Annex

Statistical measures of discriminatory power

ative
ıcy Profile

The CAP curve provides a graphical illustration of the discriminatory power of a rating process. For this purpose, the creditworthiness indicator (score) of every borrower is established for the dataset to be used to examine the rating model's discriminatory power. This score can be continuous, for instance the result of a discriminant analysis or a Logit regression, or it may be an integer which represents the rating grade to which the borrower has been assigned. In the following analysis, it is assumed that a high score is a reflection of a good rating. In a first step the borrowers are arranged in an ascending order of scores. The CAP curve is then determined by plotting the cumulative percentage of all borrowers ("alarm rate") on the horizontal axis and the cumulative percentage of all defaulters ("hit rate") on the vertical axis. This is shown in the adjacent chart. If, for example, those 30% of all debtors with the lowest rating scores include 70% of all defaulters, the point (0.3;0.7) lies on the CAP curve. The steeper the CAP curve at the beginning, the more accurate the rating process. Ideally, the rating process would give all defaulters the lowest scores. The CAP curve would then rise linearly at the beginning before becoming horizontal. The other extreme would be a purely random rating classification. Such a rating process would not have any discriminatory power. The expected CAP curve would, in this case, be identical to the diagonal. In reality, rating classifications are neither perfect nor random. The corresponding CAP curve therefore runs between these two extremes. Using the CAP curve, the discriminatory power of a rating process can be aggregated into a single figure, the so-called "Gini coefficient"[1] (GC). In the above chart, the area between the perfect rating and the random rating is denoted by a_P and the area between the actual rating and the random rating is denoted by a_R. The Gini coefficient is defined as the ratio of a_R to a_P, which means

$$GC = \frac{a_R}{a_P}.$$

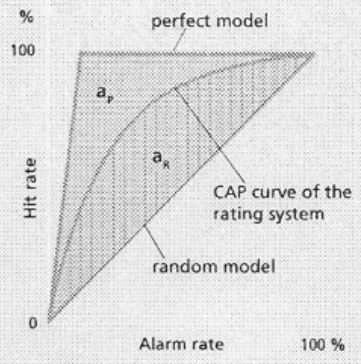
The Gini coefficient is always between minus one and one. A rating system is the more accurate the closer it is to one.

oefficient

Receiver Operating Characteristic (ROC)

The ROC curve is a concept related to the CAP curve. In order to plot this curve, the empirical score distribution for defaulters, on the one hand, and for non-defaulters, on the other, is deter-

Cumulative Accuracy Profile (CAP)* and Gini coefficient**



* For each rating score the alarm rate measures the fraction of borrowers with a lower-than-specified score within all borrowers. The hit rate gives, for each rating score, the fraction of defaulters with a lower-than-specified score within all defaulters. Processing all possible scores yields the points of the CAP curve. — ** The Gini coefficient is the quotient of the area a_R, between the CAP curve and the diagonal, and the area a_P of the shaded triangle. The larger the quotient, the more accurate the rating process.

Deutsche Bundesbank

1 The Gini coefficient is often termed the "accuracy ratio".