paradox can hold in a similar way under negative dependence, provided β_X and β_Y are both positive. If they are both negative, the paradox may hold with all inequalities reversed.

4 Independent covariates

A straightforward calculation in (1.1) shows that a necessary condition for the existence of Simpson's paradox is that the conditioning events E and F be dependent. In the Cox

model with dichotomous covariates, a necessary condition for the paradox (1.2) is that given $\{T > t\}$, X and Y be dependent. In particular, for the paradox to hold at $t = 0$ X and Y cannot be independent. Nevertheless it is possible to have the paradox at some time $t > 0$ even when the covariates are independent at time 0, due to the fact that independence is not preserved under conditioning.

Proposition 4.1. *There exists a choice of parameters in the Cox model* (2.3) *such that the paradox* (1.2) *holds for some $s, t > 0$ although the covariates X and Y are independent.*

5 Continuous covariates

In this section we consider the case of continuous covariates and provide some sufficient conditions for the paradox to hold.

Theorem 5.1. *Consider the Cox model described in* (2.3)*. Let the covariate Y have a conditional density*

$$f_Y(y|X = x) = g(y - \mu - bx), \tag{5.1}$$

where g is a strictly log-concave function and $b > 0$. Then Simpson's paradox (1.2) *holds for all $t, s > 0$ if and only if*

$$\beta_Y < 0 < \beta_X, \quad \textit{and} \quad b > \frac{\beta_X}{|\beta_Y|}. \tag{5.2}$$

The assumption on the density of Y given X used in Theorem 5.1 is equivalent to assuming $Y = X + V$ with X and V independent, and the density of V log-concave, such as normal or gamma with shape parameter ≥ 1. Under (5.1) the joint density of (X, Y) is TP_2, which is a strong notion of positive dependence, see, e.g., Joe (1997). It can be shown that when (5.2) and (5.1) hold then the joint conditional density of (X, Y) given $T > t$ is TP_2 for all $t > 0$, so that the TP_2 property is preserved in time.

Condition (5.2) is interesting since it intertwines the strength of dependence between X and Y and their relative influence on T. The first is expressed by the parameter b, the second by the ratio $\beta_X/|\beta_Y|$.

We further demonstrate the relation between the parameters of the Cox model and the dependence structure of the covariates in the special case of normal covariates, in which the dependence is simply captured by the correlation coefficient. We will see that the condition required is that the correlation be sufficiently large with respect to the ratio of the β's.

Corollary 5.2. *Consider the case where the conditional distribution of Y given $X = x$ is normal with mean $\mu + \rho x$ and variance $1 - \rho^2$, which happens for instance when the joint distribution of (X, Y) is bivariate normal with unit variances and correlation ρ. Then Simpson's paradox* (1.2) *holds for all $t, s > 0$ if and only if $\beta_Y < 0 < \beta_X$, and $\rho > \beta_X/|\beta_Y|$.*

6 Omitting covariates

The problem that we have considered in this paper is related to the issue of omitting covariates, but it has different features. It is true that omitting a covariate in a Cox model may lead to a change in the influence of a remaining covariate. For instance it is possible that expression (2.3) be decreasing in x, whereas the Cox model obtained by omitting the covariate Y

$$\exp\left\{\left(-\int_t^{t+s} h_0(u)\ \mathrm{d}u\right)\exp\{\beta_X x\}\right\} \tag{6.1}$$

is is increasing in x.

The issue of misspecification by omitting covariates in the Cox model has been studied by several authors (see, e.g., Gail et al. (1984), Solomon (1984), Lagakos and Schoenfeld (1984, 1986), Bretagnolle and Huber-Carol (1985, 1988), Morgan (1986), Struthers and Kalbfleisch (1986), Lin and Wei (1989), Anderson and Fleming (1995), Gerds and Schumacher (2001), DiRienzo and Lagakos (2001a,b, 2004), Chen (2002)). Recently Sane and Kharshikar (2001) have connected the issue to Simpson's paradox.

As we pointed out before, the class of Cox models is not closed under marginalization, therefore, if (2.3) holds, then (6.1) does not represent $P(T > t + s|T > t, X = x)$.

That misspecification of a model can lead to paradoxes is not entirely surprising. What we showed in our paper is more subtle. We showed that a Simpson-type paradox can arise even when the model is perfectly specified, just due to marginalization. It is interesting to see that, due to the proportional hazard feature of the Cox model, two conditional survival functions of the form (2.3), conditional on two different values of X, never cross. The same obviously applies to (6.1). Therefore misspecification through omission of a covariate can lead only to a a very drastic form of Simpson paradox, namely (2.3) is decreasing in x, and (6.1) is increasing in x *for all* s and t.

