Supplementary Material to:

A permutation test for the two-sample right-censored model

by Grzegorz Wyłupek, Institute of Mathematics, University of Wrocław

APPENDIX A: DERIVATION OF $\ell_j^0(t)$, $j = 1, 2, 3, \ldots$

Recall that, by (7), (9), and (10), we have

$$L_j(x) = \frac{e^x}{j!}\frac{d^j}{dx^j}(x^j e^{-x}) = \sum_{k=0}^{j} \frac{(-1)^k}{k!}\binom{j}{k} x^k, \quad j = 0, 1, 2, \ldots, \quad x \in [0, +\infty),$$

$$\ell_j(t) = L_j(-\log(1-t)) = \sum_{k=0}^{j} \frac{1}{k!}\binom{j}{k} [\log(1-t)]^k, \quad j = 0, 1, 2, \ldots, \quad t \in [0, 1),$$

$$\ell_j^0(t) = \ell_j(t) - \frac{1}{1-t}\int_t^1 \ell_j(s)ds, \quad j = 0, 1, 2, \ldots.$$

Then, for $j = 1, 2, 3, \ldots,$

$$\int_t^1 \ell_j(s)ds = \int_{-\log(1-t)}^{\infty} L_j(x)e^{-x}dx = \frac{1}{j!}\int_{-\log(1-t)}^{\infty} \frac{d^j}{dx^j}(x^j e^{-x})dx =$$

$$= \frac{1}{j!}\int_{-\log(1-t)}^{\infty} \left[\frac{d^{j-1}}{dx^{j-1}}(x^j e^{-x})\right]' dx = -\frac{1}{j!}\left[\frac{d^{j-1}}{dx^{j-1}}(x^j e^{-x})\right]_{x=-\log(1-t)} =$$

$$= [1-t]\sum_{k=0}^{j-1} \frac{1}{(k+1)!}\binom{j-1}{k}[\log(1-t)]^{k+1} = [1-t]\sum_{k=1}^{j} \frac{1}{k!}\binom{j-1}{k-1}[\log(1-t)]^k,$$

resulting in

$$\ell_j^0(t) = \sum_{k=0}^{j} \frac{1}{k!}\binom{j}{k}[\log(1-t)]^k - \sum_{k=1}^{j} \frac{1}{k!}\binom{j-1}{k-1}[\log(1-t)]^k = 1 + \sum_{k=1}^{j} \frac{1}{k!}\left[\binom{j}{k} - \binom{j-1}{k-1}\right][\log(1-t)]^k =$$

$$= 1 + \sum_{k=1}^{j-1} \frac{1}{k!}\binom{j-1}{k}[\log(1-t)]^k = \sum_{k=0}^{j-1} \frac{1}{k!}\binom{j-1}{k}[\log(1-t)]^k = \ell_{j-1}(t).$$

APPENDIX B: ADDITIONAL INTERPRETATION OF THE *WLR* STATISTIC

Hereunder, we present another possible interpretation of $\mathcal{L}_j$. Namely, we will express it in terms of the weighted differences between observed and conditionally expected failures (cf., for instance, Tarone and Ware, 1977, p. 158). For this purpose, let $X_{1:1}, \ldots, X_{n_1:n}, X_{n_1+1:n}, \ldots, X_{n:n}$ be the order statistics and $\Delta_{1:1}, \ldots, \Delta_{n_1:n}, \Delta_{n_1+1:n}, \ldots, \Delta_{n:n}$ be their concomitants, which are also called the induced order statistics (i.e., $\Delta_{i:n}$ is the censoring status of $X_{i:n}$). Let O_{1i} be the observed numbers of failures in the first group at $X_{i:n}$ and $E_{1i} = (O_{1i} + O_{2i})\,Y_1(X_{i:n})/Y(X_{i:n})$ be their indirect standardization, where O_{2i} is defined similarly to O_{1i}, $i = 1, \ldots, n$. Actually, E_{1i} is the expectation of the hyper-geometric distribution with the parameters $(1, Y_1(X_{i:n}), Y(X_{i:n}))$, which is the conditional distribution of O_{1i} given $O_i = O_{1i} + O_{2i}$, $Y_1(X_{i:n})$, and $Y(X_{i:n})$. It means that E_{1i} is strictly related to the contingency table (see Table S1) corresponding to $X_{i:n}$ such that $\Delta_{i:n} = 1$, $1 \le i \le n$.

Table S1 Contingency table corresponding to the uncensored observation $X_{i:n}$

	Sample		
Failure	1	2	Total
Yes	O_{1i}	O_{2i}	$O_{1i} + O_{2i} = 1$
No	$Y_1(X_{i:n}) - O_{1i}$	$Y_2(X_{i:n}) - O_{2i}$	$Y(X_{i:n}) - 1$
Total	$Y_1(X_{i:n})$	$Y_2(X_{i:n})$	$Y(X_{i:n})$

