Journal of Risk Model Validation **15**(1), 1–18
DOI: 10.21314/JRMV.2020.231

Research Paper

Bifractal receiver operating characteristic curves: a formula for generating receiver operating characteristic curves in credit-scoring contexts

Błażej Kochański

Faculty of Management and Economics, Gdańsk University of Technology,
Gabriela Narutowicza 11/12, 80233 Gdańsk, Poland;
email: bkochanski@zie.pg.gda.pl

(Received September 4, 2019; revised March 20, 2020; accepted October 29, 2020)

ABSTRACT

This paper formulates a mathematical model for generating receiver operating characteristic (ROC) curves without underlying data. Credit scoring practitioners know that the Gini coefficient usually drops if it is only calculated on cases above the cutoff. This fact is not a mathematical necessity, however, as it is theoretically possible to get an ROC curve that keeps the same Gini coefficient no matter how big a share of lowest score cases are excluded from the calculation (a "right-hand" fractal ROC curve). Analogously, a left-hand fractal ROC curve would be a curve that keeps its Gini coefficient constant below any cutoff point. The model proposed here is a linear combination of left- and right-hand ROC curves. A bifractal ROC curve is drawn with just two parameters: one responsible for the shape of the curve and the other responsible for the area under the curve (a Gini coefficient). As is shown in this paper, most real-life credit-scoring ROC curves lie between the two fractal curves.

Print ISSN 1753-9579 | Online ISSN 1753-9587

In consequence, the Gini coefficient will be consistently lower when computed only on approved loans.

Keywords: credit scoring; Gini coefficient; lending; receiver operating characteristic (ROC) curves.

1 INTRODUCTION

Receiver operating characteristic (ROC) curve models, mathematical formulas for generating ROC curves without underlying data, have not so far been the focus of credit-scoring researchers. However, they have been researched in other domains where such curves are used, especially in signal processing and biostatistics (Egan 1975; Fawcett 2006; Swets 1996). In medical decision making, the need for ROC curve models arose from the scarcity of data: the models for the curve shape are used to enable estimation and obtain confidence intervals for the curve (Gonçalves *et al* 2014; Pepe *et al* 2009). In the credit-scoring domain, scarcity of data is usually not a key problem. Still, ROC curve formulas could be useful for simulation or stress testing ("How much will my portfolio improve if I increase the Gini coefficient of my credit scoring by x percentage points?", "What level of new loan volume increase can I expect if I improve my Gini?", "What if, due to regulatory changes, my model deteriorates?").

The literature on credit scoring does not mention models for ROC curves. The formulas introduced in biostatistics and other domains include, among others, the bilogistic (Pepe *et al* 2009), bibeta (Chen and Hu 1975) and bigamma (Dorfman *et al* 1997) models, but the binormal model, despite some problems (Bandos *et al* 2017), is considered optimal and is usually the default choice (Gonçalves *et al* 2014; Hanley 1988; Swets 1986).

The model introduced in this paper is based on two "fractal ROC curves". The idea of modeling ROC curves and the Gini coefficient in such a way came from a question that arose a few years ago during a credit-scoring workshop in Poland. An attendee at the workshop, a risk modeling manager, complained about a recommendation he had received from the audit department in his bank. An employee in the audit department raised an issue related to a new scoring model that had recently been introduced in one of the retail loan business lines. The model, when introduced, had a sufficient Gini coefficient ($= 0.45$) on both the development and validation samples and after implementation. However, when a cutoff point was introduced and the rejection rate increased, the Gini coefficient on the accepted population dropped significantly (down to around 0.32).

For this manager, the explanation was obvious: the observed Gini coefficient tends to drop when it is only calculated on those accounts that remain above the cutoff. He knew this from experience, having observed it many times, and it had never raised doubts. Industry experts are aware of this phenomenon. It is mentioned in Scallan (2013), where it is referred to as "truncation". But the audit employees were not convinced. They even claimed, based on the decreasing Gini, that the model had significantly deteriorated since implementation, and that a new scoring model was needed. The question the manager asked during the workshop was, "Can it be somehow proved that it needs to be like this and the Gini coefficient drops when the cutoff point is increased?".

One answer to this question would be to take an empirical data set and produce the ROC curve without loans below some new assumed cutoff point, and then compare the Gini coefficients before and after introducing the new cutoff. Most likely, the Gini will decrease after the new rejects are excluded from the calculation.

But does it have to be like this? As will be shown in the next section, it is not a mathematical necessity. An ROC curve does exist for any conceivable value of the Gini coefficient, such that the Gini neither decreases nor increases after introducing a cutoff. In this paper, we will call such a curve a "right-hand fractal" ROC curve. At the same time, as shown in Section 3, real-life ROC curves usually have a different shape.

