IDENTIFYING NONPROPORTIONAL COVARIATES IN THE COX MODEL

David Kraus

Institute of Information Theory and Automation

Pod Vodárenskou věží 4

182 08 Praha 8

Czechia

david.kraus@matfyz.cz

Key Words: Cox model; goodness of fit; proportional hazards assumption; time-varying coefficients.

ABSTRACT

The problem of testing whether an individual covariate in the Cox model has a proportional (i.e., time-constant) effect on the hazard is dealt with. Two existing methods are considered: one is based on the component of the score process and the other is a Neyman type smooth test. Simulations show that when the model contains both proportional and nonproportional covariates, these methods are not reliable tools for discrimination. A simple, yet effective solution is proposed based on smooth modeling of the effects of the covariates not in focus.

1. INTRODUCTION

Consider the Cox proportional hazards regression model (Cox, 1972) for right-censored survival data in the form

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\{\beta^\mathsf{T} Z_i\} \tag{1}$$

(Andersen and Gill, 1982). Here $\lambda_i(t)$ is the intensity process of the i-th component of an n-variate counting process $N(t) = (N_1(t), \ldots, N_n(t))^\mathsf{T}$, $t \in [0, \tau]$, $Y_i(t)$ denotes the risk indicator process, Z_i is a p-vector of covariates, $\lambda_0(t)$ stands for an unknown baseline hazard function, and β is a vector of unknown regression coefficients.

In the Cox model the key assumption is proportionality of the effects of the covariates which means that the hazard ratio for two individuals does not depend on time. The assumption is not satisfied, for instance, when some of the coefficients $\beta_1, \ldots, \beta_p$ varies with time.

The aim of the paper is to study methods of assessment of the proportional hazards assumption for a single covariate, say the p-th one. More specifically, we wish to test the hypothesis that the coefficient β_p is constant against the alternative of time-varying $\beta_p(t)$.

The problem of existing methods (namely score process based tests and Neyman-type smooth tests, described later on) is that they cannot distinguish reliably which covariates are proportional and which not. In models with both proportional and nonproportional covariates, the hypothesis of proportionality is often rejected even for the proportional covariate, that is, the size of the test dramatically exceeds the nominal level. Therefore, they serve as a tool for the overall assessment of proportionality rather than for individual covariate checks. I propose an improvement that consists of modeling the effects of the other covariates, which are not of interest, as linear combinations of some smooth functions. This makes the test more precise in identifying nonproportional covariates.

Section 2 surveys some of existing methods for the proportional hazards problem. Simulation results of Section 3 warn against the use of the methods for testing proportionality of individual covariates. In Section 4 I present my solution whose performance is investigated through simulations in Section 5. The conclusions are summarised in Section 6.

2. EXISTING METHODS

2.1 TESTS BASED ON THE SCORE PROCESS

Lin et al. (1993) proposed to base tests on the score process

$$U(t;\hat{\beta}) = \sum_{i=1}^{n} \int_0^t Z_i dN_i(s) - \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i \exp\{\hat{\beta}^\mathsf{T} Z_i\}}{\sum_{i=1}^n Y_i(s) \exp\{\hat{\beta}^\mathsf{T} Z_i\}} d\bar{N}(s),$$

where $\bar{N} = \sum_{i=1}^n N_i$ and $\hat{\beta}$ is the maximum partial likelihood estimate of β, i.e., the solution to $U(\tau;\hat{\beta}) = 0$. Each component of the process reflects deviations from proportionality of the respective covariate. If the assumption is satisfied, the component fluctuates around zero, otherwise it differs. Thus, one can either perform a global test of proportionality (using the whole vector process) or assess validity of the assumption for an individual covariate (using a single component). Lin et al. (1993) use tests of the Kolmogorov–Smirnov type based on suprema of test processes. Other functionals, mainly those leading to the test of the Anderson–Darling and Cramér–von Mises type, may be used.

The asymptotic distribution of the process $n^{-1/2}U(\cdot;\hat{\beta})$ is generally intractable (the limit is a Gaussian process which is neither a martingale nor a well-known process). Therefore, Lin et al. (1993) proposed a simulation technique to approximate this distribution. The technique is now commonly used. Alternatively, one may use the martingale transformation of Khmaladze (1981), see Marzec and Marzec (1997) for its application to the score process.

