Three Measures of Explained Variation for Correlated Survival Data under the Proportional Hazards Mixed-effects Model

Supplemental Materials

Gordon Honerkamp-Smith and Ronghui Xu

March 2, 2016

A second estimate of Ω^2 can be based on the sample variance of the $\hat{\eta}_{ij} = \hat{\boldsymbol{\beta}}' \boldsymbol{Z}_{ij} + \hat{\boldsymbol{b}}'_{ij} \boldsymbol{W}_{ij}$'s:

$$R_1^2 = 1 - \frac{\pi^2/6}{\sum_{i,j}(\hat{\eta}_{ij} - \bar{\hat{\eta}})^2/(N-1) + \pi^2/6}, \tag{1}$$

where $N = \sum_{i=1}^m n_i$ is the total number of observations. As discussed in the main paper for the $\boldsymbol{b}_i$'s to be well estimated, the cluster sizes n_i need to be reasonably large, and this is when we expect R_1^2 to be a reasonable estimate of Ω^2. We include R_1^2 in the simulations here.

All the simulations below used 25 clusters with 25 observations each (25×25), and the setup is otherwise the same as described in the main paper.

Figure 1 shows how the performance of the measures depends on the amount of censoring in the data. The measures R^2 and R_1^2 are mostly unaffected by the level of censoring, though slightly more variation is seen in R_1^2. For R_{res}^2, the amount of censoring has a modest effect, with the value of R_{res}^2 increasing as the amount of censoring increases for a fixed value of β; this effect appears to be independent of the magnitude of β. The measure $\hat{\rho}^2$ exhibits the same relationship but also appears to be more sensitive to the amount of censoring. As discussed in O'Quigley *et al.* (2005), this dependence of $\hat{\rho}^2$ on the censoring appears to a large extent affected by the finite follow up time τ.

Also of interest is whether the measures depend on the distribution of the covariates Z. Here we fix the variance of Z so that we are comparing the discreteness, skewness and higher moments of the distribution; note that otherwise rescaling Z by a non-zero constant just leads to multiplying β by a constant. In Figure 2, simulations were carried out using four different distributions for the covariate Z: Normal, Uniform, Bernoulli, and Exponential. The parameters of these distributions were chosen so that $E[Z_{ij}] = 0.5$ and $\mathrm{Var}[Z_{ij}] = 0.25$ in each case. While all four sample based measures appear to be largely unaffected by the covariate distribution, the plots suggests that R^2_{res} and $\hat{\rho}^2$ show a slight dependence on the distribution of the covariates.

Finally, we investigate whether the measures are sensitive to the baseline survival distribution. We carried out simulations using the baseline hazard function $\lambda_0(t) = a/b \cdot (t/b)^{a-1}\exp\{-(x/b)^a\}$, where $a, b > 0$. This corresponds to a Weibull distribution with shape parameter a and scale parameter b. In Figure 3, we fixed $b = 1$ and varied the shape parameter a; note that $a = 1$ corresponds to the standard exponential distribution that was used in the other simulations studies. The results in Figure 3 suggest that all four measures are not sensitive to the baseline survival distribution.