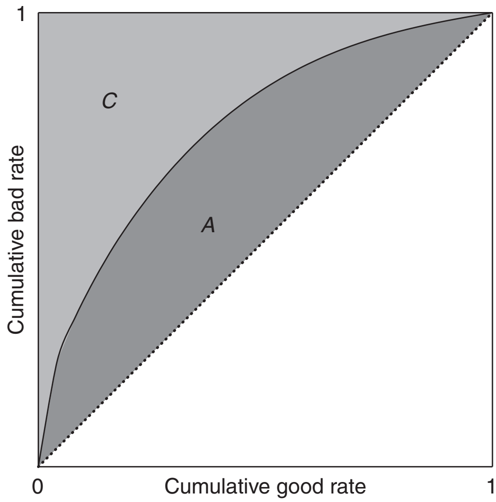
At this point, it is worthwhile mentioning that the Gini coefficient, although widely used to assess the discriminatory power of credit scorecards, is not a perfect measure. Hand and Anagnostopoulos (2013) show that the Gini is fundamentally incoherent from the perspective of misclassification costs, which results in different classifiers being treated differently. An alternative, the H-measure, is provided instead (Hand 2009). In contrast to the Gini coefficient, the proposed H-measure is not based solely on the shape of the ROC curve; it requires an assumption on the probability distribution of possible relative misclassification severities. Scallan (2013) discusses practical situations, including the shape of the ROC curve, the impact of extreme values, the reject inference method and instance sampling, when the Gini coefficient may be misleading for credit portfolio managers. Still, as the Gini coefficient remains the industry's go-to measure, we will use it in the following text despite its limitations.

In the next section, the right- and left-hand fractal curves are derived. In Section 3 the linear combination of the two curves is fitted to empirical data. Section 4 contains our conclusions and briefly discusses the limitations of the model.

2 DERIVATION OF A FRACTAL RECEIVER OPERATING CHARACTERISTIC CURVE

An ROC curve is a graphical representation of the separation power of a scoring model, or more generally, the discrimination power of some diagnostic tool. The name "receiver operating characteristic" derives from the area of electronic signal

FIGURE 1 An ROC curve.



The Gini coefficient is the ratio of the areas A and $(A + C)$.

detection, but it is also used in many other domains, including the atmospheric sciences, biosciences, experimental psychology, finance, geosciences and sociology (Gonçalves *et al* 2014). In the context of banking scoring models, the graph of an ROC curve is obtained by plotting the cumulative distribution of bad scores $F(s \mid B) = \Pr(\text{score} \leqslant s \mid B)$ on the y-axis against the cumulative distribution of good scores $F(s \mid G) = \Pr(\text{score} \leqslant s \mid G)$ on the x-axis (Anderson 2007; Siddiqi 2017; Thomas 2009).

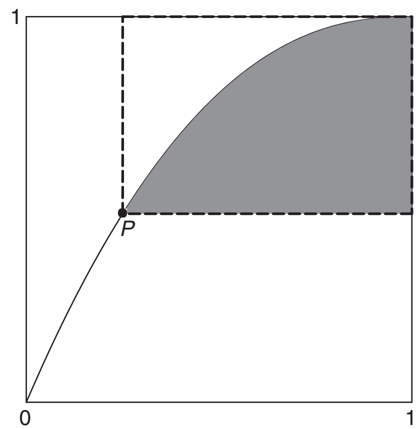
The Gini coefficient can be defined as a linear function of the area under the ROC curve (AUC) (Fawcett 2006; Hand *et al* 2001):

$$\text{Gini} = 2\text{AUC} - 1 = 2 \int F(s \mid B)\, \mathrm{d}F(s \mid G) - 1. \tag{2.1}$$

The Gini coefficient can also be expressed graphically as the ratio $A/(A + C)$ (Figure 1), where A is the area between the ROC curve (solid line) and a diagonal $y = x$ (dotted line), while $(A+C)$ is the total area of a triangle above the $y = x$ line.

If we want the Gini coefficients to be equal after introducing a cutoff no matter where we set the cutoff point, then, as shown in Figure 2, the area under the ROC curve in the 1×1 square should be equal to the area under the same curve within the rectangle defined by point P and the upper right-hand corner of the 1×1 square expressed as a fraction of the area of this rectangle.

FIGURE 2 A fractal ROC definition.



The shaded area as a fraction of the rectangle surrounded by the dashed line should be equal to the area under the curve for any cutoff point P chosen from this curve.

In other words, we are looking for a (continuous) function $f : [0, 1] \to [0, 1]$ such that $f(0) = 0$, $f(1) = 1$ and, for all $x \in [0, 1)$,

$$\frac{1}{1 - f(x)} \left\{ \frac{1}{1 - x} \int_x^1 f(\xi)\, \mathrm{d}\xi - f(x) \right\} = a, \tag{2.2}$$

where a is the AUC.