2.2 SMOOTH TESTS

Another way of testing proportionality is based on the parametric modeling of the time-varying effect. The procedure consists of introducing new artificial time-dependent covariates and testing their significance. The original idea dates back to Cox (1972) who used one new covariate. See also Grambsch and Therneau (1994). Here I shall work with an extension described in Kraus (2007) which is inspired by Neyman's smooth goodness-of-fit tests. The original model (1) is embedded in the k-dimensional model

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\{\beta^\mathsf{T} Z_i + \theta^\mathsf{T}\varphi(F_0(t)/F_0(\tau))Z_{ip}\}, \tag{2}$$

where F_0 is the distribution function associated with the baseline hazard λ_0 (in practice, F_0 is replaced by an estimator $\hat{F}_0$), and $\varphi = (\varphi_1, \ldots, \varphi_k)^\mathsf{T}$ are some bounded functions in $L_2[0,1]$ such that $\{1, \varphi_1, \ldots, \varphi_k\}$ is a set of linearly independent functions (for instance, the orthonormal Legendre polynomials on $[0,1]$ or the cosine basis $\varphi_j(u) = \sqrt{2}\cos(\pi j u)$).

The smooth test of proportionality is then a test of significance of the new covariates $\varphi_1(F_0(t)/F_0(\tau))Z_{ip}, \ldots, \varphi_d(F_0(t)/F_0(\tau))Z_{ip}$, i.e., a test of $\theta = 0$ against $\theta \neq 0$. The partial likelihood score test is most convenient as it does not involve estimating θ. The test statistic is asymptotically χ^2-distributed with k degrees of freedom. A data-driven version using a BIC-like selection rule for the choice of k has been considered in Kraus (2007).

3. WARNING AGAINST INDIVIDUAL COVARIATE TESTS

It may be misleading and dangerous to draw conclusions about proportionality of individual covariates from the tests of the previous section in models with multiple covariates. To illustrate this, I performed a simulation study.

Let us consider three models, both of which have two covariates. In the first model of the form

$$\lambda(t) = \exp\{0.7Z_1 + 0.3Z_2\} \tag{3}$$

both covariates have proportional effects. The other two models have one covariate (Z_1) with a nonproportional effect and one (Z_2) proportional. The models follow the form

$$\lambda(t) = \exp\{0.5tZ_1 + Z_2 - 8\} \tag{4}$$

and

$$\lambda(t) = \exp\{\beta(t)Z_1 + Z_2 - 8\}, \tag{5}$$

where $\beta(t) = 0.4 + 0.7 \times 1_{[1.2,2]}(t)$. The coefficient of Z_1 in (4) is monotonic, whereas in (5) it is not. In both models the variables Z_1, Z_2 are jointly normal with expectation 4, variance 1 and various values of correlation ρ. Censoring times were $\mathrm{U}(0,5)$ distributed in the models (3) and (4) (giving for all of the values of correlation the censoring rate about 24 % and 45 % in (3) and (4), respectively) and constant equal to 5 in the model (5) (about 31 % censoring).

I repeatedly generated samples of $n = 200$ observations and estimated rejection probabilities. The number of Monte Carlo runs in each situation was 5000 giving estimates of rejection probabilities with standard deviation at most $\sqrt{0.5 \times (1 - 0.5)/5000} = 0.007$. The smooth test (denoted in the tables as T_k) was used with the dimension $k = 3, 4, 5, 6$. The Kolmogorov–Smirnov (KS), Cramér–von Mises (CM), and Anderson–Darling test were performed using the simulation approximation of Lin et al. (1993) (with 1000 simulated paths of the score process).

Results for the model (3) are reported in Table 1 (results for Z_1 are given only, for Z_2 they are similar). The null hypothesis of proportionality is satisfied and the estimates of

Table 1: Estimated rejection probabilities on the nominal level of 5 % for the covariate Z_1 in the model $\lambda(t) = \exp\{0.7Z_1 + 0.3Z_2\}$ with $\mathrm{cor}(Z_1, Z_2) = \rho$.

