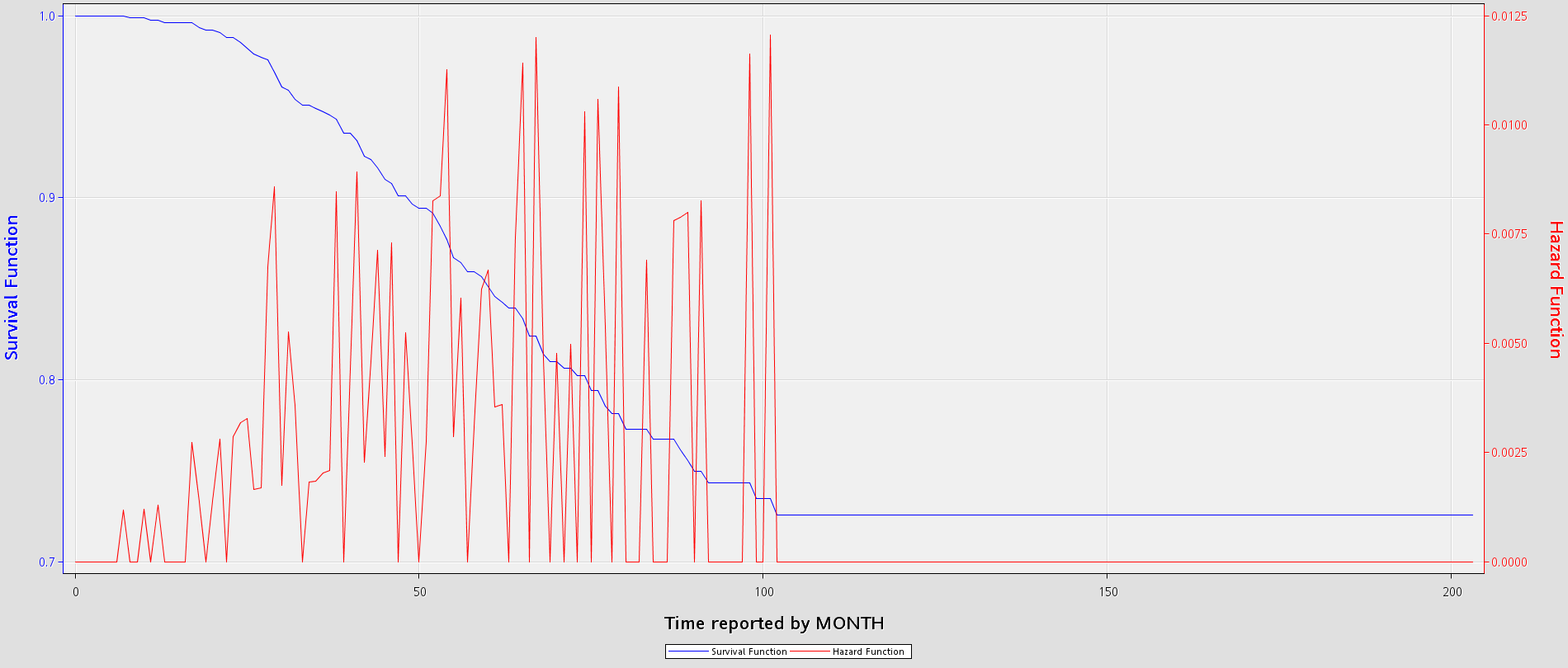
Bobby Fisher

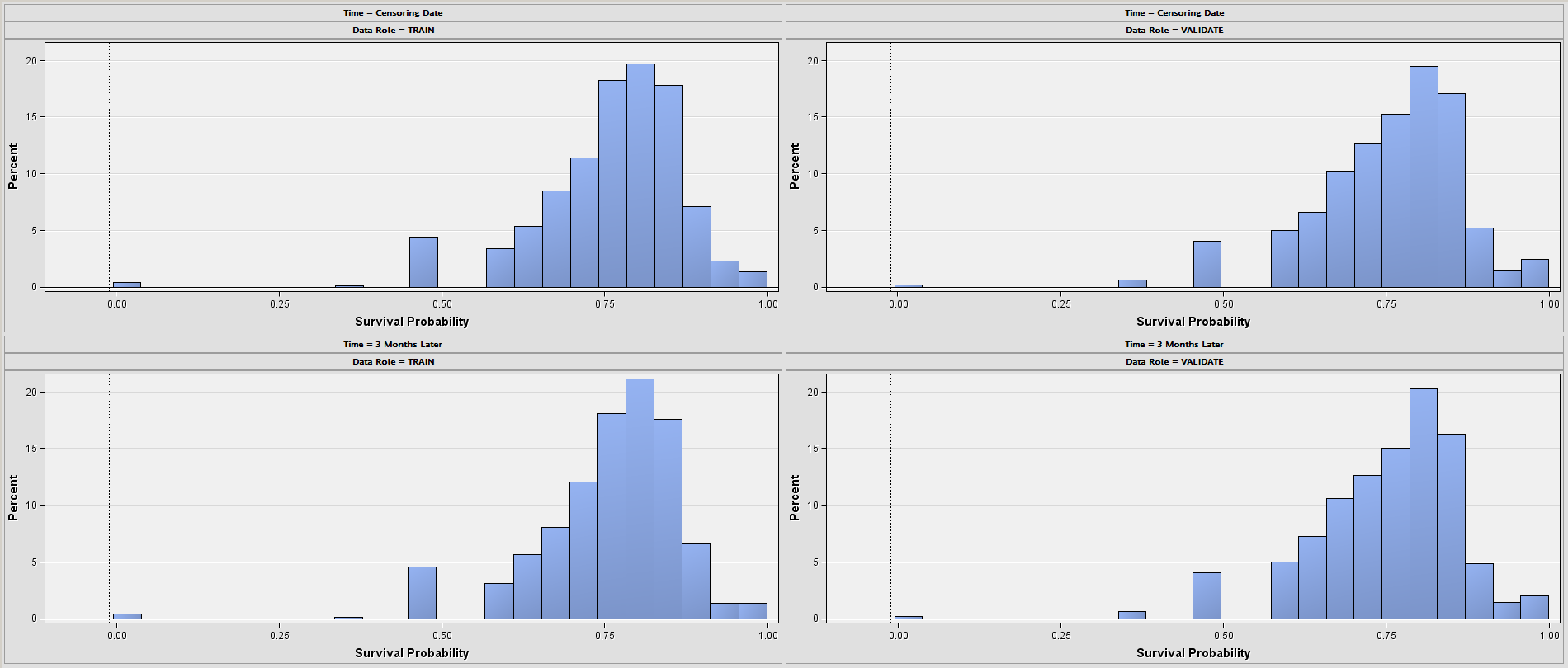
11187333

Loan Data Extra Credit

1.

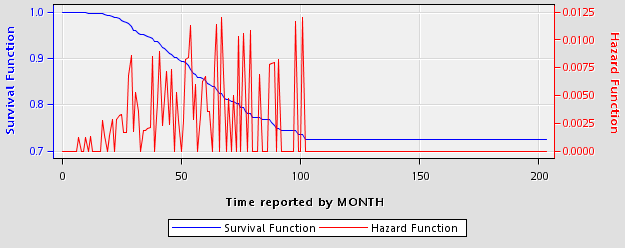
The survival curve declines steadily until month 105 when it flat lines at about 72%. Hazard is highest around month 55, 65, 67, 98, and 101. The probability of surviving month 50 is about 89% whereas surviving month 105 is about 72.5%.