We can make the integral equation (2.2) into a differential equation by introducing a function F such that $F' = f$. Then the equation can be rewritten as

$$F'(x)(1 - a) + \frac{F(x)}{1 - x} = \frac{F(1)}{1 - x} - a. \tag{2.3}$$

Let $F(1) = c$. Consequently, we have an ordinary differential equation with two parameters, a and c:

$$F'(x)(1 - a) + \frac{F(x)}{1 - x} = \frac{c}{1 - x} - a. \tag{2.4}$$

A solution to (2.4) is given by

$$F(x) = B(1 - x)^{1/(1-a)} + c - (1 - x), \tag{2.5}$$

where B is a parameter. Since $f = F'$, we obtain

$$f(x) = \frac{B}{1 - a}(1 - x)^{a/(1-a)} + 1. \tag{2.6}$$

From the assumption $f(0) = 0$ we get $B/(1-a) = -1$:

$$f(x) = 1 - (1-x)^{a/(1-a)}, \quad x \in [0, 1]. \tag{2.7}$$

The function satisfies the initial condition $f(1) = 1$ and has just one parameter, a, which is the AUC. As the Gini coefficient is $2\text{AUC}-1$, we can rewrite the function f so that it is a function of one parameter, γ, which is equivalent to the Gini coefficient:

$$f(x) = 1 - (1-x)^{(1+\gamma)/(1-\gamma)}, \quad x \in [0, 1]. \tag{2.8}$$

Equation (2.8) shows that, for each $\gamma \in (0, 1)$ (in other words, for each value of the Gini coefficient), there exists an ROC curve that satisfies the constant Gini coefficient assumption. The Gini coefficient above any cutoff point x is equal to the Gini coefficient for the total scored population. We note that the curve is self-similar in the sense that (referring once again to Figure 2) if we "stretched" the rectangle defined by point P back to the 1×1 square, the formula for the curve would not change, irrespective of which point P is selected. More formally, for any x_0 in $(0, 1)$, for any x in $(x_0, 1)$ and for any $0 \leqslant \gamma < 1$, the following equality will be true:

$$\frac{f(x) - f(x_0)}{1 - f(x_0)} = f\left(\frac{x - x_0}{1 - x_0}\right). \tag{2.9}$$

This feature may bring to mind the concept of fractals, which remain self-similar however they are magnified (Mandelbrot 1983). It may also be noted that the Gini coefficient remains the same above, but not below, any cutoff point. That is why the term "right-hand fractal curve" seems appropriate for such a curve. Figure 3 plots the fractal ROC curves for several values of γ.

As can be shown, an analogical left-hand fractal ROC curve can be derived. The integral equation for it is

$$\frac{1}{xg(x)}\left\{\int_0^x g(\xi)\,\mathrm{d}\xi\right\} = a = \frac{\gamma + 1}{2}, \tag{2.10}$$

and we get the following solution:

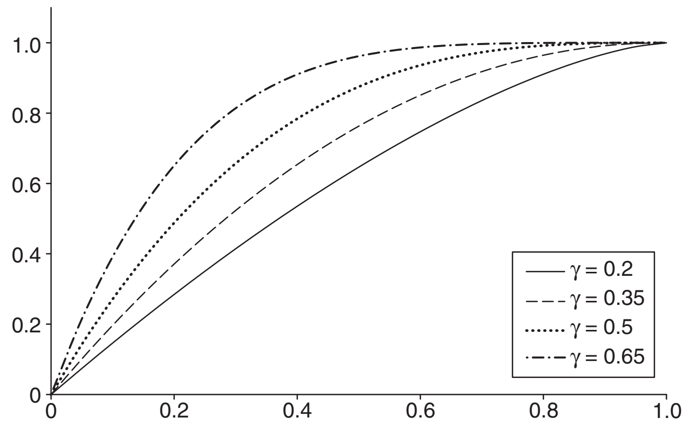
$$g(x) = x^{(1-\gamma)/(1+\gamma)}, \quad x \in [0, 1]. \tag{2.11}$$

We may show the self-similarity property is analogous to (2.9):

$$\frac{g(x)}{g(x_0)} = g\left(\frac{x}{x_0}\right). \tag{2.12}$$

Note that the function $g(x)$ specifying a left-hand fractal ROC curve is an "algebraic ROC", namely a "power ROC" as described, among other things, by Swets (1996). The right-hand fractal ROC could also be included in the category of "algebraic ROCs" as defined by Swets (1996), but, to the best of the author's knowledge, has not appeared in the literature before.

FIGURE 3 A right-hand fractal ROC curve for several gammas (Gini coefficients).



3 MODELING EMPIRICAL RECEIVER OPERATING CHARACTERISTIC CURVES

In practice, empirical ROC curves are neither right- nor left-hand fractal but somewhere in between. For example, take the data provided in Řezáč and Řezáč (2011) from an unnamed European financial institution. Figure 4 plots empirical data points against two fractal ROC curves. (Note that in order to draw the fractal ROC curves, the Gini coefficient (γ parameter) had to be assumed. The Gini coefficient calculated based on the data was taken to be $\gamma = 0.46$.) As can be seen, the empirical ROC curve lies between the two extremes.