In our model it is possible to have (2.3) decreasing in x, and (2.4) increasing in x *for some* values of s and t. Furthermore, the key element of the paradox is the dependence of the covariates X and Y, and this dependence is explicitly modelled via their joint conditional distribution at different times t.

7 Example: survival in gene therapy

Gene therapy is a form of molecular medicine which treats genetic diseases by replacing a defective gene, responsible for the pathology, with a functional one. The basic principle is to introduce a piece of genetic material into cells via a vector, which is typically a virus. The virus integrates with the cell DNA and thus delivers the genetic material into the cell nucleus.

It has been observed that when the virus integrates in certain gene regions (close to the starting point of the transcription), deregulation of the gene transcription may induce insertional mutagenesis. In this case genotoxicity occurs and as a consequence a subject may develop cancer. The integration process is then defined unsafe. Searching for a so called safety vector is now a major goal in gene therapy.

In the application considered here identically inbred mice are made tumor prone by knocking out the oncosuppressor related gene Cdkn2a. These mice develop a variety of tumors with a predictable onset time of 300 days. Bone marrow cells are then extracted from the mice and different vectors are inoculated in them. These cells are transplanted back in the mice, whose survival is then observed. One goal of this study is to investigate the influence of some covariates related to the integration process on the survival of mice. For details see Montini et al. (2006).

We consider two covariates which are related to how much and where the vector integrates in the genome. These variables are: Y = CIS (common integration sites) and X = NUCLEUS.

The value $Y = 1$ indicates that integration occurred in a low density area, and $Y = 0$ means that integration occurs in a high integration density area, that is, in a genomic region that was targeted twice or more in close proximity.

The value $X = 0$ indicates that integration occurred in a gene which codifies a protein produced within the cell nucleus. This represents a further risk factor for the integration process.

In this data set we consider bone marrow cellular DNA integrations of 60 mice classified with respect to the above two dichotomous variables. Survival time of each mouse is given in days.

The joint distribution of (X, Y) is shown in the following table, indicating that in the present coding they are positively dependent:

	X		
Y	0	1	Total
0	9	6	15
1	13	32	45
Total	22	38	60

Consider now a Cox Model as in (2.1). The standard estimation procedure that we use is based on maximizing the partial likelihood as a function of the β parameters and using the Breslow (1972, 1974) estimator for the cumulative baseline hazard function.

The proportionality-of-hazard assumption of Cox model was tested and was not rejected ($p = 0.23$). See Grambsch and Therneau (1994) for details about the test. The β coefficients of the regression were proved significant by the likelihood ratio test ($p = 0.00003$). The estimated coefficients and their 0.95-confidence intervals are

	β	$\exp(\beta)$	Confidence interval
β_X	0.405	1.499	$1.01 - 2.91.$
β_Y	-2.01	0.134	$0.056 - 0.321.$

Therefore condition (3.2) is met.

Note that condition (3.7) at time $t = 0$ in Theorem 3.1 is now satisfied since $P(Y = 1|X = 0) = 13/22 = 0.59$ and

$$P(Y = 1|X = 1) = 0.84 > \frac{\exp(0.405) - 1}{\exp(0.405)(1 - \exp(-2.01))} + \frac{0.59}{\exp(0.405)} = 0.78$$