	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$
T_3	0.054	0.057	0.055	0.050	0.055
T_4	0.052	0.057	0.060	0.050	0.059
T_5	0.053	0.055	0.055	0.054	0.059
T_6	0.060	0.055	0.055	0.056	0.059
KS	0.054	0.051	0.053	0.050	0.052
CM	0.052	0.044	0.050	0.049	0.052
AD	0.050	0.043	0.046	0.047	0.050

rejection probabilitites are close to the nominal level of 5 %, so everything seems to be all right.

Results for the models (4) and (5) are shown in Tables 2 and 3 (results for Z_2 which satisfies the null hypothesis are reported only). As the hypothesis of proportionality of Z_2 is valid, the figures in Tables 2 and 3 should be close to the nominal level of 5 %. However, it turns out that in some cases the level is dramatically exceeded. Some of the figures are really alarming, especially (but not only) in the case of highly associated covariates.

The reason is that the score process method and the smooth method are valid only under the assumption of time-constancy of the effects of all the other covariates. When proportionality is violated for some (nuisance) covariate that is not of interest, the techniques become unreliable. The procedures can indicate that proportionality is not valid but are not capable to distinguish which covariate is 'guilty' and which not. This phenomenon has been previously pointed out, e.g., by Scheike and Martinussen (2004, pp. 58–59). My simulations show that the problem is serious even if the covariates are independent which was not seen so markedly in their simulation study.

4. IMPROVEMENT

Table 2: Estimated rejection probabilities on the nominal level of 5 % for the proportional covariate Z_2 in the model $\lambda(t) = \exp\{0.5tZ_1 + Z_2 - 8\}$ with $\mathrm{cor}(Z_1, Z_2) = \rho$.

	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$
T_3	0.136	0.070	0.063	0.118	0.265
T_4	0.125	0.068	0.067	0.116	0.231
T_5	0.116	0.066	0.064	0.105	0.212
T_6	0.112	0.067	0.065	0.098	0.190
KS	0.149	0.061	0.085	0.181	0.382
CM	0.157	0.057	0.080	0.192	0.430
AD	0.149	0.053	0.076	0.186	0.425

Table 3: Estimated rejection probabilities on the nominal level of 5 % for the proportional covariate Z_2 in the model $\lambda(t) = \exp\{\beta(t)Z_1 + Z_2 - 8\}$ ($\beta(t) = 0.4 + 0.7 \times 1_{[1.2,2]}(t)$) with correlation ρ between Z_1 and Z_2.

	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$
T_3	0.091	0.068	0.086	0.152	0.297
T_4	0.085	0.060	0.078	0.142	0.276
T_5	0.086	0.060	0.072	0.136	0.256
T_6	0.078	0.058	0.075	0.128	0.244
KS	0.055	0.069	0.099	0.210	0.392
CM	0.052	0.068	0.101	0.200	0.373
AD	0.053	0.066	0.095	0.188	0.341

The tests for individual covariates work correctly provided the model is correct for all the other covariates. Therefore, a simple idea seems to be reasonable: make the model for all the other covariates correct enough to remove or diminish the influence of their potential nonproportionality on the test. This means to model the time-varying effects of the other covariates. To this end, I use smooth functions similarly as in Neyman's smooth tests.

Recall that the aim is to test proportionality of the p-th covariate.

Instead of the original model

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\left\{\sum_{j=1}^{p} Z_{ij}\beta_j\right\}, \tag{6}$$

the null model now follows the form

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\left\{\sum_{j=1}^{p-1} Z_{ij}\left(\beta_j + \sum_{k=1}^{d_j}\theta_{jk}\varphi_k(F_0(t)/F_0(\tau))\right) + Z_{ip}\beta_p\right\}, \tag{7}$$

which allows for smoothly time-varying coefficients of all the covariates but the p-th one. This large model is an ordinary Cox model with artificial time-dependent covariates and with parameters $\beta_1, \dots, \beta_p$ and $\theta_{jk},\ j = 1, \dots, p-1,\ k = 1, \dots, d_j$. The unknown time-transformation F_0 is estimated from the original model (6) and viewed as given when working with the large model (7). That is we in fact work with

