Web-based Supplementary Materials for *Testing Goodness-of-Fit of a Uniform Truncation Model* by Micha Mandel and Rebecca Betensky.

A Web Appendix A

Lemma. Let X be a random variable having a density f over (0,b) and let $g(\cdot) > 0$ be a continuous function on (0,b). If $\mathbb{E}g(sX) = c$ for all $0 < s < 1$ for some constant c, then $g(x) = c$ for $0 < x < b$.

Proof. Let $\alpha = \inf\{x \in (0,b) : |g(x) - c| > 0\}$. If such an α does not exist the lemma is proved, otherwise choose $\beta \in (\alpha, b)$ such that $|g(x) - c| > 0$ for all $\alpha < x < \beta$. Note that since g is continuous, either $g(x) > c$ or $g(x) < c$ for all $\alpha < x < \beta$. Suppose that $g(x) > c$ on (α, β), then we have $0 = \mathbb{E}[g\{(\beta/b)X\} - c] = \mathbb{E}|g\{(\beta/b)X\} - c|$ hence $g\{(\beta/b)X\} = c$ almost everywhere on $(0, \beta)$. Repeating this argument for $g(x) < c$ on (α, β) and using the continuity of F on (0,1) we conclude that $g(x) = c$ on $(0, \beta)$. This contradicts the definition of α and completes the proof.

Proof of Theorem 2.1. From (1) and (3), $P(T \le t|X = x) = G^*(t)/G^*(x)$, hence

$$P(Q \le s|X = x) = \frac{G^*(sx)}{G^*(x)} \underbrace{=}_{G^*=\mathrm{U}(0,b)} \frac{sx/b}{x/b} = s. \tag{A.1}$$

Using $P(Q \le s) = \mathbb{E}P(Q \le s|X) = s$ completes the first part of the proof. For the second part note that when $Q \sim \mathrm{U}(0,1)$

$$s = P(Q \le s) = \mathbb{E}P(Q \le s|X) = \int_0^b \frac{G^*(sx)}{G^*(x)} \frac{G^*(x)f^*(x)dx}{\mu} = \int_0^b G^*(sx)f^*(x)dx/\mu, \tag{A.2}$$

and by differentiation, the condition

$$\int_0^b xg^*(sx)f^*(x)dx = \mu$$

must hold for all $0 < s < 1$. Let X^{SB} be a random variable with density $xf^*(x)/\mathbb{E}X^*$, then the condition above can be written as

$$\mathbb{E}g^*(sX^{SB}) = c \quad 0 < s < 1,$$

where $c = \mu/\mathbb{E}X^*$, and the assertion follows from the Lemma above.

Proof of Theorem 2.3. Recall that X has the density $G^*(x)f^*(x)/\mu$ at x and note that T, conditional on $X = x$, has the distribution $G^*(\cdot)/G^*(x)$. For a decreasing (increasing) g^*, and for all $t < x$,

$$P(T > t|X = x) = \int_t^x \frac{g^*(u)}{G^*(x)}du < (>) \int_0^{x-t} \frac{g^*(u)}{G^*(x)}du = P(X - T > t|X = x),$$

which gives

$$P(T > t) = \mathbb{E}P(T > t|X) < (>)\mathbb{E}P(X - T > t|X) = P(X - T > t)$$

as to be proved.

Proof of Theorem 2.4. By integrating by part the right side of (A.2)

$$P(Q \le s) = s \int_0^b g^*(sx)\bar{F}^*(x)dx/\mu.$$

Subtract the distribution of U

$$P(Q \le s) - P(U \le s) = s\mu^{-1} \int_0^b [g^*(sx) - g^*(x)]\bar{F}^*(x)dx \tag{A.3}$$

and notice that for a monotone $g^*(\cdot)$, the integral in (A.3) is monotone in s and equals 0 for $s = 1$. Thus, for decreasing (increasing) g^*, $P(Q \le s) - P(U \le s) \ge (\le) 0$ for all s.

B Web Appendix B

This appendix presents by graphs several features of the left truncation model, and reports the results of the simulation study in detail. The statistics used and the sections in the paper where they are described are listed below. The framework is the random left truncation model where $T^* \sim G^*$ and $X^* \sim F^*$ are independent, and observations are restricted to the region $T^* < X^*$. The residual lifetime is denoted by $R^* = X^* - T^*$, where by truncation it is restricted to the region $R^* > 0$. The residual lifetime is subject to censoring by an independent variable $C \sim F_C$. Variables without asterisks denote elements in the truncated space.

- KS - Kolmogorov-Smirnov test that compares the ratio T/X to the uniform distribution (Section 3.1). Asymptotic P-values were calculated.
- WSR - Wilcoxon Sign-Rank test that compares the distribution of T and R (uncensored case only, Section 3.2). Asymptotic P-values were calculated.
- GW - Gehan Wilcoxon test that compares the distribution of T and R. It is of the form of the tests suggested in Section 3.2 with $w(u) = n^{-2}[m_T(u) + m_R(u)]$.
- PLR - Paired Log-Rank test that compares the distribution of T and R (Section 3.2).
- PW - Prentice Wilcoxon test that compares the distribution of T and R (Section 3.2). PW and GW are equivalent for uncensored data and in that case reduce to the Wilcoxon rank-sum test (with variance estimate that take into consideration the dependency).
- CLR - Conditional likelihood ratio test that embeds the distribution of T in the Beta family and uses the likelihood of $T|X = x$ (Section 3.3.1).
- CH2a - Chi-square test with asymptotic variance estimate that compares the nonparametric maximum likelihood of G^* to the uniform distribution (Section 3.3.2).
- CH2b - Chi-square test with bootstrap variance estimate that compares the nonparametric maximum likelihood of G^* to the uniform distribution (Section 3.3.2).