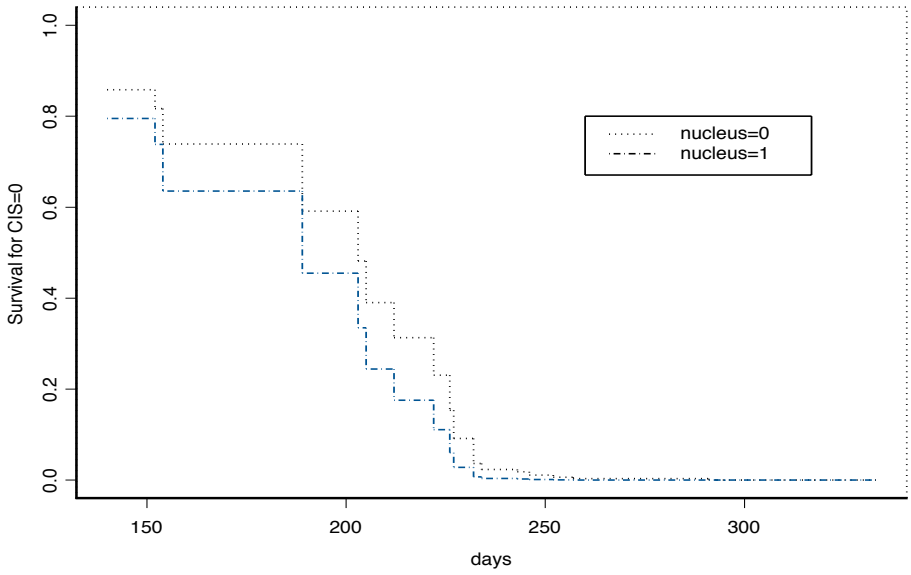

Figure 1: Estimated survival functions when CIS = 0

We next provide the survival plots based on the above estimates. Let $\widehat{S}(t, Z_{ij})$ be the estimated survival function corresponding to the covariate vector $Z = (X = i; Y = j)$, with $i, j \in \{0, 1\}$ in the Cox model. Figure 1 compares $\widehat{S}(t, Z_{00})$ to $\widehat{S}(t, Z_{10})$, whereas Figure 2 compares $\widehat{S}(t, Z_{01})$ to $\widehat{S}(t, Z_{11})$. Figure 3 shows the mixture survival functions (when averaging over CIS). The paradox occurs for values smaller than 290.

8 Discussion

In this paper we consider a Cox regression model and provide conditions on its parameters for the occurence of a Simpson's paradox. Our results and the motivating example show that Simpson's paradox may occur naturally and the range of values where it occurs can be described. The paradox is not due to misspecification of the model but just to marginalization.

From a biomedical point of view it is fundamental to properly evaluate the role of each covariate on the distribution of a failure time. Therefore it is important to recognize that under certain assumptions, such as the Cox model, and the presence of certain dependence relations among covariates, a reversal of their impact on failure times as expressed by the model may occur. In such cases, classification of the covariates as protective or detrimental may be subtle, requiring deeper understanding of causality relations among them.

One can study more complex situations of covariates which assume more than two values, and similar techniques show that the paradox may occur in disjoint intervals whose number

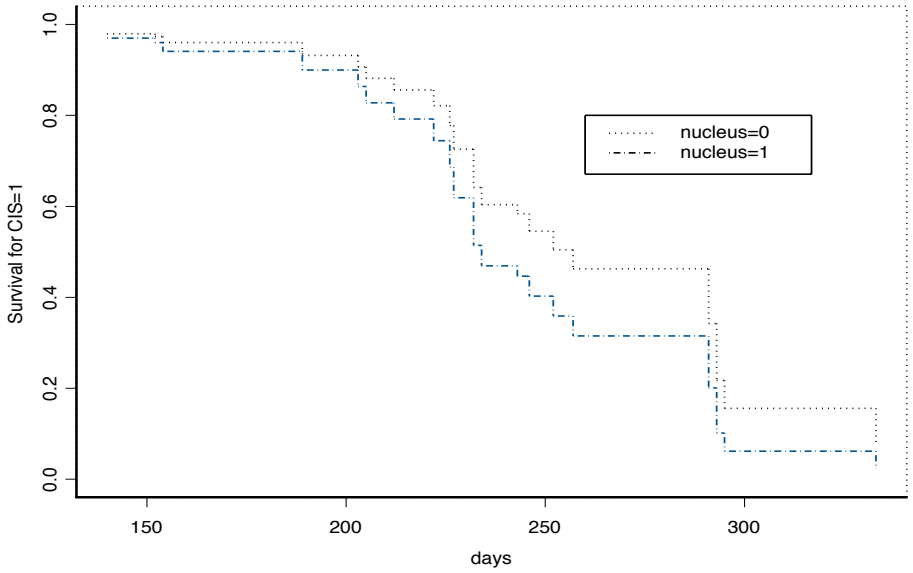

Figure 2: Estimated survival functions when CIS = 1

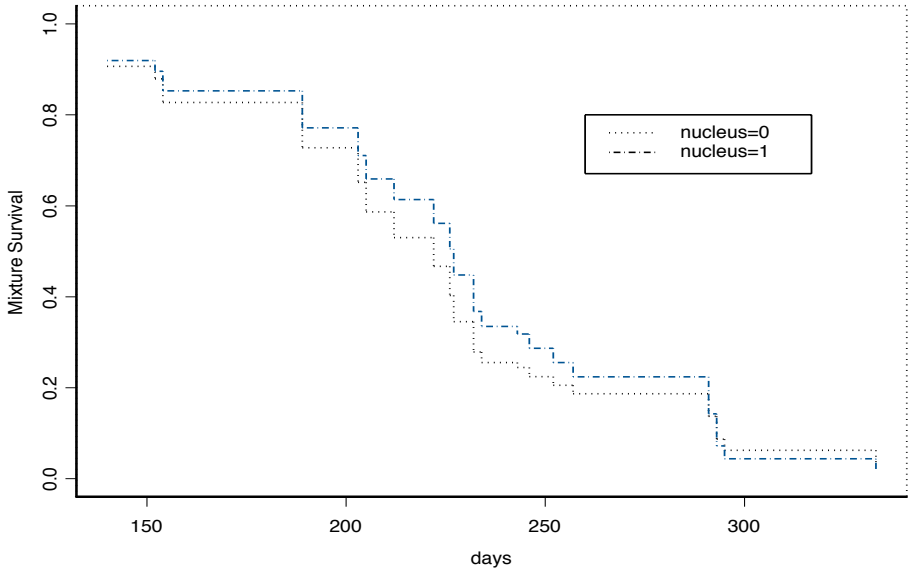

Figure 3: Mixed survival function

is directly related to the range of the covariates. Also, different dependence structures of covariates could be studied.