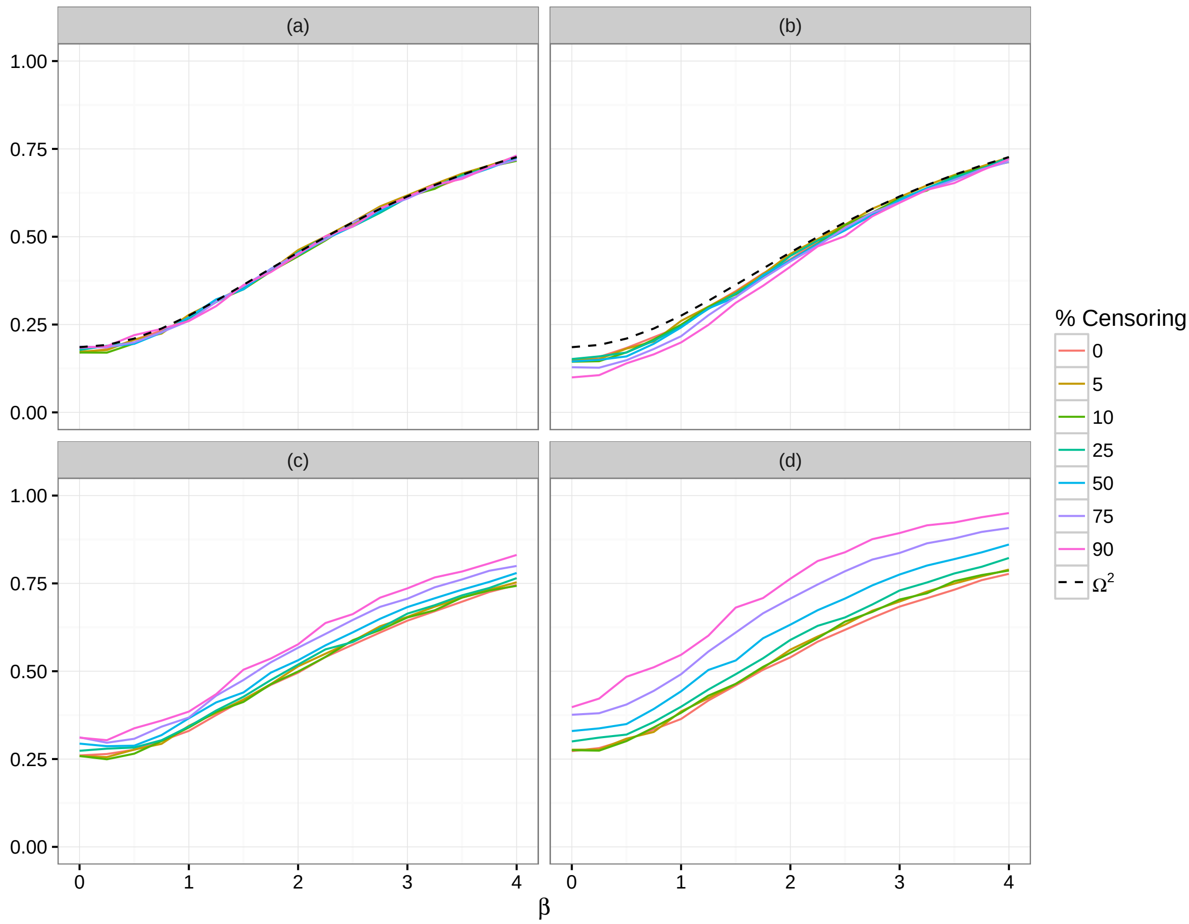

Figure 1: Performance of the measures as the percentage of censored failure times varies: (a) R^2, (b) R_1^2, (c) R_{res}^2, (d) $\hat{\rho}^2$.

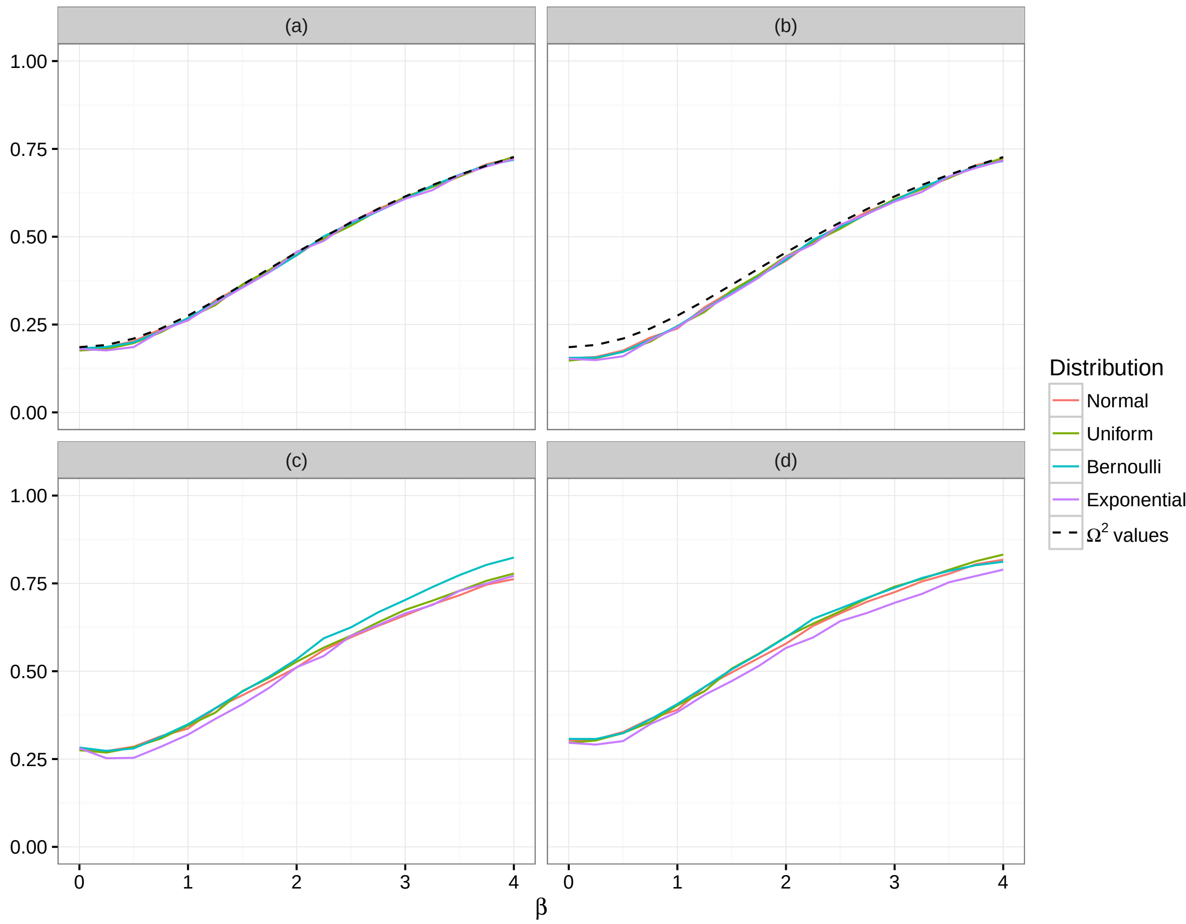

Figure 2: Simulations for the measures with different covariate distributions: (a) R^2, (b) R_1^2, (c) R_{res}^2, (d) $\hat{\rho}^2$.

