

## Splice Plots for Proportional Hazards Models

David Nelson\*

Siamak Noorbaloochi†

### Abstract

Under a proportional hazards assumption, the time to event outcome and the predictive measures are independent given the hazard function. Provided the model specification is valid, for large samples and consistent estimates of the hazard, the outcome and predictors should appear independent within subsamples with similar values for the estimated hazard. We use this result to develop a graphical diagnostic method. For a given sample point, we can fit a local survival model in the neighborhood of sample points with similar values for the estimated hazard. For a well specified model, the estimated local hazard function at this sample point would be expected to be close to the hazard function estimated from the fitted model. We can splice together the differences between the log estimated hazard for the fitted model and the local survival model in plots against individual predictors. A well specified model should yield a scattering of points centered on the origin. If the plot exhibits curvature then the fitted model would appear to be inadequate or inappropriate. We demonstrate how these plots can identify necessary or useful transformations of the predictors.

**Key Words:** Model diagnostics, sufficient summary, model selection

### 1. Introduction

Model diagnostics are an important element of a regression analysis as the veracity of inferences drawn from the analysis depends upon the appropriateness of the model assumptions. Nelson and Noorbaloochi [6] developed splice plots for diagnosing the fit of generalized linear regression models and compared the performance of splice plots to several graphical diagnostic methods including residual plots, partial residual plots, CERES plots, and smoothed residual plots. Based on the conditional independence of the outcome variable and the explanatory measures given the regression function, splice plots compare the fitted model with local models fit within neighborhoods of sample points with similar values for the estimated regression function. Plotting the differences between the estimated regression function and the locally estimated regression function can be used to assess model fit and, in particular, can be useful for assessing the functional form of the parametric regression model and identifying needed transformations of the explanatory measures.

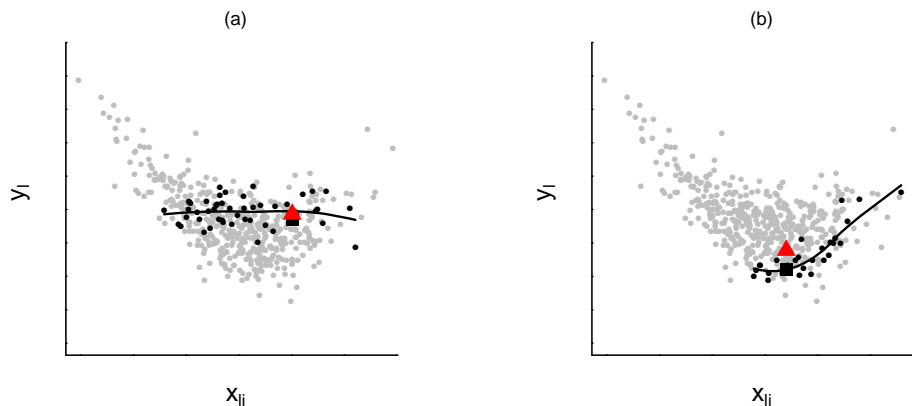
A number of different diagnostic residuals have been developed for use with proportional hazards models. Schoenfeld residuals [8] are used to assess the veracity of the proportional hazards assumption, score residuals are used to investigate the influence of individual observations on the estimated model parameters, and both martingale residuals [1], [10] and deviance residuals [10] are used to investigate needed transformations of explanatory variables. Here we adapt splice plots for use with proportional hazards models and compare these plots with martingale residual plots.

### 2. Splice Plots Background

Let  $A \perp\!\!\!\perp B$  denote independence of  $A$  and  $B$  and  $A \perp\!\!\!\perp B \mid C$  denote conditional independence of  $A$  and  $B$  given  $C$ . For a response variable  $Y$  and vector of continuous explanatory variables  $\mathbf{X} = (X_1, \dots, X_p)$  consider a sample of  $n$  independent observations

\*Minneapolis VA Health Care System, University of Minnesota Medical School

†Minneapolis VA Health Care System, University of Minnesota Medical School



**Figure 1:** Conditional regression of  $Y$  against  $X_j$  in the neighborhood  $S_{i,\epsilon}$  for (a)  $\mathbf{S}$  a sufficient summary or  $\hat{\eta} \approx \eta$  or (b)  $\hat{\eta} \neq \eta$ .  $\blacksquare$  represents the pair  $(x_{ij}, y_i)$  for the sample element  $(\mathbf{x}_i, y_i)$ ,  $\blacktriangle$  represents  $(x_{ij}, \hat{m}(x_i))$ , and the  $\bullet$  represent the sample elements falling in  $S_{i,\epsilon}$ . The solid curve presents the local regression estimate  $\hat{m}(x_{lj} | S_{i,\epsilon})$ .