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\left\{\sum_{j=1}^{p-1} Z_{ij}\left(\beta_j + \sum_{k=1}^{d_j}\theta_{jk}\varphi_k(\hat{F}_0(t)/\hat{F}_0(\tau))\right) + Z_{ip}\beta_p\right\}. \tag{8}$$

The proposed test procedure is as follows. First, one estimates the coefficients in the large model (8) by the standard partial likelihood method and then performs some of the tests mentioned in Section 2. That is either the test based on the component of the score process corresponding to Z_{ip} in (8), or the smooth test which is the score test of significance of $\theta_{p1}, \dots, \theta_{p,d_p}$ in

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\left\{\sum_{j=1}^{p-1} Z_{ij}\left(\beta_j + \sum_{k=1}^{d_j}\theta_{jk}\varphi_k(\hat{F}_0(t)/\hat{F}_0(\tau))\right) + Z_{ip}\left(\beta_p + \sum_{k=1}^{d_p}\theta_{pk}\varphi_k(F_0(t)/F_0(\tau))\right)\right\}.$$

Here in the Z_{ip}-part F_0 is replaced either by $\hat{F}_0$ (the same as for the $Z_{i1}, \ldots, Z_{i,p-1}$-part) or by $\tilde{F}_0$ computed from (8) (which is closer to the true distribution).

The tests are carried out in the same way as in the original versions of the tests: for the score process based test the simulation approximation can be used, and the distribution of the statistic of the smooth test is approximated by the $\chi^2_{d_p}$ distribution. A data-driven choice of d_p is possible.

The proposed approach practically works as is observed in the simulation study in the next section. To justify it theoretically one would have to let $d_1, \ldots, d_{p-1}$ tend to infinity at a suitable rate as n grows. The convergence must be fast enough to control the approximation error but not too fast to guarantee stability of estimation. It calls for further research to give conditions under which the Z_{ip}-component of the score process computed in (7) converges to a zero-mean Gaussian process. A similar problem was previously considered by Murphy and Sen (1991) who dealt with a histogram sieve estimator in the Cox model with time-varying coefficients.

5. SIMULATIONS

The aim of the Monte Carlo study is to explore whether the proposed improvement works (i.e., whether the level is preserved) and how it influences the power.

The design of the study is the same as in Section 3. The tests were used with various numbers of smooth functions describing the effect of the covariate that is not tested (various values of d_2 for tests of Z_1 and d_1 for Z_2). Results for the models (3), (4) and (5) are displayed in Tables 4, 5 and 6.

Table 4 shows that the proposed modification works correctly in the situation where even the original test (without smooth modeling of the other covariates) was valid. In Tables 5 and 6, results for Z_2 show that the prescribed level of 5 % is preserved when the effect of the covariate Z_1 is modeled smoothly. In the models of the study, it was enough to use two smooth functions because the time-varying coefficient of Z_1 has a relatively simple form in both models. Generally, it may be necessary to use more basis functions.

The power results for Z_1 show that if one uses more smooth functions than necessary,

Table 4: Estimated rejection probabilities on the nominal level 5 % in the model $\lambda(t) = \exp\{0.7Z_1 + 0.3Z_2\}$ with $\mathrm{cor}(Z_1, Z_2) = \rho$. Various numbers of smooth functions for the other covariate.