This brings us to the idea of a linear combination of the two fractal ROC curves as a simple way to model empirical ROC curves. If we combine (2.8) and (2.11), we get

$$\begin{aligned} r_{\beta,\gamma}(x) &= \beta f(x) + (1-\beta)g(x) \\ &= \beta\{1-(1-x)^{(1+\gamma)/(1-\gamma)}\} + (1-\beta)x^{(1-\gamma)/(1+\gamma)}. \end{aligned} \tag{3.1}$$

In this setup, a real-life ROC curve can be modeled with just two parameters. One of them (γ) represents the Gini coefficient, and the other (β) is responsible for the shape of the curve. Formula (3.1) returns a left-hand fractal ROC curve for $\beta = 0$ and a right-hand fractal ROC curve for $\beta = 1$.

Once we have (3.1), we can fit the curve to empirical data points. To obtain this goal, a minimal average weighted distance algorithm is proposed. To fit the curve we use the "bobyqa" function (Powell 2009) from the R package `minqa` (Bates *et al*

FIGURE 4 An empirical ROC curve from Řezáč and Řezáč (2011) versus right- and left-hand fractal ROC curves.

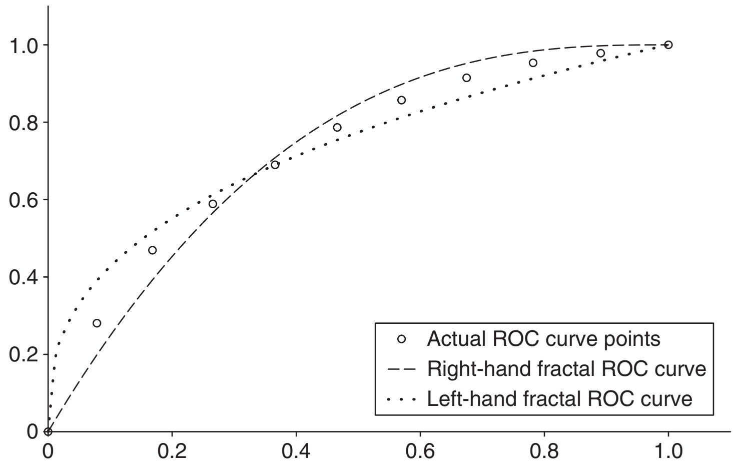


FIGURE 5 Setting weights for the "bobyqa" objective function.

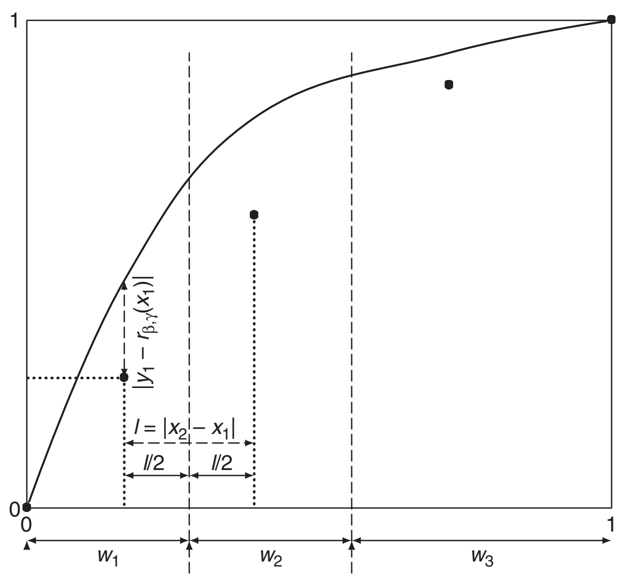
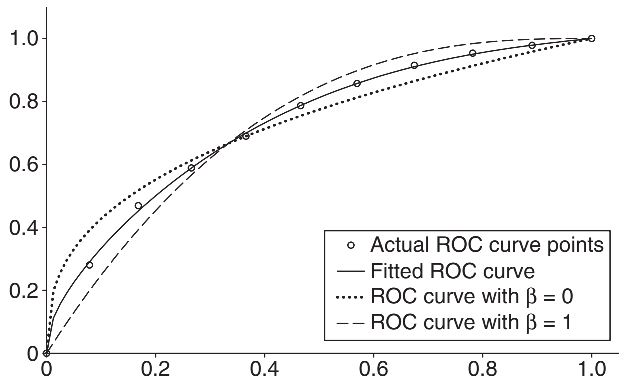


FIGURE 6 A model ROC curve from (3.1) fitted to the data from Řezáč and Řezáč (2011).



$\beta = 0.5241703$, $\gamma = 0.4603697$. $f_{\mathrm{obj}} = 0.004799454$.