Then, by (5) and (12),

$$\mathcal{L}_j = \sqrt{\frac{n}{n_1 n_2}} \left\{ \sum_{i=1}^{n_1} \int_0^\infty w_j(x) \frac{Y_2(x)}{Y(x)}\, dN_{1i}(x) - \sum_{i=1}^{n_2} \int_0^\infty w_j(x) \frac{Y_1(x)}{Y(x)}\, dN_{2i}(x) \right\} =$$

$$= \sqrt{\frac{n}{n_1 n_2}} \left\{ \sum_{i=1}^{n_1} \int_0^\infty w_j(x)\, dN_{1i}(x) - \sum_{l=1}^{2} \sum_{i=1}^{n_l} \int_0^\infty w_j(x) \frac{Y_1(x)}{Y(x)}\, dN_{li}(x) \right\} =$$

$$= \sqrt{\frac{n}{n_1 n_2}} \left\{ \sum_{i=1}^{n} w_j(X_{i:n})\, O_{1i} - \sum_{i=1}^{n} w_j(X_{i:n}) \frac{Y_1(X_{i:n})}{Y(X_{i:n})} [O_{1i} + O_{2i}] \right\} =$$

$$= \sqrt{\frac{n}{n_1 n_2}} \sum_{i=1}^{n} w_j(X_{i:n}) [O_{1i} - E_{1i}] = \sqrt{\frac{n}{n_1 n_2}} \sum_{i=1}^{n} \ell_j^0(\hat{F}(X_{i:n}))\, \mathbb{1}\big(X_{i:n} \le X_n^\star\big) [O_{1i} - E_{1i}], \qquad (S1)$$

where $X_n^\star = \min\{\max_{1 \le i \le n_1} X_i, \max_{n_1 < i \le n} X_i\}$. In that sense, $\mathcal{L}_j$ is the rescaled sum of the weighted differences between observed and conditionally expected failures, while the number of not vanishing summands is the number of the uncensored observations in both samples smaller than or equal to $X_n^\star$, and thereby, the number of distinct failure times in the pooled sample, which does not exceed $X_n^\star$.

APPENDIX C: PROOFS

Proof of Theorem 1.

First, we shall show that $C_j \xrightarrow{\mathcal{D}} N(0,1)$, $j = 1, 2, 3 \ldots$. For this purpose, we exploit part (a) of Corollary 2.5, p. 52, Janssen and Neuhaus (1997). Recall that $F = \eta F_1 + (1-\eta) F_2$, with $\eta = \lim_{n\to\infty}(n_1/n)$, $\eta \in (0,1)$, $H(x) = 1 - \eta[1 - F_1(x)][1 - G_1(x)] - (1-\eta)[1 - F_2(x)][1 - G_2(x)]$, $x \in [0, \infty)$, and $H^{-1}(t) = \inf\{\, x : H(x) \ge t \,\}$, $t \in [0, 1)$. Since the assumptions of Theorem 2.2 (a), ibidem, are satisfied with the weight $w(\cdot) = \ell_j(F(H^{-1}(\cdot)))$, cf. (10.40), (10.42) in supplement to Brendel et al. (2014), and the assertions of part (a) of the Corollary also hold, we obtain that $C_j \xrightarrow{\mathcal{D}} N(0,1)$, $j = 1, 2, 3 \ldots$.

Now, it is enough to show that $J(n,c) \xrightarrow{\mathcal{P}} 1$ for any fixed $c>0$, and that $S \xrightarrow{\mathcal{P}} 1$. For this purpose, by (19),

$$P(\,J(n,c)=0\,) = P(\max_{1\le j\le d}\big|C_j\big| > \sqrt{c\log n}\,) \le \sum_{j=1}^{d} P(\,\big|C_j\big| > \sqrt{c\log n}\,) \to 0,\ \text{as } n\to\infty,$$

because d is fixed, while $C_j = O_P(1)$. The same argument entails that

$$P(S\ge 2) = P(\cup_{k=2}^{d}\{S=k\}) = P(\cup_{k=2}^{d}\{W_k - k\log n > W_1 - \log n\}) \le$$

$$\le \sum_{k=2}^{d} P\Big(\sum_{j=2}^{k} C_j^2 > (k-1)\log n\Big) \to 0,\ \text{as } n\to\infty,$$

and completes the proof.

Proof of Theorem 2.

Let j_0 be the first index such that $\gamma_{j_0} := \int_0^\tau \ell_{j_0}^0(F(x))\,\frac{\pi_1(x)\,\pi_2(x)}{\pi(x)} d[\Lambda_1(x)-\Lambda_2(x)] \ne 0$. Then $\big|\mathcal{L}_{j_0}\big| = \sqrt{n_1 n_2/n}\,\big|\hat\gamma_{j_0}\big| \xrightarrow{\mathcal{P}} +\infty$, as $n\to+\infty$, because $\hat\gamma_{j_0} \xrightarrow{\mathcal{P}} \gamma_{j_0}$. Since $\hat\sigma_{j_0} \xrightarrow{\mathcal{P}} v_{j_0}$, which is positive and finite, where $v_{j_0}^2 = \int_0^\tau \big[\ell_{j_0}^0(F(x))\big]^2 \frac{\pi_1(x)\,\pi_2(x)}{\pi(x)}\,\frac{dH^1(x)}{\pi(x)}$, while H^1 is the subdistribution function of the form $H^1(x) = \eta\int_0^x(1-G_1(y))dF_1(y) + (1-\eta)\int_0^x(1-G_2(y))dF_2(y),\ y\in[0,\infty)$, we have $\big|C_{j_0}\big| \xrightarrow{\mathcal{P}} +\infty$. As a result,

$$P(\,J(n,c)=0\,) = P(\max_{1\le j\le d}\big|C_j\big| > \sqrt{c\log n}\,) \ge P(\,\big|C_{j_0}\big| > \sqrt{c\log n}\,) =$$

$$= P(\,|\hat\gamma_{j_0}|/\hat\sigma_{j_0} > \sqrt{c\log n}/\sqrt{n_1 n_2/n}\,) \to 1,\ \text{as } n\to\infty.$$

Thereby,

$$P(\,T\ge j_0\,) = P(\,T\ge j_0, J(n,c)=0\,) + o_P(1) = P(\,A\ge j_0\,) + o_P(1) =$$

$$= P(\,\cup_{k\ge j_0}\cap_{l<j_0}\{W_k - 2k \ge W_l - 2l\}\,) \ge P(\,\cap_{l<j_0}\{W_{j_0} - 2j_0 \ge W_l - 2l\}\,) \ge$$

$$\ge P(\,W_{j_0} - 2j_0 \ge W_{j_0} - 2\,) = P(C_{j_0}^2 \ge 2(j_0-1)) \to 1,\ \text{as } n\to\infty,$$

and the proof is complete.