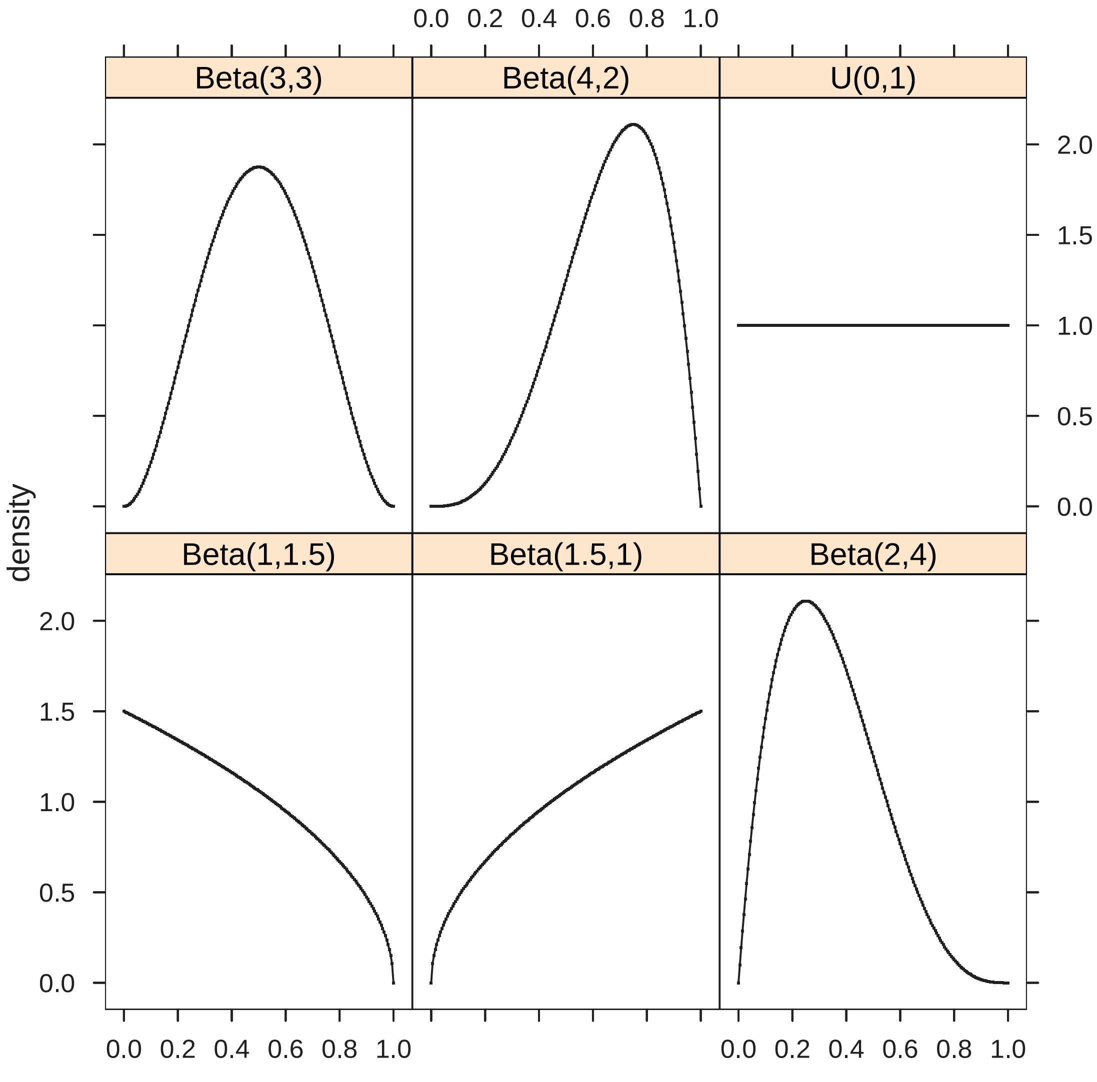

Figure 1: Densities used in the simulation study for the truncation and life times.