While the Simpson phenomenon is demonstrated here in the Cox model, it can be investigated and generalized to other regression models.

It should be clear that interpretation of the model in terms of the nature of the dependence between covariates, which may change in time, and their influence on failure time, is a delicate issue, which requires further analysis.

9 Proofs

Total Positivity

For the proofs of our results we need some background in Total Positivity. We now provide without proofs the required basic results in this area. The reader is referred to Karlin (1968) for further definitions, results and proofs, and to Brown et al. (1981) for a useful formulation for statistical applications of the theory.

Definition 9.1. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is said to be SR_k if there exist $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, 1\}$ such that for all for $m = 1, \ldots, k$ and for all *for all* $x_1 < \cdots < x_m$, $y_1 < \cdots < y_m$ we have

$$\varepsilon_m \det \begin{bmatrix} \phi(x_1, y_1) & \cdots & \phi(x_1, y_m) \\ \vdots & & \vdots \\ \phi(x_m, y_1) & \cdots & \phi(x_m, y_m) \end{bmatrix} \geq 0. \tag{9.1}$$

The condition simply means that all determinants of the above type of any given order up to k have the same sign. If $\varepsilon_m = 1$ for $m = 1, \ldots, k$, then ϕ is said to be TP_k, and RR_k if the sign sequence is $\varepsilon_m = (-1)^{m(m-1)/2}$. SSR_k, STP_k, and SRR_k are defined in the same way with strict inequalities in (9.1). If ϕ is SR_k for all $k = 1, 2, \ldots$, then ϕ is said to be SR. TP, RR, SSR, STP and SRR are defined similarly. In the above SR stands for *sign regular*, TP for *totally positive*, and RR for *reverse rule*, and when S is added these properties are said to hold *strictly*.

Example 9.2. The function $\psi(x, y) = \exp\{xy\}$ is STP, whereas the function $\phi(x, y) = \exp\{-xy\}$ is SRR. Also, g is a strictly log-concave function if and only if $\phi(x, y) = g(x - y)$ is STP_2.

Proposition 9.3. *If $\phi(x, y)$ is SSR_k and ζ and ξ are both strictly monotone, then $\phi(\zeta(x), \xi(y))$ is SSR_k.*

If both ζ, ξ are strictly increasing or both strictly decreasing then $\phi(x, y)$ and $\phi(\zeta(x), \xi(y))$ have the same sign sequence, whereas if one is strictly increasing and the other strictly decreasing the sign sequence of $\phi(\zeta(x), \xi(y))$ is obtained from that of $\phi(x, y)$ by multiplying its m-th sign by $(-1)^{m(m-1)/2}$. In this case if $\phi(x, y)$ is STP_k then $\phi(\zeta(x), \xi(y))$ is SRR_k.

Proposition 9.4 (Composition formula)**.** *If ϕ and ψ are both SSR_k, having sign sequences ε_m and ε'_m, respectively, and σ is a nonnegative σ-finite measure, then the convolution*

$$\zeta(x,y) = \int \phi(x,z)\psi(z,y)\ \mathrm{d}\sigma(z)$$

is SSR_k with sign sequence $\varepsilon_m\varepsilon'_m$ for $m = 1,\dots,k$.

Definition 9.5. Let g be a function defined on a totally ordered finite set $\mathcal{X} = \{x_1,\dots,x_n\} \subset \mathbb{R}$, where $x_1 < x_2 < \cdots < x_n$. Then $S^-(g)$ denotes the number of sign changes of the sequence $g(x_1),\dots,g(x_n)$, when zeros are deleted; $S^+(g)$ denotes the maximum number of sign changes of the sequence $g(x_1),\dots,g(x_n)$ that can be obtained by counting zeros as either $+$ or $-$.

If $\mathcal{X}$ is any subset of $\mathbb{R}$, not necessarily finite, then $S^-(g) = \sup_{V\in\mathcal{V}(\mathcal{X})} S^-(g_V)$, where $\mathcal{V}(\mathcal{X})$ is the class of finite subsets of $\mathcal{X}$, and g_V is the restriction of g to V. Analogously for S^+.

The following Proposition is somewhat weaker than Theorem 3.1 p. 233 of Karlin (1968).