Proof of (ii) and (iii) of Lemma 2.

Since the assumptions of Theorem 1 of Janssen and Mayer (2001) are satisfied, asymptotic standard normality of the permutation version of the components C_j, $j=1,2,3,\ldots$, together with the reasoning like in the proof of Theorem 1 above, provide the asymptotic chi-square distribution with one degree of freedom of the permutation version of the test statistic $W_T(\boldsymbol{D};(\boldsymbol{X}_{()},\boldsymbol{\Delta}_{()}))$. The above and Lemma 1 in Janssen and Pauls (2003) yield the property (ii). In consequence, it also entails (iii).

APPENDIX D: ALGORITHM FOR CALCULATION AN EMPIRICAL POWER FUNCTION

1. Generate (X_{li}^o, U_{li}), $i = 1, \ldots, n_l$, from (F_l, G_l), $l = 1, 2$.

2. Define $(X_{li}, \Delta_{li}) = (\min\{X_{li}^o, U_{li}\}, \mathbb{1}(X_{li}^o \le U_{li})),\ i = 1, \ldots, n_l,\ l = 1, 2$.

3. On their basis, calculate a value of the statistic W_T.

4. Denote the pooled sample from step 2 as $(\boldsymbol{X}, \boldsymbol{\Delta}) = ((X_1, \Delta_1), \ldots, (X_n, \Delta_n))$.

5. Draw without replacement n_1 pairs $(\boldsymbol{X^{(1)}}, \boldsymbol{\Delta^{(1)}}) = ((X^{(1)}, \Delta^{(1)}), \ldots, (X^{(n_1)}, \Delta^{(n_1)}))$, which constitute the first group, from $(\boldsymbol{X}, \boldsymbol{\Delta})$, and treat the remaining observations $(\boldsymbol{X^{(2)}}, \boldsymbol{\Delta^{(2)}}) = ((X^{(n_1+1)}, \Delta^{(n_1+1)}), \ldots, (X^{(n)}, \Delta^{(n)}))$, as the second group.

6. On the basis of $(\boldsymbol{X^{(1)}}, \boldsymbol{\Delta^{(1)}})$ and $(\boldsymbol{X^{(2)}}, \boldsymbol{\Delta^{(2)}})$, calculate a value of the test statistic W_T, say, $W_T^{(1)}$.

7. Repeat steps 5-6, `npr` [number of permutation runs, in our simulation `npr` $= 1000$] times, obtaining $W_T^{(1)}, \ldots, W_T^{(\texttt{npr})}$.

8. Find a permutation critical value, i.e., the permutation $(1-\alpha)$-quantile of the W_T statistic

$$q_{W_T}^{(1)}(1-\alpha) = W_T^{(a:\texttt{npr})} + (1-b)[W_T^{(a+1:\texttt{npr})} - W_T^{(a:\texttt{npr})}],$$

   where $a = \texttt{npr} - \lfloor \alpha(\texttt{npr}+1) \rfloor$, $b = \alpha(\texttt{npr}+1) - \lfloor \alpha(\texttt{npr}+1) \rfloor$, while $W_T^{(a:\texttt{npr})}$ is the ath order statistic from the sample $W_T^{(1)}, \ldots, W_T^{(\texttt{npr})}$.

9. If $W_T > q_{W_T}^{(1)}(1-\alpha)$, reject $\mathcal{H}$ and remember 1, otherwise, accept $\mathcal{H}$ and remember 0.

10. Repeat steps 1-9 `nr` [number of runs, in our simulation `nr` $= 1000$] times.

11. Estimate the power function in the point $(F_1, G_1; F_2, G_2)$ as

$$\widehat{Powerfunction}(F_1, G_1; F_2, G_2) = \frac{\text{number of 1s}}{1000} = \frac{\text{number of rejections of } \mathcal{H}}{\text{number of runs}}.$$

Remark S1.

The above algorithm works well if at least one observation in each sample is uncensored. This is a natural silent assumption imposed in the survival analysis setting. In the case when $\boldsymbol{\Delta^{(1)}} = \mathbf{0}_{n_1}$ or $\boldsymbol{\Delta^{(2)}} = \mathbf{0}_{n_2}$, we skip such a sample and repeat step 5, once more. Also, conducting the simulation study one can observe that all the observations in one sample are censored, cf. point 4, above. Then, we also need to skip such a sample and repeat steps 1–4, again. In our numerical experiment demonstrated in the paper, such a situation only occurs under small sample sizes, i.e., $n_1 = n_2 = 13$. Such a slight modification of the algorithm also concerns the remaining solutions.

APPENDIX E: DESCRIPTION OF THE COMPETITIVE SOLUTIONS

First, recall that $\hat{\tau} = \inf\{x : Y_1(x)Y_2(x) = 0\}$.