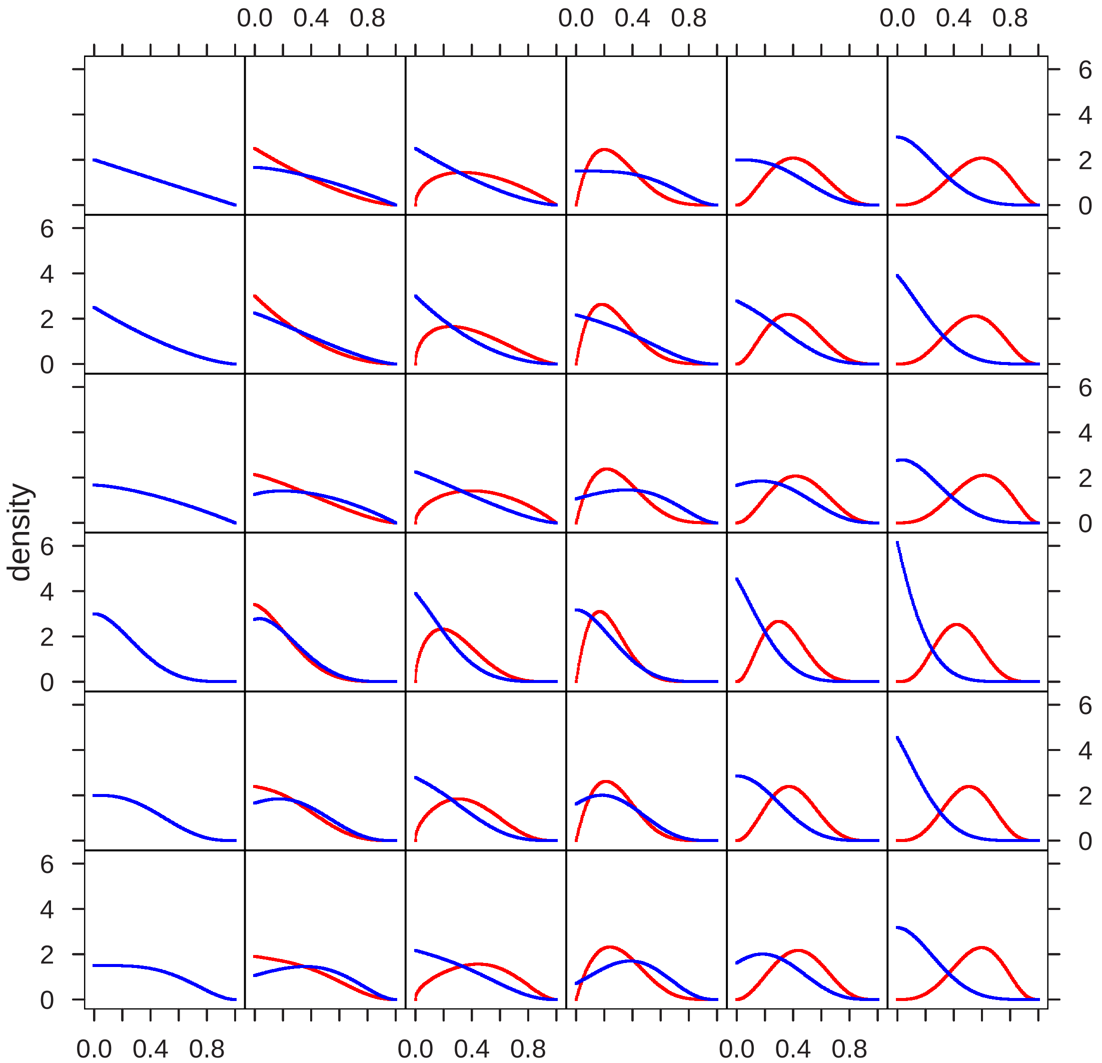

Figure 2: Densities of truncation (red solid line) and residual lifetime (blue dotted line) in the truncation model. The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

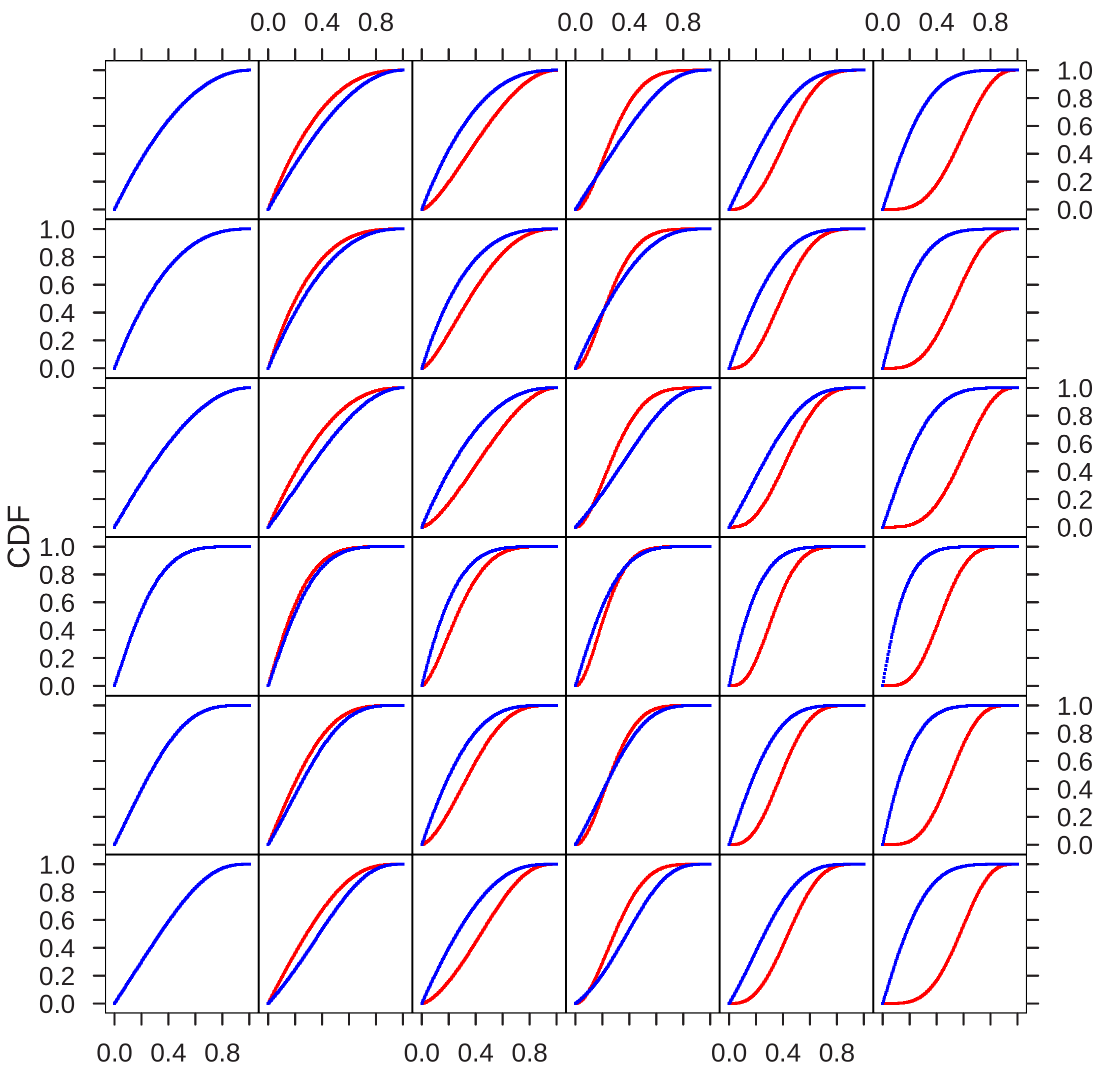

Figure 3: Distributions of truncation (red solid line) and residual lifetime (blue dotted line) in the truncation model. The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