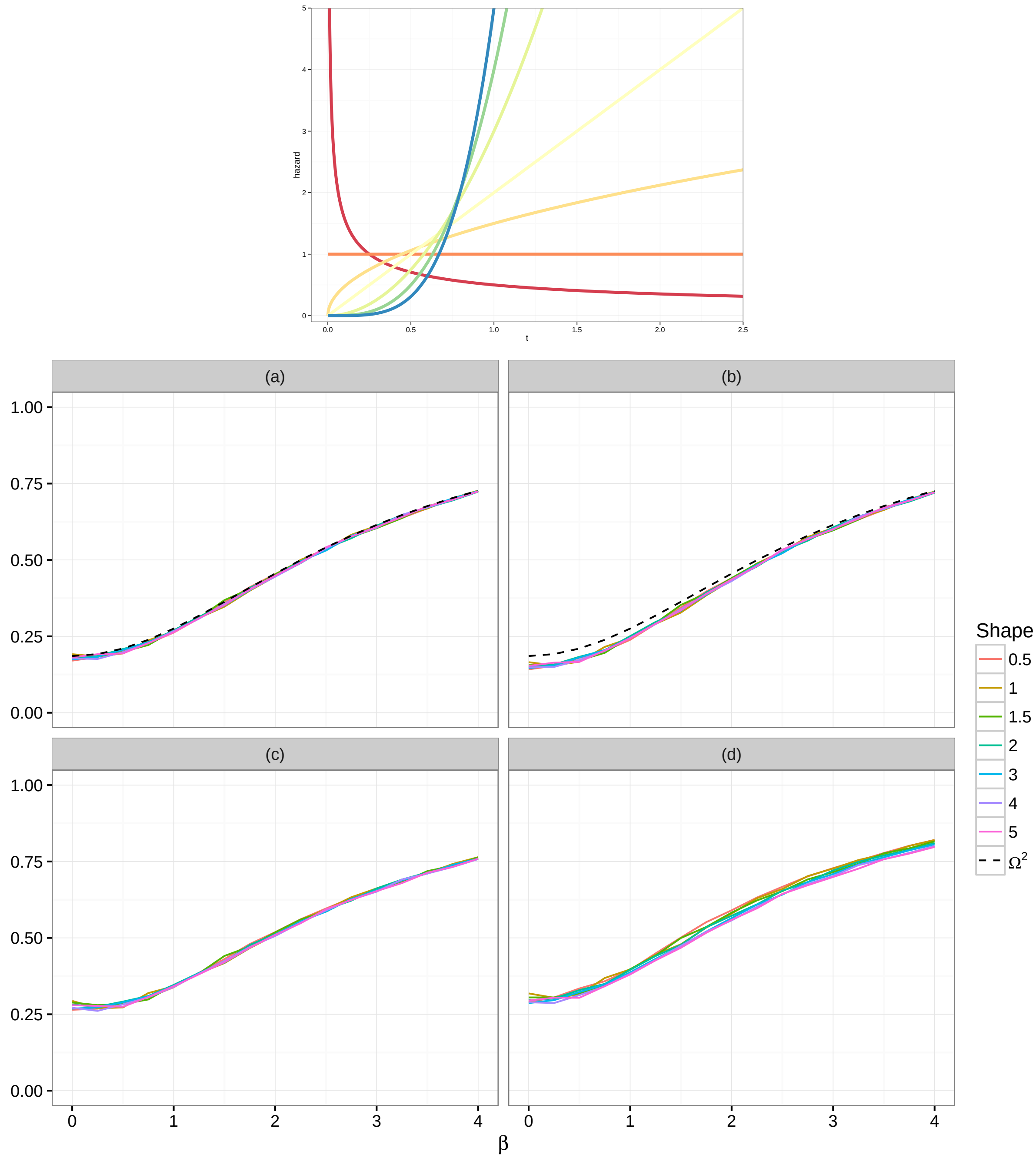

Figure 3: Top: plots of the hazard function for the Weibull distribution with varying shape parameter a and fixed scale parameter $b = 1$. Bottom: simulations for the R^2 measures with different baseline survival distributions: (a) R^2, (b) R_1^2, (c) R_{res}^2, (d) $\hat{\rho}^2$.