Proposition 9.6 (Variation diminishing property)**.** *Let g be a real valued function defined on $\mathbb{R}$, and let σ be a nonnegative sigma-finite measure on $\mathbb{R}$. If ϕ is SSR_k and $S^-(g) \le k-1$, then $S^+(f) \le S^-(g)$ where*

$$f(x) = \int g(y)\phi(x,y)\ \mathrm{d}\sigma(y). \tag{9.2}$$

Proofs of Section 3

In the sequel we will use the following notation

$$C_{t,s} = \int_t^{t+s} h_0(u)\ \mathrm{d}u. \tag{9.3}$$

Proof of Theorem 3.1. (a) The statement of (3.3) follows readily from (2.3).

Using (2.4), (3.1) and (9.3) we have

$$P(T > t+s|T > t, X = 0) = p^t_{0|0}\exp\{-C_{t,s}\} + p^t_{1|0}\exp\{-C_{t,s}B\}, \tag{9.4a}$$
$$P(T > t+s|T > t, X = 1) = p^t_{0|1}\exp\{-C_{t,s}A\} + p^t_{1|1}\exp\{-C_{t,s}AB\}. \tag{9.4b}$$

For $x, y \in \{0,1\}$,

$$p^t_{y|x} = \frac{p^0_{y|x}P(T > t|X = x, Y = y)}{p^0_{1|x}P(T > t|X = x, Y = 1) + p^0_{0|x}P(T > t|X = x, Y = 0)}, \tag{9.5}$$

and, since the support of T is unbounded, we have

$$p^t_{y|x} > 0 \quad \text{if and only if} \quad p^0_{y|x} > 0. \tag{9.6}$$

By (9.4), (3.4) is equivalent to $f(s) < 0$, where

$$f(s) := p^t_{0|0} \exp\{-C_{t,s}\} + p^t_{1|0} \exp\{-C_{t,s}B\} - p^t_{0|1} \exp\{-C_{t,s}A\} - p^t_{1|1} \exp\{-C_{t,s}AB\}. \quad (9.7)$$

We can write $f(s) = \int g(y)\phi(s,y)\, \mathrm{d}\sigma(y)$, where σ is the measure assigning unit mass to each $y \in \{B, AB, 1, A\}$, $g(y)$ is a function assigning values $p^t_{1|0}$, $-p^t_{1|1}$, $p^t_{0|0}$, $-p^t_{0|1}$ to $y = B, AB, 1, A$, respectively, and $\phi(s,y) = \exp\{-C_{t,s}y\}$. Since $C_{t,s}$ is strictly increasing in s, and $\exp\{-xy\}$ is SRR, it follows by Proposition 9.3 that $\phi(s,y)$ is SRR, hence SRR_4.

Our goal is to show that the set where $f(s)$ is negative is an interval. Whatever the order in the set $\{B, AB, 1, A\}$, we have $S^-(g) \le 3$. Using the variation diminishing property of ϕ (see Proposition 9.6), we conclude that $S^+(f) \le 3$.

Consider first the case $p^t_{1|0} > 0$. We start by showing that $f(s) > 0$ for large s. Multiply (9.7) by $\exp\{C_{t,s}B\}$ to obtain the expression

$$p^t_{0|0} \exp\{-C_{t,s}(1-B)\} + p^t_{1|0} - p^t_{0|1} \exp\{-C_{t,s}(A-B)\} - p^t_{1|1} \exp\{-C_{t,s}A(1-B)\}. \quad (9.8)$$

As $s \to \infty$, since $B < 1 < A$ by (3.2), and $C_{t,s} \to \infty$ by (2.2), the expression (9.8) tends to $p^t_{1|0}$ and therefore $f(s) > 0$ for large s. Using the fact that $f(0) = 0$, we see that if there were two or more disjoint intervals where the function f is negative, then we would have $S^+(f) > 3$, which is a contradiction.

If $p^t_{1|0} = 0$, then clearly $S^-(g) \le 2$, which implies $S^+(f) \le 2$. Then it is easy to see that having two disjoint intervals where f is negative would imply $S^+(f) > 2$, which again is a contradiction.

Finally suppose $p^0_{0|0} = 0$ and therefore, by (9.6), $p^t_{0|0} = 0$, and $p^t_{1|0} = 1$. Since $\exp\{-C_{t,s}B\} > \exp\{-C_{t,s}A\}, \exp\{-C_{t,s}AB\}$, in the present case (9.7) readily implies $f(s) > 0$ for all $s > 0$.

(b) The condition $s_2 = \infty$ is equivalent to $f(s) < 0$ for large enough s. Suppose (3.5) and (3.6) hold. Clearly (3.6) implies $p^0_{0|0} = 1$ and, since Y is not degenerate, $p^0_{1|1} > 0$ also follows. By (9.6) we have $p^t_{0|0} = 1$ and $p^t_{1|1} > 0$, and by (9.7) we obtain

$$f(s)\exp\{C_{t,s}\} = 1 - p^t_{0|1} \exp\{C_{t,s}(1-A)\} - p^t_{1|1} \exp\{C_{t,s}(1-AB)\}. \quad (9.9)$$

By (3.5) we have $AB < 1$ and clearly the expression in (9.9) goes to $-\infty$ and therefore $f(s) < 0$ as $s \to \infty$.