2014). We choose the following objective function to be minimized:

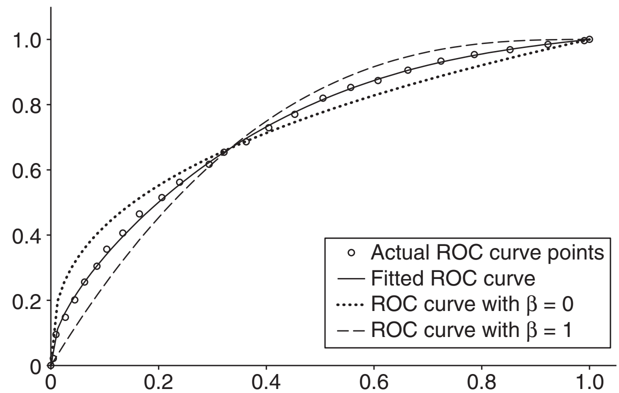
$$f_{\mathrm{obj}}(\boldsymbol{x}, \boldsymbol{y}, \beta, \gamma) = \sum_{i=1}^{n} |y_i - r_{\beta,\gamma}(x_i)| w_i, \tag{3.2}$$

where $|\cdot|$ denotes the absolute value; w_i are weights, as explained below; x_i and y_i are the coordinates of points from the empirical ROC curve; $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors containing these coordinates; and $r_{\beta,\gamma}(x_i)$ is the vertical coordinate of point x_i calculated with (3.1). Points $(0, 0)$ and $(1, 1)$ are not included in the calculation. The weights w_i are based on vector $\boldsymbol{x}$ in the way illustrated in Figure 5. Namely, for each element x_i of vector $\boldsymbol{x}$, w_i is the length of the section consisting of those x-axis points between 0 and 1 that are closer to x_i than to any other element of $\boldsymbol{x}$. In this way, if only $\boldsymbol{x}$ has no duplicates, all the w_i are greater than zero and $\sum_i w_i = 1$.

Figure 6 contains an ROC curve fitted to the data from Řezáč and Řezáč (2011) with the method described above. The parameter β turns out to be around 0.52, so the real ROC curve is almost exactly in the middle of the left- and right-hand fractal curves. Based on the plot, we may infer that the goodness-of-fit is quite good in this case. The objective function (f_{obj}) is below 0.005, which means that the weighted average absolute deviation between the empirical data and the fitted model is below this level.

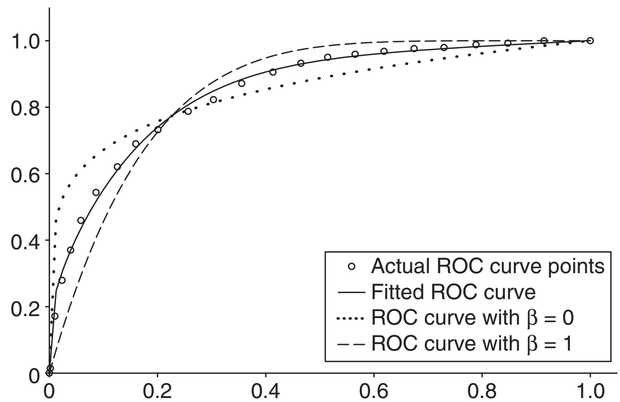
Eleven data points are available in this chart. We may use the chart with a Lorenz curve from Řezáč and Řezáč (2011) to read more data points regarding the same scoring model. Figure 7 presents the results. It seems that both the parameters (β

FIGURE 7 A curve fitted to more data points read from the Lorenz curve in Řezáč and Řezáč (2011).



$\beta = 0.5200344$, $\gamma = 0.4608602$. $f_{\text{obj}} = 0.005274939$.

FIGURE 8 A curve fitted to the empirical ROC curve found in Wójcicki and Migut (2010).

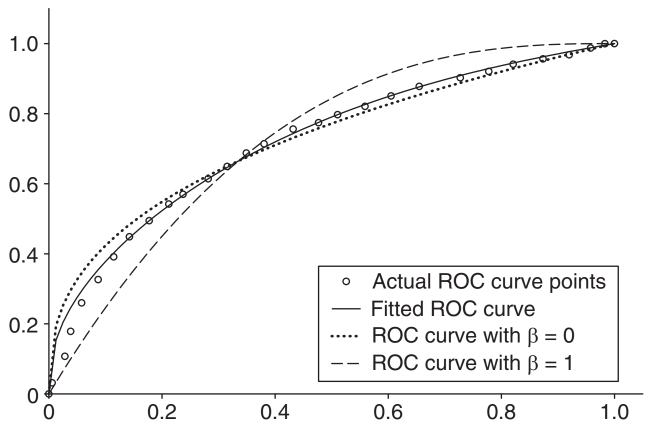


$\beta = 0.5550773$, $\gamma = 0.705437$. $f_{\text{obj}} = 0.01095339$.

and γ) and the objective function are almost the same as in the previous calculation, which consists of a smaller number of data points.