		Z_1				Z_2			
		$d_2 = 0$	$d_2 = 2$	$d_2 = 3$	$d_2 = 4$	$d_1 = 0$	$d_1 = 2$	$d_1 = 3$	$d_1 = 4$
$\rho = 0$	T_3	0.054	0.057	0.058	0.058	0.056	0.060	0.062	0.063
	T_4	0.052	0.061	0.064	0.063	0.057	0.059	0.059	0.060
	T_5	0.053	0.061	0.065	0.064	0.054	0.058	0.060	0.062
	T_6	0.060	0.062	0.064	0.065	0.057	0.060	0.060	0.062
	KS	0.054	0.057	0.055	0.053	0.051	0.051	0.050	0.050
	CM	0.052	0.053	0.052	0.050	0.049	0.047	0.049	0.046
	AD	0.050	0.050	0.049	0.048	0.048	0.047	0.048	0.046
$\rho = 0.5$	T_3	0.055	0.063	0.064	0.066	0.054	0.059	0.063	0.064
	T_4	0.060	0.065	0.068	0.066	0.052	0.056	0.057	0.060
	T_5	0.055	0.062	0.064	0.066	0.056	0.059	0.059	0.060
	T_6	0.055	0.061	0.061	0.062	0.054	0.058	0.058	0.056
	KS	0.053	0.055	0.056	0.056	0.057	0.060	0.059	0.058
	CM	0.050	0.055	0.057	0.054	0.053	0.052	0.049	0.051
	AD	0.046	0.052	0.054	0.055	0.050	0.046	0.047	0.050
$\rho = 0.9$	T_3	0.055	0.058	0.055	0.059	0.048	0.051	0.057	0.058
	T_4	0.059	0.058	0.060	0.058	0.054	0.054	0.058	0.061
	T_5	0.059	0.060	0.061	0.061	0.051	0.051	0.056	0.057
	T_6	0.059	0.063	0.064	0.064	0.056	0.054	0.054	0.055
	KS	0.052	0.053	0.054	0.051	0.056	0.052	0.053	0.049
	CM	0.052	0.050	0.051	0.052	0.050	0.045	0.051	0.049
	AD	0.050	0.047	0.048	0.050	0.048	0.041	0.043	0.046

Table 5: Estimated rejection probabilities on the nominal level 5 % in the model $\lambda(t) = \exp\{0.5tZ_1 + Z_2 - 8\}$ with $\mathrm{cor}(Z_1, Z_2) = \rho$. Various numbers of smooth functions for the other covariate.

		Z_1				Z_2			
		$d_2 = 0$	$d_2 = 2$	$d_2 = 3$	$d_2 = 4$	$d_1 = 0$	$d_1 = 2$	$d_1 = 3$	$d_1 = 4$
$\rho = 0$	T_3	0.771	0.690	0.685	0.683	0.136	0.061	0.060	0.062
	T_4	0.726	0.644	0.638	0.640	0.125	0.062	0.059	0.061
	T_5	0.698	0.611	0.612	0.609	0.116	0.064	0.059	0.061
	T_6	0.663	0.577	0.574	0.576	0.112	0.059	0.055	0.059
	KS	0.797	0.738	0.735	0.736	0.149	0.048	0.050	0.050
	CM	0.855	0.808	0.809	0.810	0.157	0.047	0.048	0.047
	AD	0.861	0.813	0.817	0.814	0.149	0.045	0.047	0.046
$\rho = 0.5$	T_3	0.657	0.626	0.625	0.615	0.063	0.063	0.062	0.059
	T_4	0.605	0.582	0.577	0.572	0.067	0.060	0.060	0.060
	T_5	0.566	0.536	0.534	0.533	0.064	0.059	0.061	0.060
	T_6	0.529	0.494	0.498	0.494	0.065	0.056	0.059	0.061
	KS	0.698	0.675	0.672	0.674	0.085	0.060	0.053	0.052
	CM	0.774	0.754	0.748	0.746	0.080	0.055	0.051	0.049
	AD	0.777	0.758	0.755	0.754	0.076	0.053	0.048	0.047
$\rho = 0.9$	T_3	0.467	0.238	0.235	0.225	0.265	0.071	0.063	0.061
	T_4	0.414	0.210	0.200	0.200	0.231	0.068	0.066	0.062
	T_5	0.374	0.181	0.182	0.179	0.212	0.064	0.063	0.061
	T_6	0.341	0.164	0.174	0.165	0.190	0.061	0.059	0.059
	KS	0.554	0.266	0.254	0.263	0.382	0.075	0.059	0.057
	CM	0.633	0.340	0.325	0.334	0.430	0.070	0.054	0.052
	AD	0.641	0.336	0.323	0.335	0.425	0.061	0.049	0.048

Table 6: Estimated rejection probabilities on the nominal level 5 % in the model $\lambda(t) = \exp\{\beta(t)Z_1 + Z_2 - 8\}$ ($\beta(t) = 0.4 + 0.7 \times 1_{[1.2,2]}(t)$) with $\mathrm{cor}(Z_1, Z_2) = \rho$. Various numbers of smooth functions for the other covariate.