$(\mathbf{x}_i = (x_{i1}, \dots, x_{ip}), y_i)$  drawn from the joint distribution of  $Y$  and  $\mathbf{X}$ . A regression of the  $y_i$  on the  $\mathbf{x}_i$  estimates the conditional distribution  $f(y | \mathbf{X})$  yielding an estimate  $\hat{m}(\mathbf{x})$  of  $\mu(\mathbf{x}) = E(Y | \mathbf{X} = \mathbf{x})$ . For many types of regression models this conditional distribution is characterized by a function  $\eta(\mathbf{X})$  together with a set of constants,  $\phi$ , that do not depend on  $Y$  or  $\mathbf{X}$ . This is the case for generalized linear regression models where the conditional distribution for  $Y$  takes the form

$$f(y | \mathbf{X}) = h(\phi, \mathbf{X}) \exp \left\{ \frac{y B(\mu(\mathbf{X})) - A(\mu(\mathbf{X}))}{a(\phi)} + C(y, \phi) \right\},$$

for some functions  $a, h, A, B, C$  and constant  $\phi$ , where  $\eta(\mathbf{X}) = g(\mu(\mathbf{X}))$  is assumed to be a linear combination of the  $\mathbf{X}$ , or transformations of the  $\mathbf{X}$ , for the link function  $g$ .

Note that  $Y \perp\!\!\!\perp \mathbf{X} | f(y | \mathbf{X})$  as all of the information the covariates  $\mathbf{X}$  contain about  $Y$  is contained in this conditional distribution function. If  $f(y | \mathbf{X})$  depends on  $\mathbf{X}$  only through  $\mathbf{S}(\mathbf{X})$  then  $Y \perp\!\!\!\perp \mathbf{X} | \mathbf{S}(\mathbf{X})$ . Such an  $\mathbf{S}$  is a sufficient dimension reduction summary [5, 7] as all of the information in the explanatory variables about the distribution of the response  $Y$  is contained in the summary  $\mathbf{S}$ . For a sufficient summary  $\mathbf{S}$  and sample element  $(\mathbf{x}_i, y_i)$ , with  $\mathbf{S}(\mathbf{x}_i) = \mathbf{s}_i$ , consider the subsample of  $(\mathbf{x}_l, y_l)$  belonging to the neighborhood

$$S_{i,\epsilon} = \{(\mathbf{x}, y) : |\mathbf{S}(\mathbf{x}_i) - \mathbf{S}(\mathbf{x})| < \epsilon\}$$

for some  $\epsilon > 0$ . For  $\epsilon$  sufficiently small we would expect the  $y_l$  and  $\mathbf{x}_l$  in  $S_{i,\epsilon}$  to appear to be observations of independent random variables. If we then fit a well specified local regression of the  $y_l$  on any function  $T(\mathbf{x}_l)$  of the  $\mathbf{x}_l$  in  $S_{i,\epsilon}$ , the local regression estimate for the conditional mean function,  $\hat{m}(T(\mathbf{x}_l) | S_{i,\epsilon})$ , should be roughly constant and close to  $\mu(\mathbf{x}_i) = E(Y | \mathbf{S}(\mathbf{X}) = \mathbf{s}_i)$ . Further then, if the regression model fit to the entire sample is well-specified then we would expect  $\hat{m}(\mathbf{x}_i)$  to be near  $\hat{m}(T(\mathbf{x}_i) | S_{i,\epsilon})$ . In particular, if  $T(\mathbf{X}) = X_j$  and we regress the  $y_l$  on the  $x_{lj}$  in this neighborhood then we would expect  $\hat{m}(x_{ij} | S_{i,\epsilon})$  to be close to  $\hat{m}(\mathbf{x}_i)$  as illustrated in Figure 1(a). If, however, we splice the local differences  $\hat{m}(x_{ij} | S_{i,\epsilon}) - \hat{m}(\mathbf{x}_i)$ , or  $g(\hat{m}(x_{ij} | S_{i,\epsilon})) - \hat{\eta}(\mathbf{x}_i)$  if we are fitting a generalized linear model, together in a plot against the  $x_{ij}$  and see any systematic curvature then we would have evidence that the specified regression model is inadequate or inappropriately specified.