The Gehan (1965) test is based on the statistic

$$G = \frac{\sqrt{\frac{n}{n_1 n_2}} \int_0^{\hat{\tau}} \frac{Y(x)}{n} \frac{Y_1(x)Y_2(x)}{Y(x)} \left(\frac{dN_1(x)}{Y_1(x)} - \frac{dN_2(x)}{Y_2(x)} \right)}{\sqrt{\frac{n}{n_1 n_2} \int_0^{\hat{\tau}} \left[\frac{Y(x)}{n} \right]^2 \frac{Y_1(x)Y_2(x)}{Y(x)} \frac{dN(x)}{Y(x)}}},$$

while the Mantel (1966) test is based on the statistic

$$M = \frac{\sqrt{\frac{n}{n_1 n_2}} \int_0^{\hat{\tau}} \frac{Y_1(x)Y_2(x)}{Y(x)} \left(\frac{dN_1(x)}{Y_1(x)} - \frac{dN_2(x)}{Y_2(x)} \right)}{\sqrt{\frac{n}{n_1 n_2} \int_0^{\hat{\tau}} \frac{Y_1(x)Y_2(x)}{Y(x)} \frac{dN(x)}{Y(x)}}}.$$

The statistic M is the first component of the data-driven statistic W_T. In both cases, we reject the null hypothesis for large values of the square of the respective statistics.

The Renyi-type Kolmogorov-Smirnov (Gill, 1980) statistic of the form

$$R = \sup_{t \in [0, \hat{\tau})} \frac{\left| \int_0^t \frac{Y_1(x)Y_2(x)}{Y(x)} \left(\frac{dN_1(x)}{Y_1(x)} - \frac{dN_2(x)}{Y_2(x)} \right) \right|}{\sqrt{\int_0^{\hat{\tau}} \frac{Y_1(x)Y_2(x)}{Y(x)} \frac{dN(x)}{Y(x)}}}.$$

Large values of the R statistic certifies the alternative.

The Fleming and Harrington (1991), p. 257, statistic from the $G^{p,q}$ family with $p = 0, q = 1$

$$FH = \frac{\sqrt{\frac{n}{n_1 n_2}} \int_0^{\hat{\tau}} \hat{F}(x) \frac{Y_1(x)Y_2(x)}{Y(x)} \left(\frac{dN_1(x)}{Y_1(x)} - \frac{dN_2(x)}{Y_2(x)} \right)}{\sqrt{\frac{n}{n_1 n_2} \int_0^{\hat{\tau}} [\hat{F}(x)]^2 \frac{Y_1(x)Y_2(x)}{Y(x)} \frac{dN(x)}{Y(x)}}}.$$

We reject the null hypothesis when FH^2 is large enough.

The Lin and Wang (2004) proposal of the form

$$LW = \frac{\widetilde{LW} - E[\widetilde{LW}]}{\sqrt{Var[\widetilde{LW}]}}, \quad \text{where} \quad \widetilde{LW} = \sum_{i=1}^{n} [O_{1i} - E_{1i}]^2 \Delta_{i:n},$$

cf. formula (S1) in Appendix B. An explicit form of the expectation and variance of the $\widetilde{LW}$ statistic can be found in the paper ibidem p. 490. Also, cf. Koziol and Jia (2014), p. 2, where the corrected form of the variance has been provided. When LW^2 exceeds $q_{\chi_1^2}(1-\alpha)$, we infer that the alternative is true.

The Qiu and Sheng (2008) test is a two-stage procedure employing the log-rank test and a test focused on detection crossing hazard rates. To save space, we do not present here the definition of the procedure, but we send a reader to pp. 194-197 of the aforementioned paper. For verification of the null hypothesis $\mathcal{H}$, we use the function `twostage` from the package `TSHRC` in the following configuration:

```
twostage(X, Δ, Z, nboot = 1000, alpha = 0.05, eps = 0.1)[[3]],
```

where $\boldsymbol{X} = (X_1, \ldots, X_{n_1}; X_{n_1+1}, \ldots, X_n)$, $\boldsymbol{\Delta} = (\Delta_1, \ldots, \Delta_{n_1}; \Delta_{n_1+1}, \ldots, \Delta_n)$, while $\boldsymbol{Z} = (1, \ldots, 1; 0, \ldots, 0)$ is the group indicator (i.e., 1 for the first group and 0 for the second one), `nboot` is the number of the bootstrap replications, `alpha` is the fixed significance level α, and `eps` means ε, which is one of the parameters describing the second stage procedure. For more comments, see the description of the package and the related paper. The function returns p-value of the two-stage test, which we use in the sequel. We will denote this procedure as QS.

The next competitor will be the so-called improved log-rank-type test using adaptive weights proposed by Yang and Prentice (2010). Just like above, we use, for the inference, the ready function `YPmodel.adlgrk` from the package `YPmodel`. For this purpose, we built a data frame

`data.set = data.frame(V1 =` $\boldsymbol{X}$`, V2 =` $\boldsymbol{\Delta}$`, V3 =` $\boldsymbol{Z}$`).`

Then, the function

```
YPmodel.adlgrk(data=data.set)$pval
```

returns the p-value of the adaptively weighted log-rank test, calculated using the formula (19), p. 33, Yang and Prentice (2010), based on the test statistic given by the left-hand side of the inequality (15), p. 32, with Φ_1 and Φ_2 estimated as described is section 2.3 of that paper. In the sequel, we will denote this procedure as YP.