To prove the converse, suppose $f(s) < 0$ as $s \to \infty$. Note that by multiplying (9.7) by $\exp\{C_{t,s}B\}$ the expression converges to $p^t_{1|0}$ as $s \to \infty$, and therefore $f(s) < 0$ as $s \to \infty$ implies $p^t_{1|0} = 0$ for all t, and (3.6) holds. We can now use (9.9) in a similar way to show that if $f(s) < 0$ as $s \to \infty$, then $AB < 1$, which is (3.5).

(c) Consider the function

$$R(C) = (1 - p^t_{1|0}) \exp\{-C\} + p^t_{1|0} \exp\{-CB\} - (1 - p^t_{1|1}) \exp\{-CA\} - p^t_{1|1} \exp\{-CAB\},$$

and note that (3.4) is equivalent to $R(C_{t,s}) < 0$. At the origin the value of this function is $R(0) = 0$, and its derivative is $R'(0) = -(1 - p^t_{1|0}) - p^t_{1|0}B + (1 - p^t_{1|1})A + p^t_{1|1}AB$. We have

$$R'(0) < 0 \quad \text{if and only if} \quad p^t_{1|1} > \frac{A-1}{A-AB} + \frac{p^t_{1|0}}{A}. \quad (9.10)$$

Since $C_{t,s}$ is strictly increasing in s and $C_{t,0} = 0$, (9.10) shows that (3.4) holds for s in a right neighborhood of the origin if and only if (3.7) holds. Thus $s_1 = 0$ is equivalent to (3.7).

(d) If $s_1 = 0$ and $s_2 = \infty$ then (3.5), (3.6), and (3.7) hold by (b) and (c), and (3.8) follows from (3.7). To prove the converse note that (3.8) implies (3.5), and that (3.6) and (3.8) imply (3.7).

To prove the second part, define

$$D_t = \int_0^t h_0(u)\ \mathrm{d}u. \tag{9.11}$$

Then

$$\begin{aligned} p^t_{1|1} &= \frac{p^0_{1|1} P(T > t|X = 1, Y = 1)}{p^0_{1|1} P(T > t|X = 1, Y = 1) + p^0_{0|1} P(T > t|X = 1, Y = 0)} \\ &= \frac{p^0_{1|1} \exp\{-D_t AB\}}{p^0_{1|1} \exp\{-D_t AB\} + p^0_{0|1} \exp\{-D_t A\}} \\ &= \frac{p^0_{1|1}}{p^0_{1|1} + p^0_{0|1} \exp\{D_t A(B-1)\}}. \end{aligned}$$

The latter expression is increasing in t since $B < 1$. If (3.6) holds, then by (9.6) $p^t_{1|0} = 0$ for all t. Therefore the right hand side of (3.7) is constant, while the left hand side is increasing. It is easy to see that this implies that (3.7) holds for t in some upper interval. □

Proofs of Section 4

Proof of Proposition 4.1. We will show that the paradox can hold for small values of s by finding parameters in (2.3) that satisfy (3.2) and (3.7) for independent X and Y. The covariates X and Y are independent at time 0 if and only if $p^0_{1|0} = p^0_{1|1} =: p_1$. Hence, using (9.5), inequality (3.7) becomes

$$\begin{aligned} &\frac{p_1 P(T > t|X = 1, Y = 1)}{p_1 P(T > t|X = 1, Y = 1) + (1 - p_1) P(T > t|X = 1, Y = 0)} \geq \\ &\qquad \frac{A-1}{A - AB} + \frac{1}{A} \frac{p_1 P(T > t|X = 0, Y = 1)}{p_1 P(T > t|X = 0, Y = 1) + (1 - p_1) P(T > t|X = 0, Y = 0)}. \end{aligned} \tag{9.12}$$

By (2.3), (3.1), and (9.11), inequality (9.12) can be written as

$$\frac{p_1 \exp\{-D_t AB\}}{p_1 \exp\{-D_t AB\} + (1 - p_1) \exp\{-D_t A\}} \geq \frac{A-1}{A - AB} + \frac{1}{A} \frac{p_1 \exp\{-D_t B\}}{p_1 \exp\{-D_t B\} + (1 - p_1) \exp\{-D_t\}}. \tag{9.13}$$

An example of values of A, B, D_t , and p_1 that satisfy (9.13) and (3.2) is $A = 9$, $B = 0.05$, $p_1 = 0.3$, and any $-\log 0.5 < D_t < -\log 0.3$. □