If a generalized linear regression model fit to sample data is well specified then the fitted regression function  $\hat{\eta}(\mathbf{X})$  should be close to  $\eta(\mathbf{X})$  and should then function in the argument above similarly to the sufficient summary  $\mathbf{S}$ . For  $S_{i,\epsilon}$  now defined as

$$S_{i,\epsilon} = \{(\mathbf{x}, y) : |\hat{\eta}(\mathbf{x}_i) - \hat{\eta}(\mathbf{x})| < \epsilon\},$$

if we estimate  $\hat{m}(x_{ij} | S_{i,\epsilon})$  from local models fit in these neighborhoods we would expect similar behavior in plots splicing together the differences  $g(\hat{m}(x_{ij} | S_{i,\epsilon})) - \hat{\eta}(\mathbf{x}_i)$ . For a well specified model this plot should comprise a scattering of points centered around the origin with apparent independence between these differences and the predictor. If the plot displays any systematic curvature then there would appear to be a problem with the model used to estimate  $\hat{\eta}$ . In particular, consider  $\eta(\mathbf{X}) = g(E(Y | \mathbf{X})) = h(\mathbf{X}) + \delta(X_j)$  for link function  $g$ , function  $h$ , and function  $\delta$  of the  $j^{\text{th}}$  covariate but we fit a model assuming  $\eta(\mathbf{X}) = h(\mathbf{X})$ . There will be neighborhoods where the  $y_l$  and  $\mathbf{x}_l$  do not look independent, the local regression function  $\hat{m}(x_{ij} | S_{i,\epsilon})$  is not flat, and the difference  $\hat{m}(x_{ij} | S_{i,\epsilon}) - \hat{m}(\mathbf{x}_i)$  is not close to zero, as illustrated in Figure 1(b). In particular, if  $\hat{h}(\mathbf{X}) \approx h(\mathbf{X})$ , then in the subsample  $S_{i,\epsilon}$ ,  $\eta(\mathbf{x}_l) \approx c_i + \delta(x_{lj})$  for  $c_i = \hat{\eta}(\mathbf{x}_i)$  and  $g(\hat{m}(x_{lj} | S_{i,\epsilon})) - \hat{\eta}(\mathbf{x}_i) \approx \delta(x_{lj})$ . Now, if we splice these local differences together in a plot against the  $j^{\text{th}}$  covariate we would expect to see a scattering of points highlighting the component  $\delta(X_j)$  missing from the regression model fit to the data.

### 3. Splice Plots for Proportional Hazards Models

Consider an event time  $T$ , a censoring time  $C$ , and vector of explanatory variables,  $\mathbf{X} = (X_1, \dots, X_p)$ . The survival function for  $T$  given  $\mathbf{X} = \mathbf{x}$  is given by

$$S(t; \mathbf{x}) = Prob(T > t | \mathbf{X} = \mathbf{x}) = \exp\left(-\int_0^t \lambda(s; \mathbf{x}) ds\right)$$

for the conditional hazard function

$$\lambda(s; \mathbf{x}) = \lim_{\delta \rightarrow 0} \delta^{-1} Prob(t \leq T \leq t + \delta | \mathbf{x}, T \geq t).$$

It is clear then that  $T \perp\!\!\!\perp \mathbf{X} | \lambda(s; \mathbf{X})$ . All of the information the covariates  $\mathbf{X}$  contain about  $T$  is contained in this hazard function  $\lambda(s; \mathbf{X})$ . Here we focus on proportional hazards survival models of the form

$$\lambda(s; \mathbf{x}) = h_0(s) \exp(g(\mathbf{x}))$$

for a baseline hazard function,  $h_0$  and a log hazard ratio function  $g$ . For such hazard functions we then have that