The Liu and Yin (2017) statistic has the form

$$LY = \max_{1 \le j \le n} \left\{ \widetilde{LY}_j \, \Delta_{j:n} \right\},$$

where

$$\widetilde{LY}_j = \left[\sum_{i=1}^{j} (O_{2i} - E_{2i}) \Delta_{i:n} \right]^2 \left(\sum_{i=1}^{j} \frac{Y_1(X_{i:n}) Y_2(X_{i:n}) O_i [Y(X_{i:n}) - O_i]}{Y^2(X_{i:n})[Y(X_{i:n}) - 1]} \right)^{-1} + \left[\sum_{i=j+1}^{n} (O_{2i} - E_{2i}) \Delta_{i:n} \right]^2 \left(\sum_{i=j+1}^{n} \frac{Y_1(X_{i:n}) Y_2(X_{i:n}) O_i [Y(X_{i:n}) - O_i]}{Y^2(X_{i:n})[Y(X_{i:n}) - 1]} \right)^{-1},$$

cf. the definition on pp. 403-404 of the above paper. We reject the null hypothesis for large values of the LY statistic.

To define the nonparametric combination test of Arboretti et al. (2018), NPC for short, we introduce a technique of calculation of the p-values of the related test. For this purpose, we consider the Mantel (1966) statistic stressing its dependence from $\hat{\tau}$ and $\boldsymbol{\Delta}$, i.e, $M = M(\hat{\tau}, \boldsymbol{\Delta})$. Let $M_1 = M(\infty, \boldsymbol{\Delta})$ and $M_0 = M(\infty, \mathbf{1}_n - \boldsymbol{\Delta})$, where $\mathbf{1}_n$ is the vector $(1, \ldots, 1)$ of the length n. First, we calculate the permutation p-values of M_1 and M_0 as

$$p_{M_k} = \frac{0.5 + \sum_{s=1}^{\mathtt{npr}} \mathbb{1}(|M_k^{(s)}| \ge |M_k^{(0)}|)}{\mathtt{npr} + 1}, \quad k = 0, 1,$$

where `npr` is the number of the permutations runs, $M_k^{(0)}$ is a value of the M_k statistic calculated on the basis of the data, while $M_k^{(s)}$ is a value of the M_k statistic calculated on the basis of the sth permutation of the data at hand. Next, we compute the auxiliary quantities

$$\xi_k^{(s)} = \frac{0.5 + \sum_{j=1, j\neq s}^{\texttt{npr}} \mathbb{1}(|M_k^{(j)}| \geq |M_k^{(s)}|)}{\texttt{npr} + 1}, \quad s = 1, \ldots, \texttt{npr}, \quad k = 0, 1.$$

The empirical p-value of the NPC test with the Tippet combination function is defined as

$$p_{NPC} = \frac{0.5 + \sum_{s=1}^{\texttt{npr}} \mathbb{1}(\min\{\xi_1^{(s)}, \xi_0^{(s)}\} \leq \min\{p_{M_1}, p_{M_0}\})}{\texttt{npr} + 1}.$$

It should also be mentioned that if in $M_1^{(s)} = M(\infty, \boldsymbol{\Delta}^{(s)})$, a part of the vector $\boldsymbol{\Delta}^{(s)}$ corresponding to the first or second sample is equal to $\mathbf{0}_{n_l}$ or $\mathbf{1}_{n_l}$, $l = 1, 2$, respectively, we exclude such a permutation from the calculations as suggested in Arboretti et al. (2018). See the related paragraph in section 3.3 of the paper ibidem.

APPENDIX F: DESCRIPTION OF THE R CODE

In a separate `R` file, a computer program, enabling to compute values of the statistic W_T and the related permutation p-values is presented. The auxiliary functions permitting to calculate value of the Laguerre polynomial L_j and the standardized weighted log-rank statistic C_j are also given.

The notation used in the program is similar to that used in the paper. Let

$$\mathbf{X}1 = (x_{11}, \ldots, x_{1n_1}) \quad \text{and} \quad \mathbf{X}2 = (x_{21}, \ldots, x_{2n_2})$$

together with

$$\mathbf{Delta}1 = (\Delta_{11}, \ldots, \Delta_{1n_1}) \quad \text{and} \quad \mathbf{Delta}2 = (\Delta_{21}, \ldots, \Delta_{2n_2})$$

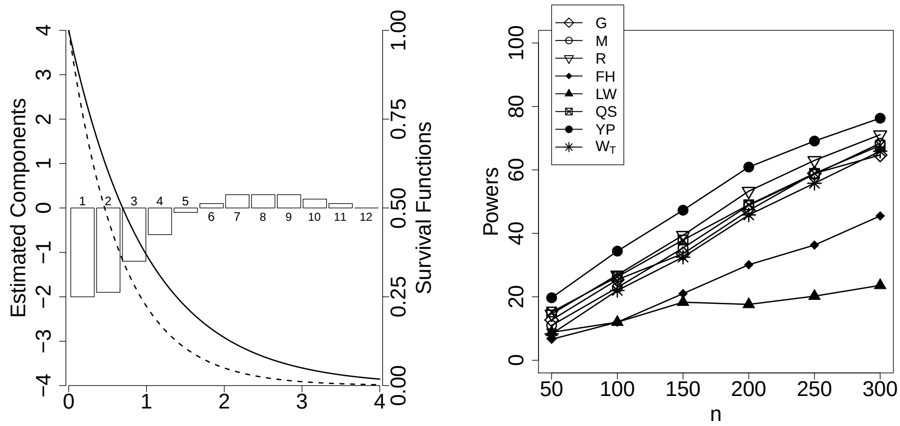
be the vectors of the observations. Let `n1` $= n_1$, `n2` $= n_2$, and `n` $= n$ be the sample size in the first, second, and pooled sample, respectively.