Proofs of Section 5

Proof of Theorem 5.1. Property (1.2a) clearly holds since $\beta_X > 0$. To see (1.2b) consider

$$P(T > t+s|T > t, X = x) = \int P(T > t+s|T > t, X = x, Y = y) f_Y(y|T > t, X = x)\ \mathrm{d}y, \tag{9.14}$$

where

$$f_Y(y|T > t, X = x) = \frac{P(T > t|X = x, Y = y) f_Y(y|X = x)}{\int P(T > t|X = x, Y = y) f_Y(y|X = x)\ \mathrm{d}y}. \tag{9.15}$$

Using (2.3), (9.3), (9.11), and (9.15), expression (9.14) becomes

$$P(T > t+s|T > t, X = x) = \frac{\int \exp\{-(C_{t,s} + D_t)\exp\{\beta_X x + \beta_Y y\}\} f_Y(y|X = x)\ \mathrm{d}y}{\int \exp\{-D_t \exp\{\beta_X x + \beta_Y y\}\} f_Y(y|X = x)\ \mathrm{d}y}. \tag{9.16}$$

We can assume $\mu = 0$ by replacing the function $g(\cdot)$ by $g(\cdot - \mu)$. We want to show that (9.16) is strictly increasing in x, which, by (5.1), is equivalent to

$$K(D, x) = \int \exp\{-D\exp\{\beta_X x + \beta_Y y\}\} g(y - bx)\ \mathrm{d}y \tag{9.17}$$

being STP_2 in (D, x). The substitution $-u = \beta_X x + \beta_Y y$ leads to

$$K(D, x) = -\frac{1}{\beta_Y} \int \exp\{-D\exp\{-u\}\} g\left(-\frac{u}{\beta_Y} - \left(b + \frac{\beta_X}{\beta_Y}\right) x\right)\ \mathrm{d}u. \tag{9.18}$$

The term $\exp\{-D\exp\{-u\}\}$ is of the form $\exp\{\zeta(D)\xi(u)\}$, with ζ, ξ strictly decreasing functions, and is STP_2 in (D, u) by Proposition 9.3, because $\exp\{xy\}$ is STP_2 in (x, y). Since $\beta_Y < 0$, $(b + \beta_X/\beta_Y) > 0$, and g is strictly log-concave, we have that

$$g\left(-\frac{u}{\beta_Y} - \left(b + \frac{\beta_X}{\beta_Y}\right) x\right) \tag{9.19}$$

is STP_2 in (u, x). By Proposition 9.4 we obtain that $K(D, x)$ is STP_2 in (D, x).

If (5.2) does not hold, then different possibilities arise. If $\beta_X < 0$, then (1.2) cannot hold.

If $\beta_X > 0$ and $\beta_Y > 0$, then $(b + \beta_X/\beta_Y) > 0$, (9.19) is RR_2 in (u, x), and by Propositions 9.3 and 9.4, $K(D, x)$ is SRR_2 in (D, x) rather than STP_2. Similarly, if $\beta_X > 0$ and $\beta_Y < 0$, then in order for (9.19) to be STP_2 we must have $(b + \beta_X/\beta_Y) > 0$, so $b > \beta_X/|\beta_Y|$. □

References

Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993) *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

Anderson, G. L. and Fleming, T. R. (1995) Model misspecification in proportional hazards regression. *Biometrika* **82**, 527–541.

Blyth, C. R. (1972a) On Simpson's paradox and the sure-thing principle. *J. Amer. Statist. Assoc.* **67**, 364–366, 373–381.

Blyth, C. R. (1972b) Some probability paradoxes in choice from among random alternatives. *J. Amer. Statist. Assoc.* **67**, 366–373, 373–381.

Blyth, C. R. (1973) Simpson's paradox and mutually favorable events. *J. Amer. Statist. Assoc.* **68**, 746.

Breslow, N. E. (1972) Contribution to discussion of paper by D. R. Cox. *J. Roy. Statist. Soc. Ser. B* **34**, 216–217.

Breslow, N. E. (1974) Covariance analysis of censored survival data. *Biometrics* **30**, 89–99.

Bretagnolle, J. and Huber-Carol, C. (1985) Sous-estimation des contrastes due à l'oubli de variables pertinentes dans le modèle de Cox pour les durées de survie avec censure. *C. R. Acad. Sci. Paris Sér. I Math.* **300**, 359–362.

Bretagnolle, J. and Huber-Carol, C. (1988) Effects of omitting covariates in Cox's model for survival data. *Scand. J. Statist.* **15**, 125–138.

Brown, L. D., Johnstone, I. M., and MacGibbon, K. B. (1981) Variation diminishing transformations: a direct approach to total positivity and its statistical applications. *J. Amer. Statist. Assoc.* **76**, 824–832.