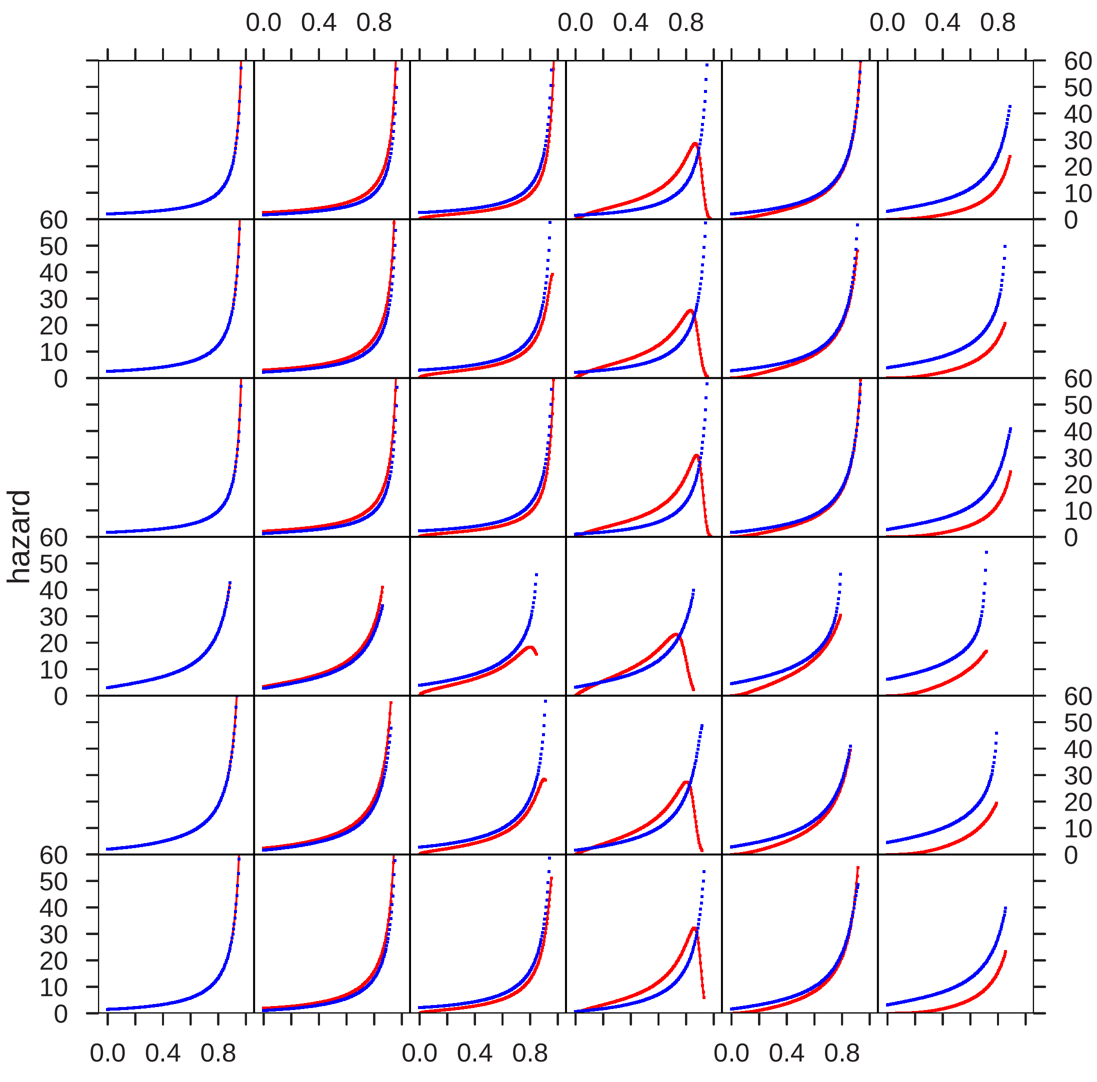

Figure 4: Hazard functions of truncation (red solid line) and residual lifetime (blue dotted line) in the truncation model. The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