We start the code with the function `pol.L` returning the value of the jth Laguerre polynomial in the point x. Cf. formula (7) in the paper. Next, we introduce the function `st.wlr.test` which returns the value of the component C_j, see formula (14), where $j =$ `nr`. Finally, given c and $d(n) = d$ (in the routine, `c` and `d`, respectively), the procedure `W.T.test` provides the value of the statistic W_T, cf. (16) and (21), while the function `p.value_W.T.test` yields the p-value of the related permutation test based on `npr` permutation runs. Cf. steps 1–7 of the algorithm described in Appendix D.

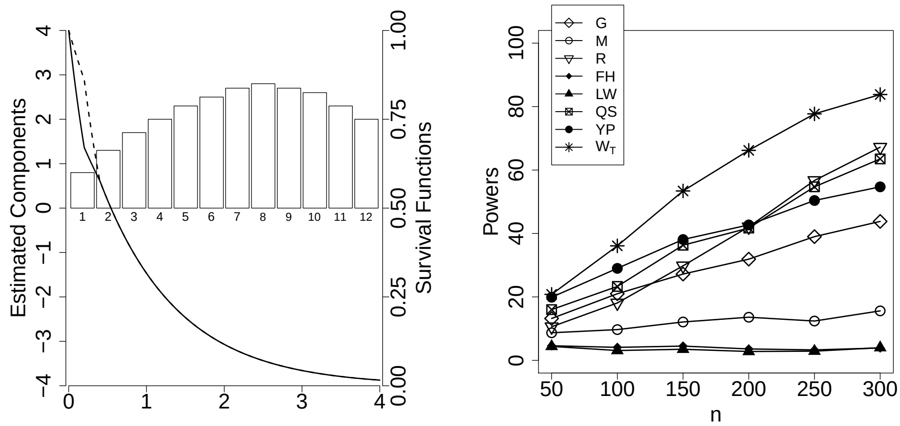
APPENDIX G: POWER COMPARISONS, $G_1 \neq G_2$

Here, we present the results of the simulation study briefly discussed in Section 4.3.2.

$\mathcal{A}_{10} : Exp(1)/Exp(1.5)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [52%, 26%]



$\mathcal{A}_{11}(0.2, 0.4) : E(2, 0.75, 1)/E(0.75, 2, 1)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [45%, 34%]



$\mathcal{A}_{12}(0.1, 0.4, 0.7) : E(2, 3, 0.75, 1)/E(2, 0.75, 3, 1)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [32%, 27%]

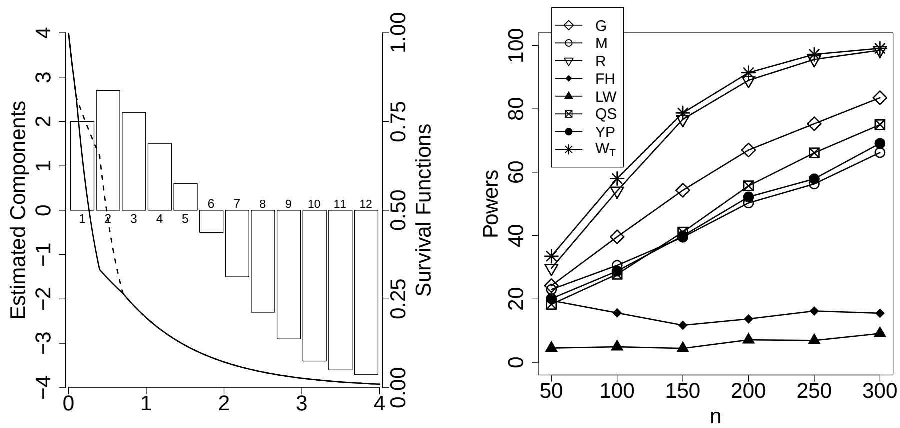
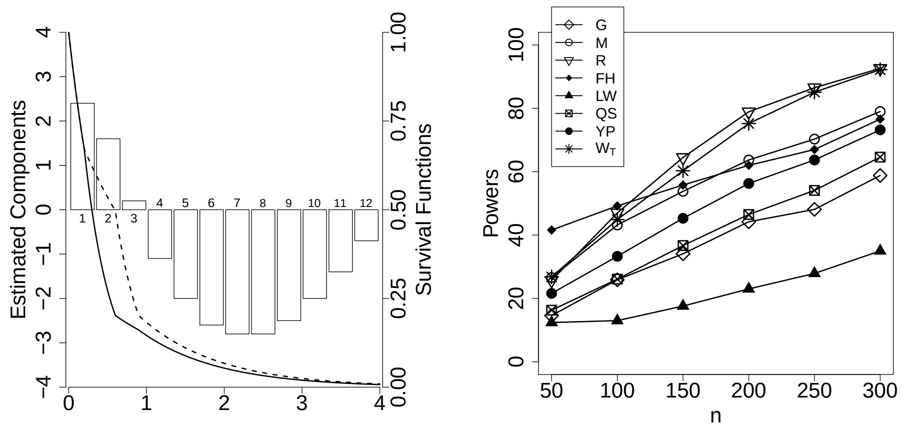
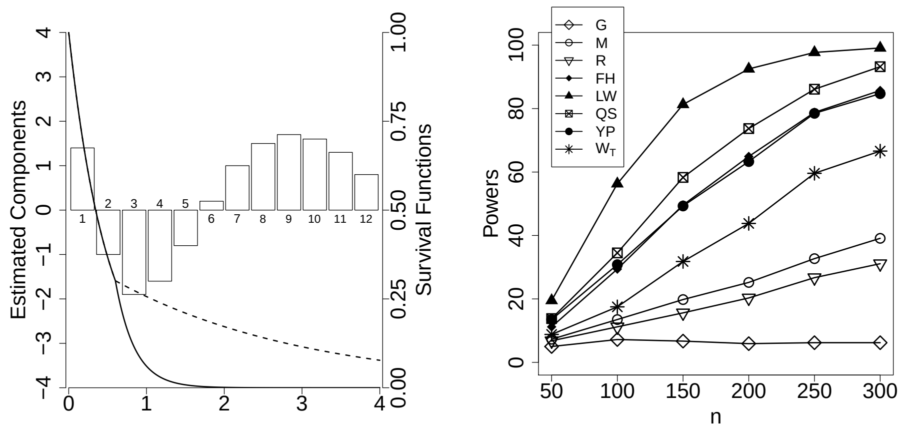