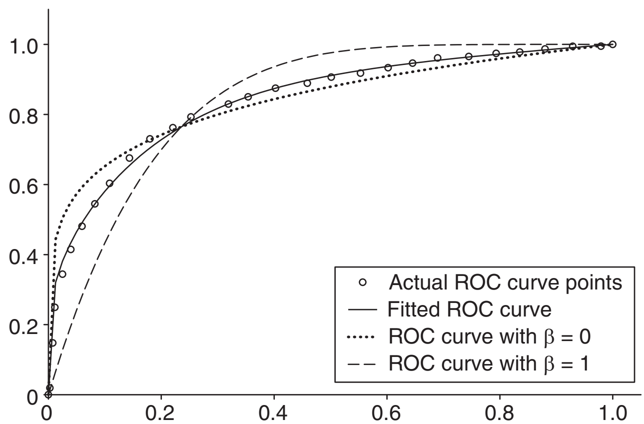
Figure 8 contains an ROC curve fitted to the empirical ROC curve from an anti-fraud scoring model developed in a Polish bank (Wójcicki and Migut 2010). Again, the fit seems to be quite good, and the beta is close to one-half (0.56).

FIGURE 9 A curve fitted to the empirical data on Altman's Z-score from Engelmann and Tasche (2003).



$\beta = 0.256409$, $\gamma = 0.4562762$. $f_{\text{obj}} = 0.01010763$.

FIGURE 10 A curve fitted to the empirical data on a logit score from Engelmann and Tasche (2003).



$\beta = 0.3269145$, $\gamma = 0.6869269$. $f_{\text{obj}} = 0.008575946$.

Engelmann and Tasche (2003) present ROC curves calculated for the Altman Z-score model and a self-developed logit model based on the Bundesbank database of small and medium enterprises (years 1987–99). Figures 9 and 10 present the

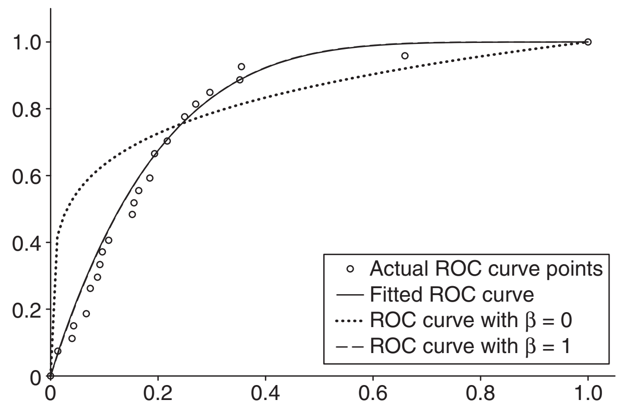
TABLE 1 Results of fitting a linear combination of two fractal ROC curves to empirical ROC curves. [Table continues on next page.]

Source	Shape parameter (β)	Gini parameter (γ)	Goal function (f_{obj})	Comments
Wójcicki and Migut (2010)	0.555	0.705	0.011	Anti-fraud scoring model developed in a Polish bank
Engelmann and Tasche (2003)	0.327	0.687	0.009	A logit model for German small and mid-size enterprises based on the Bundesbank database
Engelmann and Tasche (2003)	0.256	0.456	0.010	Altman's Z-score on German small and mid-size enterprises
Řezáč and Řezáč (2011)	0.524	0.460	0.005	Unnamed European financial institution
Řezáč and Řezáč (2011)	0.520	0.461	0.005	As above, but with more data read from a Lorenz curve graph
Beling *et al* (2005)	0.268	0.570	0.018	A "CB (credit bureau?)" score of unknown origin
Beling *et al* (2005)	0.734	0.552	0.005	An application score of unknown origin
Bernardo *et al* (2013)	0.221	0.608	0.023	Credit card model on the data from an unnamed credit card institution – genetic algorithm

TABLE 1 Continued.

Source	Shape parameter (β)	Gini parameter (γ)	Goal function (f_{obj})	Comments
Bernardo *et al* (2013)	0.317	0.671	0.014	Credit cards – neural networks
Miyamoto (2014)	0.989	0.668	0.035	A model based on logistic regression – small bank, individual entrepreneurs
Miyamoto (2014)	0.606	0.452	0.012	As above – small businesses
Based on UCI data (Lichman 2013)	0.112	0.529	0.008	Logistic regression model based on the data from Lichman (2013)
Zehori (2016)	0.419	0.721	0.008	A model built based on the data from Kaggle's "Give me some credit" competition
Anonymous 1	0.522	0.307	0.006	An ROC curve for a scoring model used by lender 1
Anonymous 2	0.545	0.442	0.006	An ROC curve for a credit-scoring model used by lender 2

FIGURE 11 A curve fitted to the empirical data on a model for individual entrepreneurs from Miyamoto (2014).



$\beta = 0.9888473$, $\gamma = 0.6683853$. $f_{\text{obj}} = 0.03521597$.