$$T \perp\!\!\!\perp \mathbf{X} | g(\mathbf{X}). \tag{1}$$

Further, if  $C \perp\!\!\!\perp (T, \mathbf{X})$  or  $C \perp\!\!\!\perp T | \mathbf{X}$  then, following Dawid [2],

$$T \perp\!\!\!\perp (C, \mathbf{X}) | g(\mathbf{X}). \tag{2}$$

The regression analysis estimates the conditional distribution of  $T$  given the covariates  $\mathbf{X}$  by modeling the hazard function  $\lambda(s; \mathbf{x})$  using the observed data  $(Y, D, \mathbf{X})$  where  $Y = \min(T, C)$  and  $D = I_{(C>T)}$  indicates whether the event time was observed or censored. If a fully parametric specification of a family of proportional hazard functions

$$\{\lambda(s; \mathbf{x}) = h_0(s; \phi) \exp(g(\mathbf{x}, \theta)) | \phi \in \Phi, \theta \in \Theta\}$$

is fit to the data then the estimated values  $\hat{\theta}$ , typically estimated using likelihood maximization, can be used to form an estimate  $\hat{\lambda}(s; \mathbf{x}) = h(s; \hat{\phi}) \exp(g(\mathbf{x}; \hat{\theta}))$  of the hazard function. If a semiparametric specification of a family of proportional hazards functions

$$\{\lambda(s; \mathbf{x}) = h_0(s) \exp(g(\mathbf{x}, \theta)) \mid \theta \in \Theta\},$$

with  $h_0$  unspecified, is fit to the data, then we can combine the partial likelihood derived estimates  $\hat{\theta}$  with available methods for estimating the baseline hazard function, which use the  $g(\mathbf{x}; \hat{\theta})$ , to obtain an estimate  $\hat{\lambda}(s; \mathbf{x}) = \hat{h}_0(s) \exp(g(\mathbf{x}, \hat{\theta}))$  for the hazard function.

Given Equations (1) and (2) we see that we can readily modify the argument in the section above. Assume  $\lambda(s; \mathbf{x}) = h_0(s) \exp(g(\mathbf{x}; \theta))$  and we fit a model assuming  $\lambda(s; \mathbf{x}) = h_0(s) \exp(g(\mathbf{x}, \beta))$ . Let

$$S_{i,\epsilon} = \{(y, d, \mathbf{x}) : |g(\mathbf{x}_i; \hat{\beta}) - g(\mathbf{x}; \hat{\beta})| < \epsilon\}$$

denote small neighborhoods constructed around a given sample value  $(y_i, d_i, \mathbf{x}_i)$ . The model fit to the data is well specified so, with sufficiently large  $n$  and small  $\epsilon$ , we should essentially have  $(Y, C) \perp\!\!\!\perp \mathbf{X} \mid g(\mathbf{X}; \hat{\beta})$  for  $(Y, D, \mathbf{X}) \in S_{i,\epsilon}$ . Let  $\hat{w}_{i,\epsilon}(\mathbf{x}_i)$  be a local regression estimate for the conditional log hazard function derived from regressing the  $(y_l, d_l)$  on the  $\mathbf{x}_l$  or any subset of the  $\mathbf{x}_l$  for  $(y_l, d_l, \mathbf{x}_l) \in S_{i,\epsilon}$ . For brevity of discussion we refer the reader to Loader [4] for a discussion of these local survival regression models. With the near conditional independence in this neighborhood we would expect  $\hat{w}_{i,\epsilon}(\mathbf{x}_i) - \log \hat{\lambda}(s; \mathbf{x}_i)$  to be near zero in value. Further then, if we join or *splice* together these individual local estimates,  $\hat{w}_{i,\epsilon}(\mathbf{x}_i) - \log \hat{\lambda}(s; \mathbf{x}_i)$ , in a plot against the respective sample values for individual covariates or functions of the covariates the constructed *splice plots* should exhibit no systematic curvature.