Figure S1 Left panel: Survival functions S_1 (—), S_2 (- -), in the first and second sample, respectively. The bars represent the average values of the components C_js, $j = 1, \dots, 12$, under $n = 200$. Right panel: Empirical powers against n; $\alpha = 0.05$; $n_1 = n_2$; $d = 12$; $c = 2$. Based on 1000 MC runs and 1000 permutation/bootstrap runs. Powers multiplied by 100.

$\mathcal{A}_{13}(0.2, 0.6, 0.9) : E(2, 3, 0.75, 1)/E(2, 0.75, 3, 1)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [30%, 26%]



$\mathcal{A}_{14}(0.6) : E(2, 4)/E(2, 0.4)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [28%, 30%]



$\mathcal{A}_{15}(0.1, 0.4, 0.5, 0.7) : E(2, 2, 2, 4, 4)/E(2, 3, 0.75, 0.75, 1)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [27%, 22%]

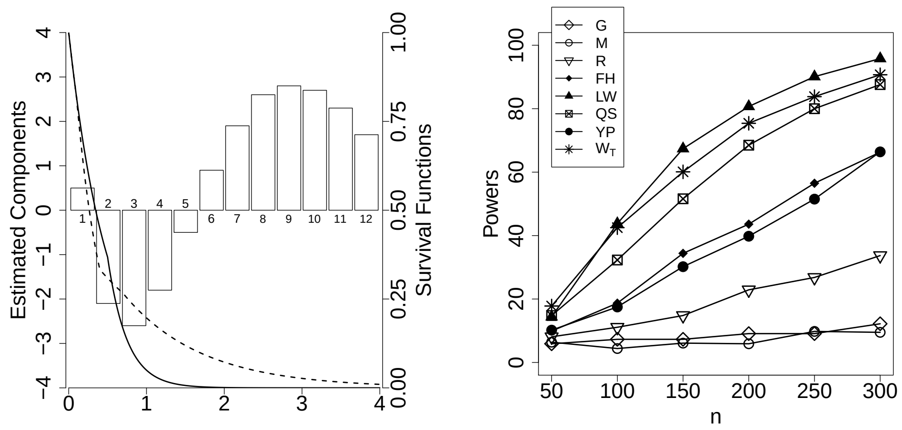
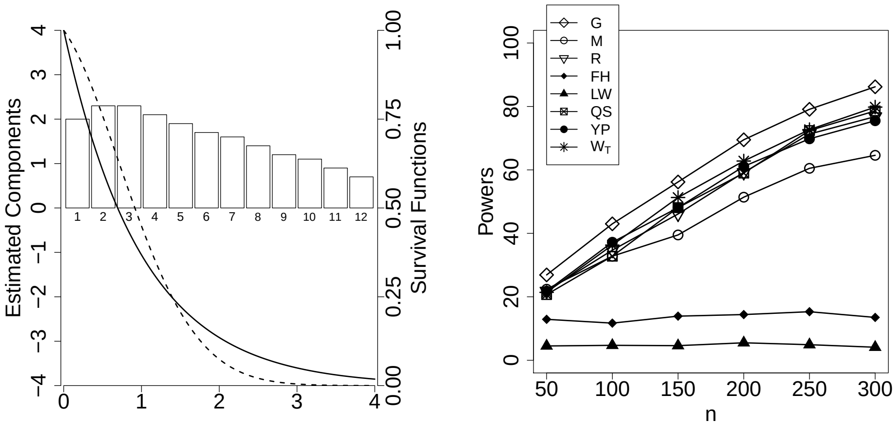
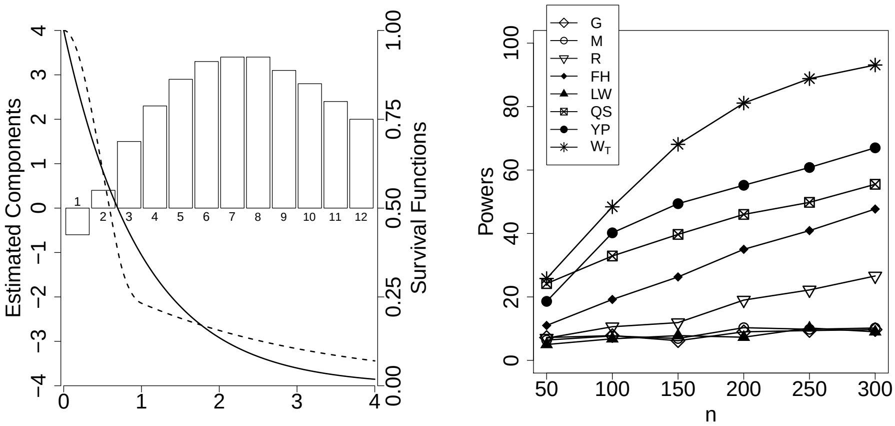