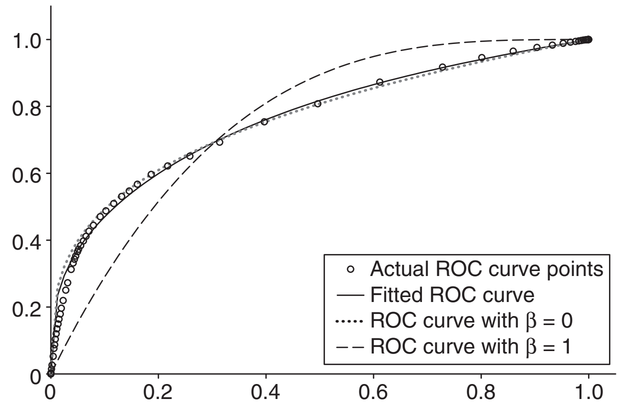
results of the curve fitting for these cases. Here, the β parameters are 0.26 and 0.33, respectively.

It is quite difficult to find publicly available data for empirical ROC curves. And even if such information is available in the literature, in most cases it is necessary to extract it from the graph itself. Two consumer-lending institutions kindly provided their real-life ROC curves for this exercise under the condition of anonymity. Also, data from Lichman (2013) was used to build a simple logistic regression classifier. The results for the above-mentioned examples and several others found in the literature are presented in Table 1. The overall fit appears to be quite good. The goal function does not exceed 0.035 in any of the cases, which means that, at worst, the weighted average vertical distance between the data points and the fitted curve is 3.5 percentage points. The median result of f_{obj} is around 1 percentage point.

The case of the Miyamoto (2014) model for individual entrepreneurs in a small bank is an interesting one. The fitted curve here is almost the right-hand fractal curve, as the shape parameter is very close to 1 ($\beta = 0.989$, Figure 11). As it appears, a right-hand fractal ROC curve, where the Gini does not drop with the introduction of the cutoff, not only is a theoretical possibility but also may occur in practice. A curve quite close to the left-hand fractal ROC curve is presented in Figure 12 ($\beta = 0.112$).

However, except for these two cases, the beta parameters lie between 0.25 and 0.75, which shows that real-life ROC curves fall somewhere in the middle of the left- and right-hand fractal models. This would be an important conclusion for the

FIGURE 12 A curve fitted to ROC points for a scorecard built on data from Lichman (2013).



$\beta = 0.1119022$, $\gamma = 0.5294155$. $f_{\mathrm{obj}} = 0.008194536$.

audit department employees mentioned in Section 1. In most situations, they should expect the Gini coefficient to drop when it is computed on approved loans after the introduction of a substantially higher cutoff point.

4 CONCLUSIONS

The relatively simple two-parameter formula derived from two "fractal" ROC curves,

$$y = \beta\{1 - (1 - x)^{(1+\gamma)/(1-\gamma)}\} + (1 - \beta)x^{(1-\gamma)/(1+\gamma)},$$

seems to be a useful and intuitive model for empirical ROC curves associated with credit-scoring Gini coefficients. The γ parameter in the model reflects the value of the Gini coefficient, while the other parameter (β) serves as a shape parameter.

When the model curve is fitted to actual data, the beta parameter turns out to be around the midpoint (between 0.25 and 0.75). The empirical ROC curves lie between the left- and right-hand fractal ROC curves. As a consequence, there is almost always a decrease in the Gini coefficient when a more strict cutoff point is introduced. The Gini will be consistently lower when computed only on approved loans.

The model can be used to simulate new, unknown ROC curves or to simplify calculations for existing ROC curves. For example, it could help answer the question of how much the profits of a lending institution may be improved after the introduction of a new scorecard with increased separation power. This is a typical scenario when the need arises to draw a hypothetical ROC curve.

Certainly, the model has some limitations. Most importantly, the usefulness of the model relies on the goodness-of-fit to actual ROC data. As shown in Section 3, the fit to empirical data is quite good, although it may differ for various cases, and alternative options (such as the binormal curve) may turn out to be more adequate.

DECLARATION OF INTEREST

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

I thank Professor Karol Dziedziul from Gdańsk University of Technology for his guidelines on solving integral equations introduced in the paper. I sincerely thank the anonymous reviewers for their helpful comments on the manuscript.

REFERENCES

Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford University Press.

Bandos, A. I., Guo, B., and Gur, D. (2017). Estimating the area under ROC curve when the fitted binormal curves demonstrate improper shape. *Academic Radiology* **24**(2), 209–219 (https://doi.org/10.1016/j.acra.2016.09.020).

Bates, D., Mullen, K. M., Nash, J. C., and Varadhan, R. (2014). Minqa: derivative-free optimization algorithms by quadratic approximation. R package, Version 1.2.4. URL: https://CRAN.R-project.org/package=minqa.