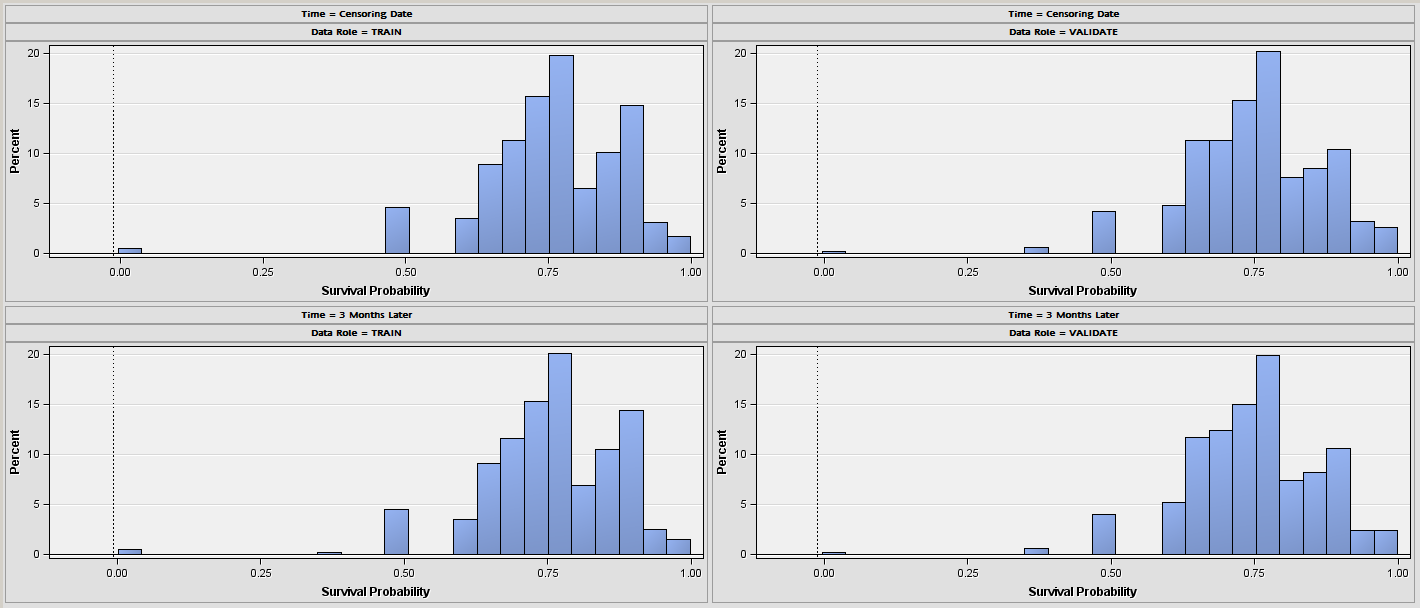




2.

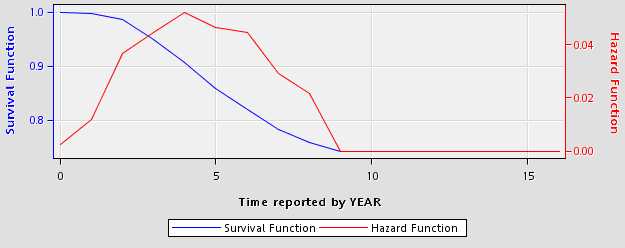
By adding rating as an input, there was very minimal improvement to the model. The first model has a KS stat of .1923 and a Gini of .1769. The second model had a KS of .1978 and a Gini of .1781. So there was improvement.

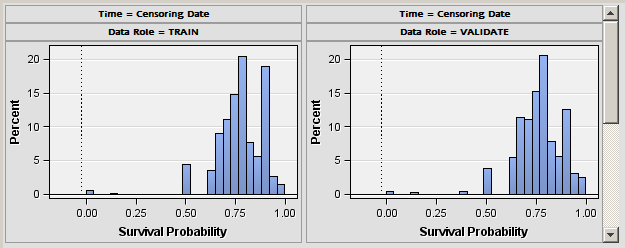




3.