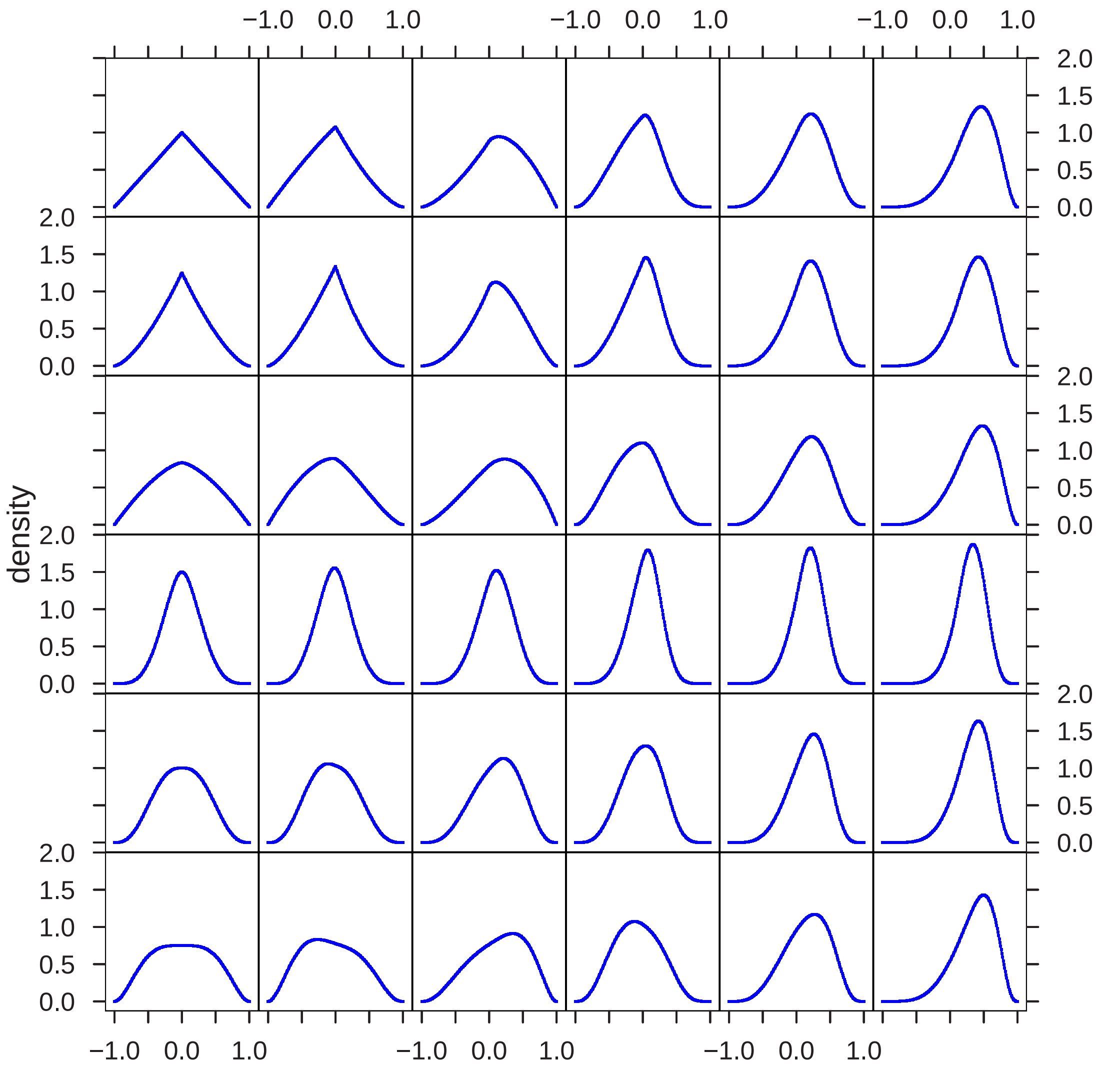

Figure 5: Density of the difference $D = T - R$ in the truncation model. The density is given by $f_D(d) = \frac{1}{2}\int_{|d|}^{\infty} g^*[(x+d)/2]f^*(x)/\mu dx$, $-1 < d < 1$ and reduces to $f_D(d) = \frac{1}{2}\bar{F}^*(|d|)/\mu$ for $G^* = \mathrm{U}(0,1)$. The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

