

Anderson, Raymond Ray.Anderson@standardbank.co.za

Scorecard Calibration - MAPA mapping

The scorecard calibration problem has haunted us for some time. We need to ensure that final score values provided by different scorecards have the same meaning, in terms of ranking ability, and/or predictive accuracy. Although a seemingly minor problem, a simple, consistent, and accurate means of doing so is elusive. Indeed, it is usually given fleeting treatment in credit scoring literature.

Initially, the credit scoring industry focused on ensuring consistent ranking ability, using the scorecard build definition, for different population splits within a scorecard suite. More recently, demands have arisen for scores to also: a) be consistent over successive generations of scorecard suites; and b) provide reliable default or loss probability estimates. Problems arise for defaults and losses that are rare events and cannot be modeled directly. The issue is even more pressing with Basel II, which demands a default/not default measure using 90 days past due (DPD), whereas we wish to maintain our existing scorecard methodology with good/indeterminate/bad definitions based largely around 60 DPD.

Standard Bank currently develops scorecards using a linear regression approach, developed originally by MDS and still used by Experian, and addressed the calibration problem by grouping scores into risk bands. In contrast, many other lenders use logistic regression and apply formulae, of the form $s_2 = a + b * s_1$ or similar, to adjust the score weights. These approaches are very simplistic, but sufficed.

More adequate solutions rely upon providing mapping tables at raw score level. While seemingly straightforward, the hurdle is the non-monotonic relationship between raw score and the target variable. A complex formulaic approach has been presented by Dr. Peter Gloessner (2003), but this can be difficult to interpret and implement.

The monotone adjacent pooling algorithm (MAPA) mapping method achieves the same end, but is simpler. After pooling, several steps are taken to provide an approximate fitted curve between the original score and a binary target variable, and the result is score-specific probability estimates for the latter. It is comparable to Gloessner's approach, yet none of the steps are conceptually demanding, and for small datasets MSExcel can be used.

Both approaches then present a further challenge, which is to ensure that mapping tables can be implemented within the scoring system. This may require system changes that are outside the scope of this paper.

Gloessner P. (2003). Calculating Basel II Risk Parameters for a Portfolio of Retail Loans. Thesis submitted for MSc in Mathematical Finance. Kellogg College, Oxford University.