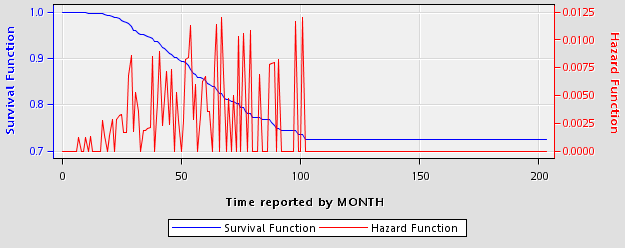
Changing the time interval didn’t help. It made things worse. It is easier to see how the hazard function changes over time. It also shows how gradual the survival function is. Year 1 has a 99.7% chance of surviving, but by year 5 that has dropped to 85.9%. The scores from the validation were a benefit of .14, a KS of .1877, and a Gini of .16. These are almost a whole point lower than the first model. The first model is better using the month time interval.

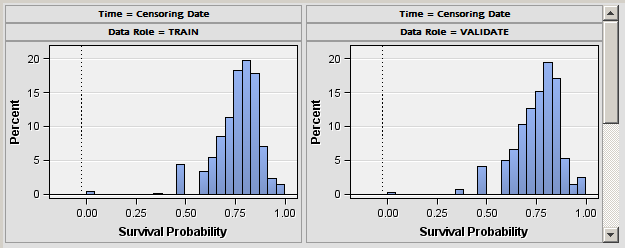




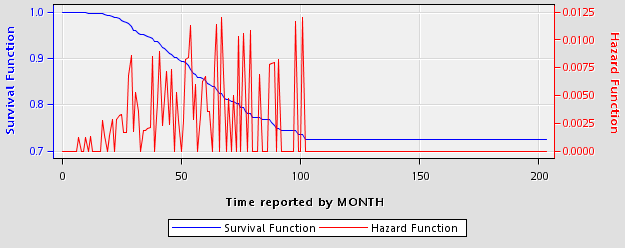
4.

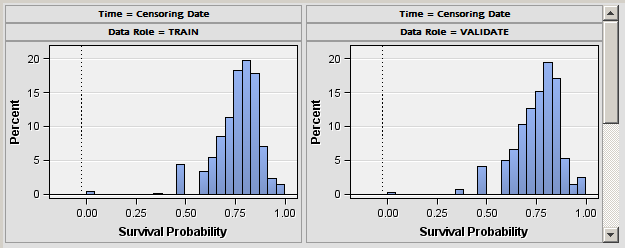
Knot size 3 was worse. Validation scores had a benefit of .158, a KS of .1965, and a Gini of .1759. These are worse than model 2.



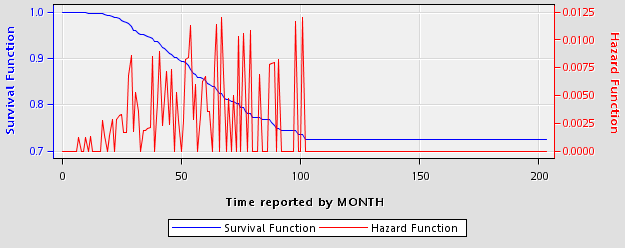


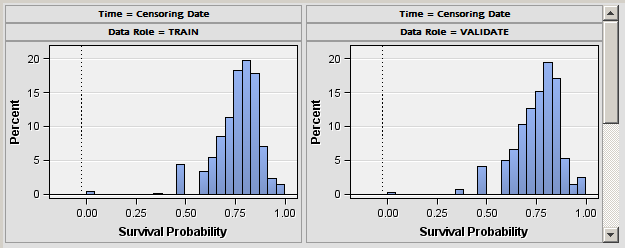
5 knots and a stepwise was worse. Benefit of .154, KS of .192, Gini of .1769.





Adding Spline Variables didn’t help either. Benefit of .154, KS of .192, Gini of .1769.





5.