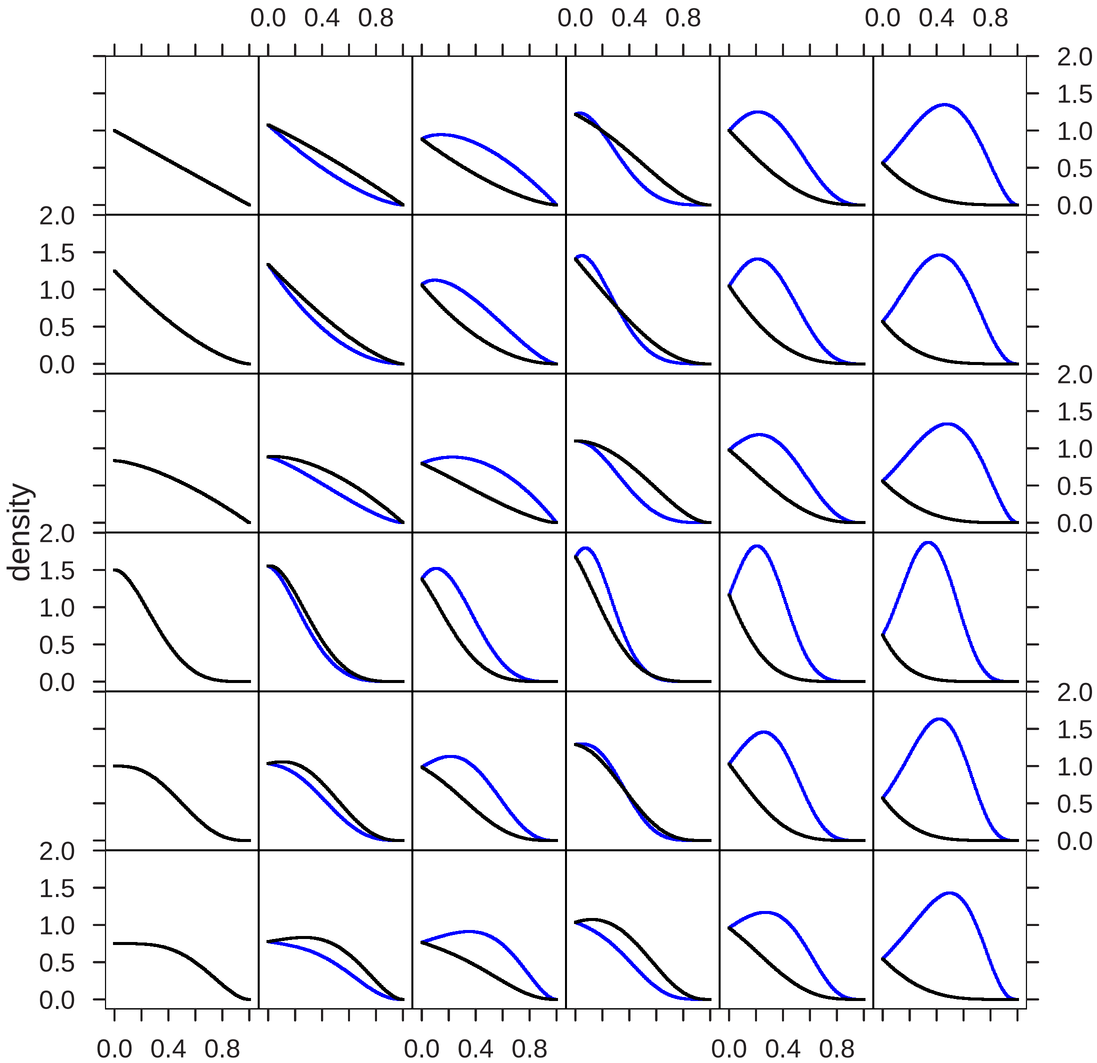

Figure 6: Density of the difference $D = T - R$ in the truncation model. Blue line the density at x and black line the density at $-x$. The density is given by $f_D(d) = \frac{1}{2}\int_{|d|}^{\infty} g^*[(x+d)/2]f^*(x)/\mu dx$, $-1 < d < 1$ and reduces to $f_D(d) = \frac{1}{2}\bar{F}^*(|d|)/\mu$ for $G^* = \mathrm{U}(0,1)$. The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

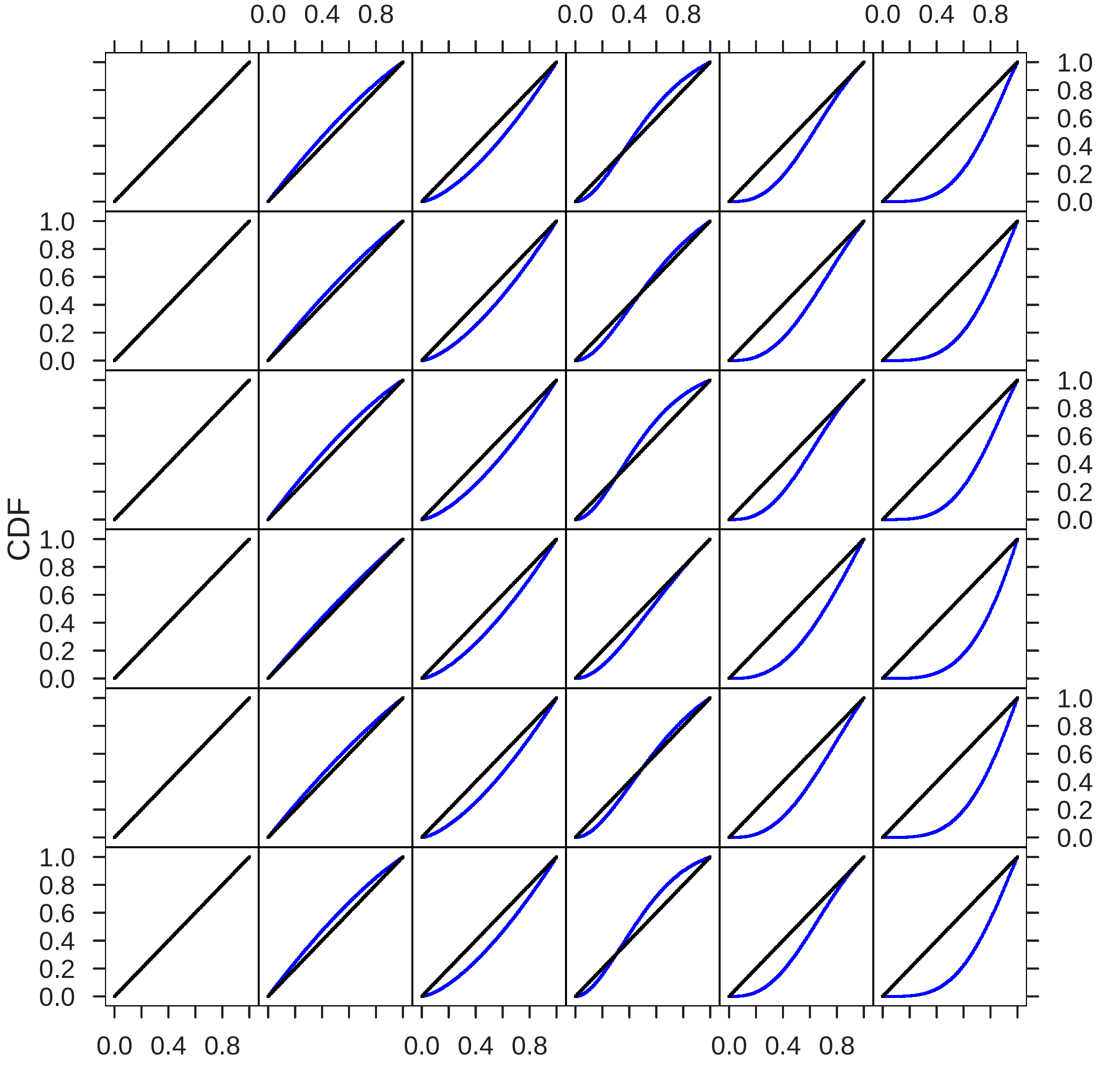

Figure 7: Distribution of the ratio $Q = T/X$ (blue line) in the truncation model. The distribution is given by $F_Q(s) = \int_0^1 G^*(sx)f^*(x)dx/\mu$, $0 < s < 1$ and reduces to $F_Q(s) = s$ for $G^* = \text{U}(0,1)$ (black line). The rows and columns represent different models for the lifetime and truncation distributions, respectively. The distribution considered are, from left to right and top to bottom, U(0,1), Beta(1,1.5), Beta(1.5,1), Beta(2,4), Beta(3,3) and Beta(4,2).