		Z_1				Z_2			
		$d_2 = 0$	$d_2 = 2$	$d_2 = 3$	$d_2 = 4$	$d_1 = 0$	$d_1 = 2$	$d_1 = 3$	$d_1 = 4$
$\rho = 0$	T_3	0.662	0.569	0.547	0.569	0.091	0.051	0.053	0.052
	T_4	0.676	0.630	0.608	0.612	0.085	0.052	0.052	0.048
	T_5	0.651	0.612	0.595	0.594	0.086	0.051	0.056	0.049
	T_6	0.634	0.599	0.587	0.579	0.078	0.053	0.059	0.052
	KS	0.596	0.605	0.612	0.617	0.055	0.053	0.052	0.049
	CM	0.525	0.541	0.538	0.546	0.052	0.049	0.047	0.048
	AD	0.559	0.567	0.563	0.567	0.053	0.047	0.048	0.046
$\rho = 0.5$	T_3	0.570	0.539	0.499	0.483	0.086	0.050	0.052	0.052
	T_4	0.569	0.529	0.497	0.485	0.078	0.052	0.048	0.052
	T_5	0.534	0.511	0.481	0.467	0.072	0.051	0.050	0.050
	T_6	0.516	0.495	0.464	0.449	0.075	0.054	0.053	0.052
	KS	0.603	0.577	0.559	0.559	0.099	0.053	0.052	0.051
	CM	0.568	0.541	0.527	0.525	0.101	0.053	0.049	0.051
	AD	0.553	0.521	0.510	0.505	0.095	0.050	0.049	0.048
$\rho = 0.9$	T_3	0.452	0.200	0.172	0.149	0.297	0.057	0.058	0.060
	T_4	0.436	0.200	0.179	0.145	0.276	0.060	0.068	0.064
	T_5	0.412	0.200	0.177	0.152	0.256	0.065	0.068	0.065
	T_6	0.394	0.195	0.175	0.150	0.244	0.063	0.067	0.062
	KS	0.541	0.261	0.242	0.238	0.392	0.075	0.063	0.057
	CM	0.522	0.261	0.239	0.233	0.373	0.060	0.055	0.053
	AD	0.481	0.209	0.196	0.191	0.341	0.051	0.052	0.053

the decrease of power is not dramatic. The factor that prevents us from including very large numbers of basis functions (i.e., new artificial covariates) is the sample size.

6. SUMMARY

I have shown that it is not advisable to use tests derived under the proportional hazards assumption for testing proportionality of individual covariates. The remedy based on modeling possibly time-varying effects of the other covariates which are not in focus has been shown to work by simulations.

ACKNOWLEDGEMENTS

The author gratefully acknowledges that the work has been supported by the GAČR grant 201/05/H007, GAAV grant IAA101120604 and MŠMT project 1M06047. Computations have been carried out in METACentrum (Czech academic supercomputer network).

BIBLIOGRAPHY

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.*, 10, 1100–1120.

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

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

Khmaladze, E. V. (1981). Martingale approach in the theory of goodness-of-fit tests. *Teor. Veroyatnost. i Primenen.*, 26, 246–265. In Russian. English translation in *Theory Probab. Appl.*, 26, 240–257.

Kraus, D. (2007). Data-driven smooth tests of the proportional hazards assumption. *Lifetime Data Anal.*, 13, 1–16.

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika*, 80, 557–572.

Marzec, L. and Marzec, P. (1997). Generalized martingale-residual processes for goodness-of-fit inference in Cox's type regression models. *Ann. Statist.*, 25, 683–714.

Murphy, S. A. and Sen, P. K. (1991). Time-dependent coefficients in a Cox-type regression model. *Stochastic Process. Appl.*, 39, 153–180.

Scheike, T. H. and Martinussen, T. (2004). On estimation and tests of time-varying effects in the proportional hazards model. *Scand. J. Statist.*, 31, 51–62.