In contrast to the above, assume that the relationship between  $T$  and the covariates is given by  $\lambda(s; \mathbf{x}) = h_0(s) \exp(\tilde{g}(\mathbf{x}; \theta))$  but we fit the model  $\lambda(s; \mathbf{x}) = h_0(s) \exp(g(\mathbf{x}, \beta))$  where

$$\tilde{g}(\mathbf{x}; \theta) = \tilde{g}(\mathbf{x}; \beta, \phi) = g(\mathbf{x}; \beta) + d(\mathbf{x}; \phi)$$

for a function  $d$ . If  $g(\mathbf{x}; \hat{\beta})$  and  $\hat{h}_0(s)$  are close to  $g(\mathbf{x}; \beta)$  and  $h_0(s)$  then in the neighborhood  $S_{i,\epsilon}$  around the  $i^{th}$  sample value with covariate values  $\mathbf{x}_i$  and  $g(\mathbf{x}_i; \hat{\beta}) = c_i$  we have  $\log \lambda(s; \mathbf{x}) \approx \log(h_0(s)) + c_i + d(\mathbf{x}; \phi)$ . Now, as above, the  $\hat{w}_{i,\epsilon}(\mathbf{x}_i) - \log \hat{\lambda}(s; \mathbf{x}_i)$  will not tend to be close to zero and scatterplots of these quantities against the values of the  $x_{ij}$  may exhibit systematic curvature. Furthermore, if  $d(\mathbf{x}; \phi) = \delta(x_j; \phi)$ , namely the specified regression function is missing a component that is a function solely of the  $j^{th}$  covariate, then in the neighborhood  $S_{i,\epsilon}$  we have

$$\log \lambda(s; \mathbf{x}) \approx \log c_i + \log h_0(s) + \delta(x_{i,j}; \phi)$$

Here, the values of the  $\hat{w}_{i,\epsilon}(\mathbf{x}_i) - \log \hat{\lambda}(s; \mathbf{x}_i)$  would be expected to be close to  $\delta(x_{ij}; \phi)$  and, hence, scatterplots of these quantities against the values of the  $x_{ij}$  will present the form of the missing functional component in the fitted regression function.

#### 4. Implementation and Empirical Performance

**Simulation Studies:** To demonstrate this diagnostic method we construct splice plots for models fit to samples drawn from the population models presented in Table 1. We drew samples of 500 elements from each population, fit a Cox proportional hazards model to the sample data modeling the log hazard function as an additive linear function of the two covariates, and constructed splice plots looking at potential transformations for each of

**Table 1:** Simulation Study Population Models

$$T | \mathbf{X} = \mathbf{x} \sim Weibull(\exp(g(\mathbf{x})), 2), \quad C | \mathbf{X} = \mathbf{x} \sim 0.5 + Exponential(0.25|x_1|)$$

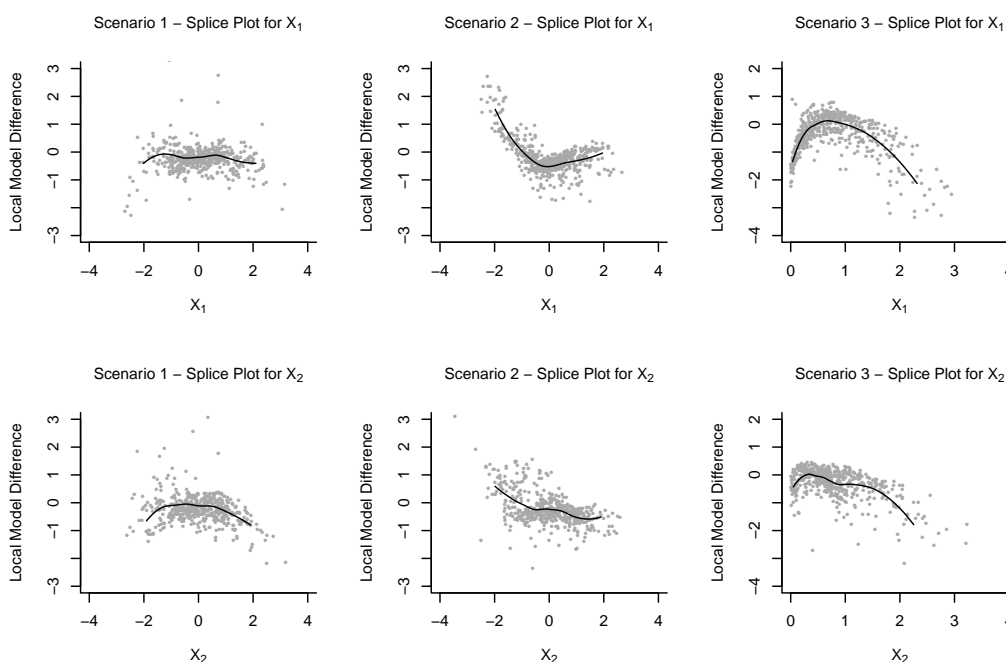
$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim MVN(\mathbf{0}, \Sigma), \quad \Sigma = \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix}$$