				G^*			
F^*	test	U(0,1)	Beta(1,1.5)	Beta(1.5,1)	Beta(2,4)	Beta(3,3)	Beta(4,2)
U(0,1)	KS	.063	.125	.640	.218	.950	1
	WSR	.070	.208	.708	.315	.870	1
	PLR	.088	.273	.688	.560	.708	1
	PW	.068	.223	.760	.313	.890	1
	CLR	.060	.200	.740	.965	1	1
	CH2a	.139	.265	.641	.830	.913	.978
	CH2b	.033	.172	.289	.538	.887	1
Beta(1,1.5)	KS	.050	.115	.575	.088	.990	1
	WSR	.058	.170	.663	.058	.955	1
	PLR	.073	.200	.628	.240	.860	1
	PW	.060	.175	.700	.050	.963	1
	CLR	.053	.135	.698	.900	1	1
	CH2a	.121	.224	.616	.562	.937	.986
	CH2b	.052	.142	.175	.247	.827	.983
Beta(1.5,1)	KS	.045	.185	.540	.310	.915	1
	WSR	.030	.258	.675	.438	.828	1
	PLR	.043	.308	.685	.700	.635	1
	PW	.038	.273	.713	.428	.838	1
	CLR	.035	.230	.690	.985	1	1
	CH2a	.075	.268	.632	.892	.875	.941
	CH2b	.041	.203	.335	.643	.844	1
Beta(2,4)	KS	.053	.080	.595	.228	1	1
	WSR	.040	.108	.703	.150	1	1
	PLR	.058	.143	.683	.073	.988	1
	PW	.058	.105	.750	.210	1	1
	CLR	.045	.118	.708	.793	1	1
	CH2a	.150	.178	.624	.258	.992	.995
	CH2b	.019	.073	.126	.133	.751	.825
Beta(3,3)	KS	.035	.110	.620	.108	.998	1
	WSR	.048	.165	.690	.073	.993	1
	PLR	.068	.200	.688	.135	.960	1
	PW	.058	.165	.728	.075	.995	1
	CLR	.058	.125	.690	.843	1	1
	CH2a	.115	.203	.642	.451	.977	.973
	CH2b	.044	.126	.237	.214	.877	.989
Beta(4,2)	KS	.040	.175	.580	.363	.958	1
	WSR	.030	.238	.708	.435	.918	1
	PLR	.065	.295	.693	.618	.825	1
	PW	.038	.240	.740	.435	.923	1
	CLR	.058	.235	.678	.963	1	1
	CH2a	.095	.253	.628	.845	.920	.900
	CH2b	.023	.141	.361	.545	.870	1

Table 1: Estimated power of tests for uncensored data with level 0.05. Based on 400 replications of sample size 50.

		G^*					
F^*	test	U(0,1)	Beta(1,1.5)	Beta(1.5,1)	Beta(2,4)	Beta(3,3)	Beta(4,2)
U(0,1)	KS	.033	.255	.915	.438	1	1
	WSR	.060	.388	.945	.560	1	1
	PLR	.078	.475	.920	.875	.953	1
	PW	.050	.385	.960	.520	1	1
	CLR	.045	.358	.980	1	1	1
	CH2a	.113	.361	.846	.985	1	1
	CH2b	.068	.316	.700	.942	.997	1
Beta(1,1.5)	KS	.048	.200	.913	.200	1	1
	WSR	.033	.270	.945	.103	1	1
	PLR	.043	.340	.910	.375	.993	1
	PW	.035	.293	.965	.080	1	1
	CLR	.050	.253	.970	1	1	1
	CH2a	.093	.261	.816	.836	.992	1
	CH2b	.041	.239	.571	.631	.987	1
Beta(1.5,1)	KS	.038	.335	.888	.628	1	1
	WSR	.043	.495	.943	.728	.988	1
	PLR	.040	.530	.940	.938	.938	1
	PW	.040	.500	.955	.705	.985	1
	CLR	.033	.415	.960	1	1	1
	CH2a	.073	.418	.855	1	.997	.992
	CH2b	.048	.368	.712	.982	.997	1
Beta(2,4)	KS	.045	.115	.905	.588	1	1
	WSR	.045	.133	.940	.320	1	1
	PLR	.050	.160	.915	.103	1	1
	PW	.035	.133	.965	.408	1	1
	CLR	.043	.135	.960	.995	1	1
	CH2a	.105	.153	.832	.409	1	.997
	CH2b	.039	.108	.516	.335	1	.989
Beta(3,3)	KS	.048	.198	.930	.228	1	1
	WSR	.043	.280	.958	.060	1	1
	PLR	.050	.338	.933	.170	1	1
	PW	.045	.275	.963	.058	1	1
	CLR	.048	.228	.958	.995	1	1
	CH2a	.093	.268	.855	.670	1	1
	CH2b	.043	.218	.676	.510	.997	1
Beta(4,2)	KS	.040	.333	.908	.668	1	1
	WSR	.040	.440	.938	.688	.998	1
	PLR	.038	.508	.938	.863	.993	1
	PW	.043	.455	.945	.678	.998	1
	CLR	.038	.415	.948	1	1	1
	CH2a	.055	.408	.863	.970	1	.997
	CH2b	.043	.326	.767	.933	.998	1

Table 2: Estimated power of tests for uncensored data with level 0.05. Based on 400 replications of sample size 100.