Data Role ID Event Probability before or at the Future Time

VALIDATE 1256511291378000 1.0

VALIDATE 2293500251899000 0.15763391191503193

VALIDATE 1230507760896000 0.10657581696564335

VALIDATE 2902544032009000 0.030744220483508448

VALIDATE 1201535688239000 0.02594712556356757

Data Role ID Survival Probability at Future Time

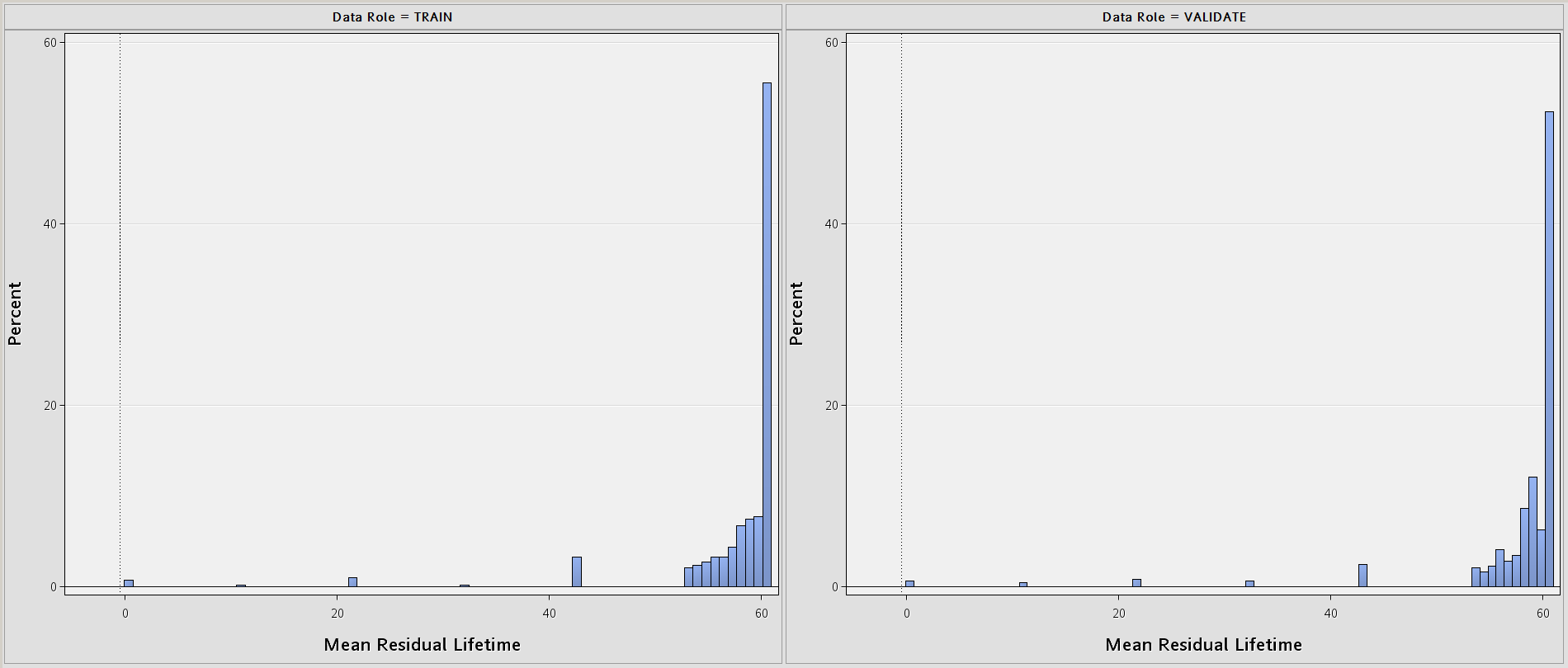
VALIDATE 1901574441129000 0.6606317254319652

VALIDATE 503913516636229000 0.6622123512223657

VALIDATE 503913514369019000 0.6622792697876659

VALIDATE 1201522670729000 0.6629527790548484

VALIDATE 1901573569659000 0.6629527790548484



Highest value for mean residual was [60.232, 61.000]