Simulation Models for  $g(\mathbf{X})$

Scenario 3:  $g(\mathbf{X}) = X_1 + X_2$

Scenario 4:  $g(\mathbf{X}) = 0.5X_1 + 0.5X_1^2 + X_2$

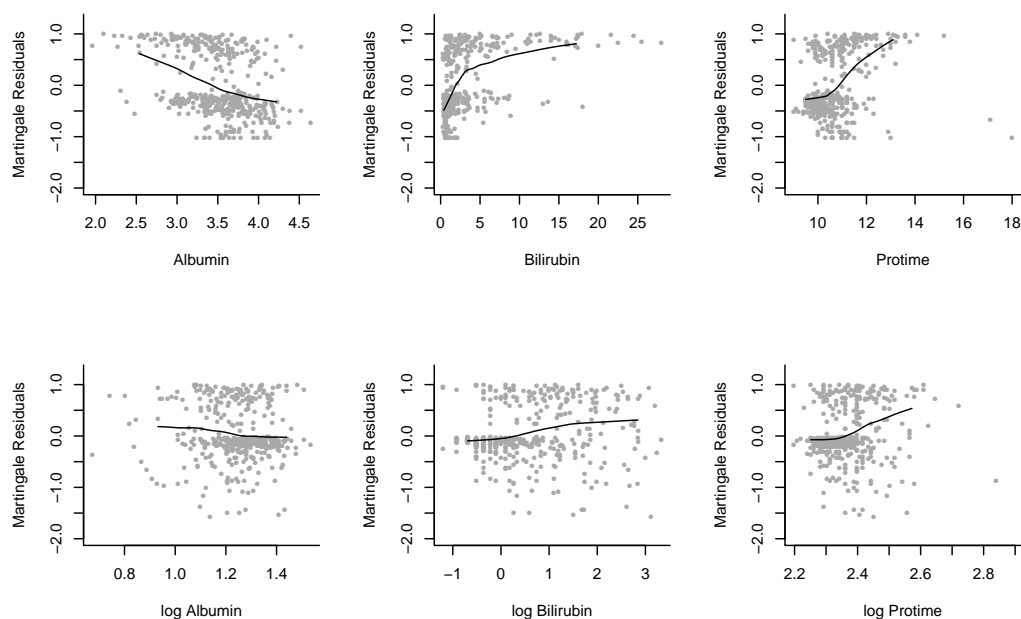
Scenario 5:  $g(\mathbf{X}) = \log |X_1| + |X_2|$



**Figure 2:** Splice for survival models fit to samples drawn from population models in Table 1.

the individual covariates. Specifically, for the  $j^{th}$  predictor, we constructed a neighborhood around a sample point  $(y_i, \delta_i, \mathbf{x}_i)$  using the one-fifth of the sample values  $(y_l, \delta_l, \mathbf{x}_l)$  with estimated values for  $\hat{g}(\mathbf{x}_l)$  closest to  $\hat{g}(\mathbf{x}_i)$ . Within this neighborhood we estimated  $\hat{w}_{i,\epsilon}(\mathbf{x}_i) = \hat{w}_{i,\epsilon}(x_{ij})$ , the conditional log hazard function at  $x_{ij}$ , using a local likelihood second order survival model [4] regressing the  $(y_l, d_l)$  on the  $x_{lj}$  in this neighborhood.