Beling, P., Covaliu, Z., and Oliver, R. M. (2005). Optimal scoring cutoff policies and efficient frontiers. *Journal of the Operational Research Society* **56**(9), 1016–1029 (https://doi.org/10.1057/palgrave.jors.2602021).

Bernardo, D., Hagras, H., and Tsang, E. (2013). A genetic type-2 fuzzy logic based system for financial applications modelling and prediction. In *Proceedings of the 2013 IEEE International Conference on Fuzzy Systems*, pp. 1–8. IEEE Press, Hoboken, NJ (https://doi.org/10.1109/fuzz-ieee.2013.6622310).

Chen, W., and Hu, N. (1975). Proper bibeta ROC model: algorithm, software, and performance evaluation. In *Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment*, Abbey, C. K., and Kupinski, M. A. (eds), pp. 97–104. SPIE Proceedings, Volume 9787. SPIE, San Diego, CA (https://doi.org/10.1117/12.2216777).

Dorfman, D. D., Berbaum, K. S., Metz, C. E., Lenth, R. V., Hanley, J. A., and Abu Dagga, H. (1997). Proper receiver operating characteristic analysis: the bigamma model. *Academic Radiology* **4**(2), 138–149 (https://doi.org/10.1016/s1076-6332(97)80013-x).

Egan, J. P. (1975). *Signal Detection Theory and ROC Analysis*. Academic Press, New York.

Engelmann, B., and Tasche, D. (2003). Testing rating accuracy. *Risk* **16**(1), 82–86.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters* **27**(8), 861–874 (https://doi.org/10.1016/j.patrec.2005.10.010).

Gonçalves, L., Subtil, A., Oliveira, M. R., and de Zea Bermudez, P. (2014). ROC curve estimation: an overview. *REVSTAT: Statistical Journal* **12**(1), 1–20.

Hand, D. J. (2009). Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Machine Learning* **77**(1), 103–123 (https://doi.org/10.1007/s10994-009-5119-5).

Hand, D. J., and Anagnostopoulos, C. (2013). When is the area under the receiver operating characteristic curve an appropriate measure of classifier performance? *Pattern Recognition Letters* **34**(5), 492–295 (https://doi.org/10.1016/j.patrec.2012.12.004).

Hand, D. J., and Till, R. J. (2001). A simple generalization of the area under the ROC curve to multiple class classification problems. *Machine Learning* **45**(2), 171–186 (https://doi.org/10.1023/A:1010920819831).

Hanley, J. A. (1988). The robustness of the "binormal" assumptions used in fitting ROC curves. *Medical Decision Making* **8**, 197–203 (https://doi.org/10.1177/0272989x8800800308).

Lichman, M. (2013). UCI Machine Learning Repository, University of California, School of Information and Computer Science Irvine, CA. URL: http://archive.ics.uci.edu/ml.

Mandelbrot, B. (1983). *The Fractal Geometry of Nature*. W. H. Freeman, San Francisco, CA.

Miyamoto, M. (2014). Credit risk assessment for a small bank by using a multinomial logistic regression model. *International Journal of Finance and Accounting* **3**(5), 327–334 (https://doi.org/10.5923/j.ijfa.20140305.07).

Pepe, M., Longton, G., and Janes, H. (2009). Estimation and comparison of receiver operating characteristic curves. *Stata Journal* **9**(1), 1–16 (https://doi.org/10.1177/1536867x0900900101).

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Technical Report DAMTP 2009/NA06, Centre for Mathematical Sciences, University of Cambridge. URL: https://bit.ly/3awxHU5.

Řezáč, M., and Řezáč, F. (2011). How to measure the quality of credit scoring models. *Finance a Uvěr: Czech Journal of Economics and Finance* **61**(5), 486–507.

Scallan, G. (2013). Why you shouldn't use the Gini. Talk presented at the ARCA Retail Credit Conference, Leura, Australia. URL: https://bit.ly/3axcMQR.

Siddiqi, N. (2017). *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards*, 2nd edn. Wiley (https://doi.org/10.1002/9781119282396).

Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. *Psychological Bulletin* **99**(2), 181–198 (https://doi.org/10.1037/0033-2909.99.2.181).

Swets, J. A. (1996). *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*. Psychology Press, New York (https://doi.org/10.4324/9781315806167).

Thomas, L. C. (2009). *Consumer Credit Models: Pricing, Profit and Portfolios*. Oxford University Press (https://doi.org/10.1093/acprof:oso/9780199232130.001.1).

Wójcicki, B., and Migut, G. (2010). Wykorzystanie skoringu do przewidywania wyłudzeń kredytów w Invest-Banku. In *Skoring w Zarządzaniu Ryzykiem*, pp. 47–57. Statsoft, Kraków. URL: https://bit.ly/2YDGmhU.

Zehori, I. (2016). Credit-score: credit scoring by predicting the probability that somebody will experience financial distress in the next two years. Data from the kaggle “give me some credit” competition. URL: https://github.com/IdoZehori/Credit_Score.