Figure S2 Left panel: Survival functions S_1 (—), S_2 (- -), in the first and second sample, respectively. The bars represent the average values of the components C_js, $j = 1, \ldots, 12$, under $n = 200$. Right panel: Empirical powers against n; $\alpha = 0.05$; $n_1 = n_2$; $d = 12$; $c = 2$. Based on 1000 MC runs and 1000 permutation/bootstrap runs. Powers multiplied by 100.

$\mathcal{A}_{16} : Exp(1)/LH(0.3, 1)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [52%, 40%]



$\mathcal{A}_{17}(0.7, 1) : Exp(1)/LH(0, 4; 8.4, -8; 0.4, 0)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [52%, 35%]



$\mathcal{A}_{18} : Cos(7, 0.6)/Cos(7, -0.6)$, $G_1 \sim U(0, 1.5)$, $G_2 \sim U(0, 2.5)$, [51%, 37%]

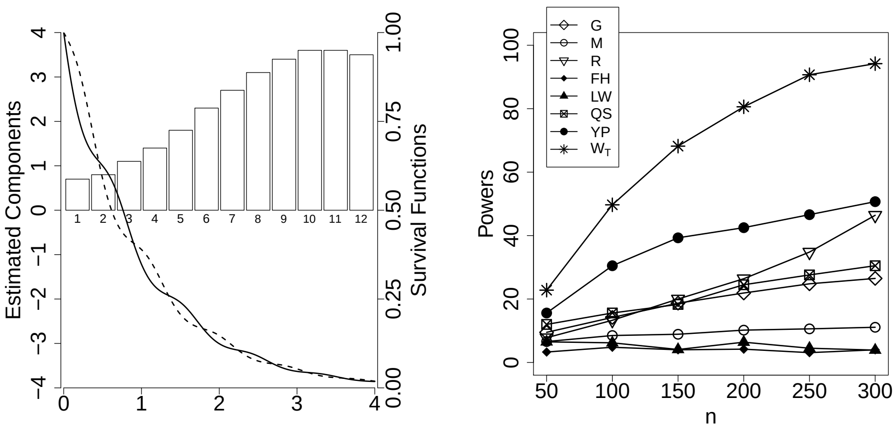


Figure S3 Left panel: Survival functions S_1 (—), S_2 (- -), in the first and second sample, respectively. The bars represent the average values of the components C_js, $j = 1, \ldots, 12$, under $n = 200$. Right panel: Empirical powers against n; $\alpha = 0.05$; $n_1 = n_2$; $d = 12$; $c = 2$. Based on 1000 MC runs and 1000 permutation/bootstrap runs. Powers multiplied by 100.

REFERENCES

Arboretti, R., Fontana, R., Pesarin, F., Salmaso, L. (2018). Nonparametric combination tests for comparing two survival curves with informative and non-informative censoring. *Statistical Methods in Medical Research*, 27, 3739–3769.

Brendel, M., Janssen, A., Mayer, C-D., Pauly, M. (2014). Weighted logrank permutation tests for randomly right censored life science data. *Scandinavian Journal of Statistics*, 41, 742–761.

Fleming, T. R., Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* Wiley, New York.

Gehan, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika*, 52, 203–223.

Gill, R. D. (1980). Censoring and stochastic integrals. *Mathematical Centre Tracts* 124. Amsterdam: Mathematisch Centrum. `http://oai.cwi.nl/oai/asset/11499/11499A.pdf`

Janssen, A., Neuhaus, G. (1997). Two-sample rank tests for censored data with non-predictable weights. *Journal of Statistical Planning and Inference*, 60, 45–59.

Janssen, A., Mayer, C-D. (2001). Conditional studentized survival tests for randomly censored models. *Scandinavian Journal of Statistics*, 28, 283–293.

Janssen, A., Pauls, T. (2003). How do bootstrap and permutation tests works? *Annals of Statistics*, 31, 768–806.

Koziol, J. A., Jia, Z. (2014). Weighted Lin-Wang tests for crossing hazards. *Computational and Mathematical Methods in Medicine*, http://dx.doi.org/10.1155/2014/643457

Lin, X., Wang, H. (2004). A new testing approach for comparing the overall homogeneity of survival curves. *Biometrical Journal*, 46, 489–496.

Liu, Y., Yin, G. (2017). Partitioned log-rank tests for the overall homogeneity of hazard rate functions. *Lifetime Data Analysis*, 23, 400–425.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, 50, 163–170.

Qiu, P., Sheng, J. (2008). A two-stage procedure for comparing hazard rate functions. *Journal of the Royal Statistical Society, Series B*, 70, 191–208.

Tarone, R. E., Ware, J. (1977). On distribution-free test for equality of survival distributions. *Biometrika*, 64, 156–160.

Yang, S., Prentice, R. (2010). Improved logrank-type tests for survival data using adaptive weights. *Biometrics*, 66, 30–38.