The constructed splice plots are presented in Figure 2. For the first of these scenarios, as expected, the splice plots for both of the  $X_i$  present a scattering of points with little indication of any curvature. In the second of these scenarios the fitted model is missing  $X_1^2$  as a predictor. In the splice plot for  $X_1$  we observe curvature consistent with this missing quadratic term. The splice plot for  $X_2$  for this scenario exhibits an unexpected distribution of points for the lower values of  $X_2$  that likely stems from the correlation

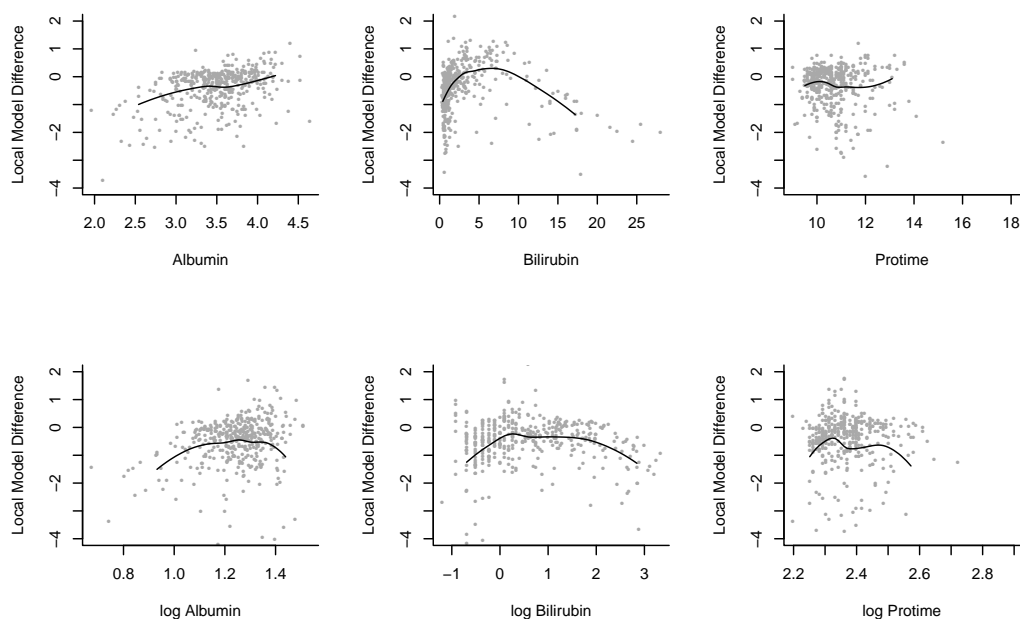


**Figure 3: PBC Martingale Residual Plots.** The first row of plots presents martingale residual plots for albumin, bilirubin, and protime from model with no predictors. The second row of plots presents martingale residual plots for log albumin, log bilirubin, and log protime from a model regressing survival on age, edema, log albumin, log bilirubin, and log protime.

between  $X_1$  and  $X_2$  affecting the estimation of the  $\hat{w}_{i,\epsilon}(x_{ij})$ . As discussed Nelson and Noorbaloochi [6], the splice plots for predictors other than those directly involved with the missing functional component can exhibit some curvature but the curvature present in the splice plots for the relevant predictors will tend to be much stronger. Lastly, for the third scenario, the splice plot for  $X_1$  exhibits curvature that is stronger than that for a simple logarithmic transformation of  $X_1$  but is consistent with the difference between a logarithmic transformation and a linear function of  $X_1$ . Here again, the splice plot for the second predictor is perhaps a little less flat and featureless than we could expect but the plot presents much less systematic curvature than is present in the plot for  $X_1$ .

**Mayo Primary Biliary Cirrhosis Dataset:** We end this discussion with an application to a well studied dataset. Therneau and Grambsch [9] discuss analysis of data from a Mayo Clinic randomized clinical trial studying primary biliary cirrhosis (PBC) of the liver. The data set comprises records for 312 individuals who participated in the trial and another 106 individuals who consented to provide study data, including study outcomes, but did not participate in the trial. Here we model the association between survival and age, presence of edema, serum albumin level, serum bilirubin level, and standardized blood clotting time (denoted here as protime). We compare the splice plots constructed for serum albumin, serum bilirubin, and protime to martingale residual plots constructed for these three predictors. Again, for brevity, we refer the reader to Therneau and Grambsch [9] for a discussion of martingale residuals.