Chen, H. Y. (2002) Double-semiparametric method for missing covariates in Cox regression models. *J. Amer. Statist. Assoc.* **97**, 565–576.

Cohen, J. E. (1986) An uncertainty principle in demograpy and the unisex issue. *Amer. Statist.* **40**, 32–39.

Cohen, M. R. and Nagel, E. (1934) *An Introduction to Logic and Scientific Method.* Harcourt, New York.

Cox, D. R. (1972) Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* **34**, 187–220.

Di Serio, C. (1997) The protective impact of a covariate on competing failures with an example from a bone marrow transplantation study. *Lifetime Data Analysis* **3**, 99–122.

Di Serio, C. (2003) Risk evaluation in dependent competing risk frameworks. Mimeo.

DiRienzo, A. G. and Lagakos, S. W. (2001a) Bias correction for score tests arising from misspecified proportional hazards regression models. *Biometrika* **88**, 421–434.

DiRienzo, A. G. and Lagakos, S. W. (2001b) Effects of model misspecification on tests of no randomized treatment effect arising from Cox's proportional hazards model. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63**, 745–757.

DiRienzo, A. G. and Lagakos, S. W. (2004) The effects of misspecifying Cox's regresssion model on randomized treatment group comparisons. In *Advances in Survival Analysis*, volume 23 of *Handbook of Statist.*, 395–409. Elsevier, Amsterdam.

Everitt, B. S. and Dunn, G. (1998) *Statistical Analysis of Medical Data: New Developments*. Arnold, London.

Gail, M. H., Wieand, S., and Piantadosi, S. (1984) Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* **71**, 431–444.

Gerds, T. A. and Schumacher, M. (2001) On functional misspecification of covariates in the Cox regression model. *Biometrika* **88**, 572–580.

Good, I. J. and Mittal, Y. (1987) The amalgamation and geometry of two-by-two contingency tables. *Ann. Statist.* **15**, 694–711.

Gouget, N. and Raoult, J.-P. (1999) Computation of predictive densities in the semiparametric Bayesian Cox-Dirichlet model. *J. Nonparametr. Statist.* **10**, 307–341.

Grambsch, P. M. and Therneau, T. M. (1994) Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 515–526.

Hougaard, P. (1984) Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika* **71**, 75–83.

Hougaard, P. (1986) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 387–396.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001) *Bayesian Survival Analysis*. Springer-Verlag, New York.

Joe, H. (1997) *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Karlin, S. (1968) *Total Positivity. Vol. I.* Stanford University Press, Stanford, Calif.

Lagakos, S. W. and Schoenfeld, D. A. (1984) Properties of proportional-hazards score tests under misspecified regression models. *Biometrics* **40**, 1037–1048.

Lagakos, S. W. and Schoenfeld, D. A. (1986) Reply to "Omitting covariates from the proportional hazards model". *Biometrics* **42**, 995–995.

Lin, D. Y. and Wei, L. J. (1989) The robust inference for the Cox proportional hazards model. *J. Amer. Statist. Assoc.* **84**, 1074–1078.

Lindley, D. V. and Novick, M. R. (1981) The role of exchangeability in inference. *Ann. Statist.* **9**, 45–58.

Mittal, Y. (1991) Homogeneity of subpopulations and Simpson's paradox. *J. Amer. Statist. Assoc.* **86**, 167–172.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi Sergi, L., Benedicenti, F., Ambrosi, A., Di Serio, C., Doglioni, C., von Kalle, C., and Naldini, L. (2006) Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat. Biotechnol.* **24**, 687–696.

Morgan, T. M. (1986) Omitting covariates from the proportional hazards model. *Biometrics* **42**, 993–995.

Rinott, Y. and Tam, M. (2003) Monotone regrouping, regression, and Simpson's paradox. *Amer. Statist.* **57**, 139–141.

Samuels, M. L. (1993) Simpson's paradox and related phenomena. *J. Amer. Statist. Assoc.* **88**, 81–88.

Sane, M. and Kharshikar, A. V. (2001) Effect of missing an influential covariate. *Comm. Statist. Theory Methods* **30**, 837–853.

Scarsini, M. and Spizzichino, F. (1999) Simpson-type paradoxes, dependence, and ageing. *J. Appl. Probab.* **36**, 119–131.

Simpson, E. H. (1951) The interpretation of interaction in contingency tables. *J. Roy. Statist. Soc. Ser. B.* **13**, 238–241.

Solomon, P. J. (1984) Effect of misspecification of regression models in the analysis of survival data. *Biometrika* **71**, 291–298.

Struthers, C. A. and Kalbfleisch, J. D. (1986) Misspecified proportional hazard models. *Biometrika* **73**, 363–369.

Yule, G. H. (1903) Notes on the theory of association of attributes in statistics. *Biometrika* **2**, 121–134.

January 3, 2007