		G^*					
F^*	test	U(0,1)	Beta(1,1.5)	Beta(1.5,1)	Beta(2,4)	Beta(3,3)	Beta(4,2)
U(0,1)	KS	.055	.515	1	.775	1	1
	WSR	.060	.695	1	.803	1	1
	PLR	.065	.760	.998	.990	.998	1
	PW	.058	.675	1	.718	1	1
	CLR	.055	.663	1	1	1	1
	CH2a	.088	.519	.990	1	1	1
	CH2b	.053	.509	.967	1	1	1
Beta(1,1.5)	KS	.035	.355	1	.580	1	1
	WSR	.040	.493	1	.158	1	1
	PLR	.035	.573	.998	.620	1	1
	PW	.033	.503	1	.095	1	1
	CLR	.043	.443	1	1	1	1
	CH2a	.070	.338	.990	.987	1	1
	CH2b	.056	.343	.944	.987	1	1
Beta(1.5,1)	KS	.038	.643	.993	.958	1	1
	WSR	.050	.773	.998	.955	1	1
	PLR	.048	.828	.995	.998	.998	1
	PW	.045	.783	1	0.950	1	1
	CLR	.033	.773	1	1	1	1
	CH2a	.070	.643	.992	1	1	1
	CH2b	.058	.639	.975	1	1	1
Beta(2,4)	KS	.068	.173	1	.960	1	1
	WSR	.063	.238	1	.518	1	1
	PLR	.060	.265	1	.165	1	1
	PW	.063	.248	1	.660	1	1
	CLR	.053	.220	1	1	1	1
	CH2a	.090	.165	.988	.890	1	1
	CH2b	.039	.171	.945	.889	1	1
Beta(3,3)	KS	.048	.358	1	.635	1	1
	WSR	.035	.475	1	.055	1	1
	PLR	.050	.510	.998	.273	1	1
	PW	.035	.490	1	.055	1	1
	CLR	.038	.423	1	1	1	1
	CH2a	.078	.345	.995	.953	1	1
	CH2b	.041	.352	.966	.955	1	1
Beta(4,2)	KS	.033	.583	.993	.960	1	1
	WSR	.035	.733	.998	.915	1	1
	PLR	.048	.780	.998	.995	1	1
	PW	.040	.738	.998	.893	1	1
	CLR	.048	.693	1	1	1	1
	CH2a	.083	.580	.995	1	1	1
	CH2b	.041	.553	.980	1	1	1

Table 3: Estimated power of tests for uncensored data with level 0.05. Based on 400 replications of sample size 200.

		F^*					
G^*	test	U(0,1)	Beta(1,1.5)	Beta(1.5,1)	Beta(2,4)	Beta(3,3)	Beta(4,2)
Beta(1,1.5)	WSR	.305	.273	.388	.150	.345	.240
	KS	.213	.203	.300	.118	.285	.200
Beta(1.5,1)	WSR	.820	.775	.790	.803	.813	.800
	KS	.768	.690	.678	.738	.698	.733

Table 4: Estimated power of one sided tests for uncensored data with level 0.05. Based on 400 replications of sample size 50.

			probability and type of censoring					
			$P(\Delta = 0) = 0.75$		$P(\Delta = 0) = 0.50$		$P(\Delta = 0) = 0.25$	
model	n	test	Fixed	Random	Fixed	Random	Fixed	Random
	50	GW	.150	.030	.048	.045	.040	.040
		PW	.150	.033	.048	.040	.040	.045
		PLR	.135	.055	.030	.108	.120	.175
		CH2b	.178	.204	.230	.258	.305	.333
G^*=Beta(2,4)	100	GW	.230	.055	.045	.043	.070	.058
F^*=Beta(1,1.5)		PW	.230	.053	.045	.053	.070	.058
		PLR	.205	.075	.043	.160	.180	.270
		CH2b	.350	.485	.453	.533	.590	.630
	200	GW	.430	.070	.063	.060	.080	.063
		PW	.430	.065	.063	.055	.080	.078
		PLR	.388	.108	.058	.238	.338	.460
		CH2b	.853	.953	.920	.935	.958	.960
	50	GW	.558	.483	.713	.645	.723	.730
		PW	.558	.488	.713	.645	.723	.723
		PLR	.550	.498	.683	.595	.663	.645
		CH2b	.086	.146	.103	.138	.125	.118
G^*=Beta(1.5,1)	100	GW	.870	.805	.940	.910	.955	.960
F^*=Beta(2,4)		PW	.870	.793	.940	.905	.955	.958
		PLR	.865	.745	.908	.850	.928	.903
		CH2b	.315	.459	.420	.523	.515	.468
	200	GW	.988	.970	1	1	1	1
		PW	.988	.960	1	.998	1	1
		PLR	.990	.935	1	.983	1	.995
		CH2b	.800	.950	.905	.958	.940	.930

Table 5: Estimated power of tests for censored data with level 0.05. Based on 400 replications.

			probability and type of censoring					
			$P(\Delta = 0) = 0.75$		$P(\Delta = 0) = 0.50$		$P(\Delta = 0) = 0.25$	
model	n	test	Fixed	Random	Fixed	Random	Fixed	Random
	50	GW	.048	.060	.060	.050	.065	.063
		PW	.048	.065	.060	.053	.065	.068
		PLR	.050	.068	.068	.065	.058	.085
		CH2b	.044	.033	.043	.035	.060	.033
G^*=Beta(1,1)	100	GW	.058	.035	.033	.035	.038	.040
F^*=Beta(1,1.5)		PW	.058	.040	.033	.035	.038	.043
		PLR	.060	.040	.035	.045	.040	.058
		CH2b	.030	.028	.033	.045	.045	.045
	200	GW	.033	.038	.043	.038	.033	.038
		PW	.033	.040	.043	.038	.033	.040
		PLR	.043	.038	.058	.033	.030	.048
		CH2b	.020	.045	.055	.043	.050	.055
	50	GW	.025	.055	.038	.053	.055	.043
		PW	.025	.050	.038	.058	.055	.035
		PLR	.025	.063	.045	.068	.055	.045
		CH2b	.036	.031	.035	.040	.033	.033
G^*=Beta(1.5,1)	100	GW	.035	.053	.043	.043	.035	.045
F^*=Beta(2,4)		PW	.035	.050	.043	.045	.035	.043
		PLR	.040	.038	.050	.060	.050	.055
		CH2b	.040	.048	.055	.040	.040	.038
	200	GW	.053	.065	.060	.058	.063	.065
		PW	.053	.060	.060	.058	.063	.068
		PLR	.050	.068	.053	.045	.070	.058
		CH2b	.055	.053	.055	.060	.055	.053

Table 6: Estimated significance levels of tests for censored data with target level 0.05. Based on 400 replications.