The first row of plots in Figure 3 presents the martingale residual plots for these three variables constructed from a model with no predictors included in the fitted model. In the plot for serum bilirubin we see curvature consistent with a logarithmic function of this predictor. The residual plots for serum albumin and protime are less clear in demonstrating



**Figure 4: PBC Data Splice Plots.** The first row of plots presents splice plots for albumin, bilirubin, and protine from model regressing survival on age, edema, albumin, bilirubin, and protine. The second row of plots presents splice plots for log albumin, log bilirubin, and log protine from a model regressing survival on age, edema, log albumin, log bilirubin, and log protine.

what, if any, transformations of these predictors would be beneficial. Therneau et al. [10] fit a model with log transformations of all three of these predictors. The second row of martingale residual plots in Figure 3 are constructed from this final fitted model; while not entirely indicative of a good fit, these plots do not indicate substantial problems with the fit of the model. Each of the plots presents a scatter of points that are generally flat though not entirely flat, and perhaps even less flat in the log protime plot. In summary then, while the fit of the model may not be ideal, the model could be expected to provide useful inference regarding the association between mortality and these predictors.

The first row of splice plots presented in Figure 4 were constructed using a model including these five predictors with no transformations. We see here as well an indication in the plot for serum bilirubin that a logarithmic transformation of this predictor may be beneficial. The splice plot for protime does not indicate a need to transform this predictor while, here as well, the splice plot for serum albumin is a little difficult to interpret. For the model fitting log transformations of these three predictors we constructed the splice plots presented in the second row of Figure 4. As with the martingale residual plots for this model, these splice plots do not clearly indicate substantial problems with the specification of the model. The splice plots, while not ideal, present generally flat scatterings of points with the plot for bilirubin suggesting that an additional or alternate transformation of the bilirubin levels may be necessary.

In both the martingale residual plots and the splice plots we see that the log transformation of the bilirubin measure leads to a better fitting model but there may be some additional improvement in fit to be found. This dataset has been examined in numerous discussions of diagnostic methods for Cox proportional hazards models. In these previous discussions a logarithmic transformation of bilirubin is commonly identified as leading to an improvement in the fit of the survival model but some of the previous examinations of this dataset

have identified an inverse square root transformation as potentially useful, see León and Tsai [3]. We will not pursue an analysis of this dataset in any greater depth as the focus here is simply to demonstrate the application of splice plots for model diagnostics.

### 5. Summary and Discussion:

For assessing fit of generalized linear models, Nelson and Noorbaloochi [6] found that splice plots offer some improvement in performance over many common residual plots and similar performance to smoothed residual plots for smaller numbers of predictors. For situations with a large number of predictors, splice plots performed better than the smoothed residual approaches. Here we found similarly good performance for splice plots in assessing the fit of proportional hazards models, with comparable performance to martingale residual plots for these models.

The approach discussed for developing splice plots for proportional hazards models and generalized linear models depends solely on the conditional independence of the outcome and predictors given the regression function and the ability to express this conditional independence as equivalent to the conditional independence of the outcome and predictors given a simple estimable univariate function of the predictors. This approach should be applicable and perform well in other situations where conditional local regression estimates can be applied to estimate the conditional mean or suitable transformation of the conditional mean for the outcome given the predictors.

### Acknowledgments

This research was supported by Department of Veterans Affairs Health Services Research and Development Service grant IIR 09-360. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

### References

- [1] E.B. Barlow and R.L. Prentice. Residuals for relative risk regression. *Biometrika*, 75:65–74, 1988.
- [2] A.P. Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society, Series B*, 41:1–31, 1979.
- [3] L.F. León and C.-L. Tsai. Functional form diagnostics for cox’s proportional hazards model. *Biometrics*, 60:75–84, 2004.
- [4] C.R. Loader. *Local Likelihood and Regression*. Springer, New York, 1999.
- [5] D. Nelson and S. Noorbaloochi. Sufficient dimension reduction summaries for balanced comparisons. *Journal of Statistical Planning and Inference*, 139:617–628, 2009.
- [6] D. Nelson and S. Noorbaloochi. Splice plots for generalized linear models. *Communications in Statistics, Theory and Methods*, 45:3524–3540, 2016.
- [7] S. Noorbaloochi and D. Nelson. Conditionally specified models and dimension reduction in the exponential families. *Journal of Multivariate Analysis*, 99:1577–1589, 2008.
- [8] D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69:239–241, 1982.
- [9] T. Therneau and P. Grambsch. *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York, 2000.
- [10] T. Therneau, P. Grambsch, and T. Fleming. Martingale-based residuals for survival models. *Biometrika*, 77:147–160, 1990.