Scientific Research Publishing

Theoretical Economics Letters, 2025, 15(5), 1251-1266
https://www.scirp.org/journal/tel
ISSN Online: 2162-2086
ISSN Print: 2162-2078

Gini Coefficient and AUC in Assessing Predictive Model Performance: Effect of Ranks

Erkki K. Laitinen

ACA Research Group, Faculty of Business Studies, University of Vaasa, Vaasa, Finland
Email: ekla@uwasa.fi

How to cite this paper: Laitinen, E. K. (2025). Gini Coefficient and AUC in Assessing Predictive Model Performance: Effect of Ranks. *Theoretical Economics Letters, 15,* 1251-1266.
https://doi.org/10.4236/tel.2025.155070

Received: March 30, 2025
Accepted: October 19, 2025
Published: October 22, 2025

Abstract

This paper deals with three traditional measures of concentration: (Corrado) Gini coefficient, adjusted Gini coefficient, and AUC. These metrics are popular methods in assessing the performance of predictive models, like credit scoring models. They are non-parametric variables and therefore only depending on the ranking of events. The three measures are closely related to each other. The adjusted Gini coefficient (Accuracy ratio, AR) is only a transformation of the Corrado Gini coefficient being in a linear relationship with AUC. It also equals to Somers' *D*. This paper also introduces the measure *E*, which is based on a classification of the ranks of events. *E* produces the same result as AUC, but is simple to calculate and interpret. The features of the metrics are discussed in three numerical examples, one of which deals with credit scoring in a large imbalanced dataset (23.533 active firms and 147 bankrupt firms). Numerical examples are used to illustrate the properties of the metrics, especially at the level of ranks.

Keywords

Credit Scoring, Concentration, Gini Coefficient, AUC, ROC, Imbalanced Sample

1. Introduction

The purpose of this paper is to analyse a formulation of the Gini coefficient and to present examples of its use in evaluating the quality of a predictive model. The research on the Gini coefficient has a long history. More than 100 years ago, Corrado Gini (1912) developed this best-known measure of concentration based on the Lorenz curve of income distribution. Thus far, there is a very large body of research on the Gini coefficient and similar measures of concentration. In their extensive literature review, Giorgi & Gigliarano (2017: p. 1144) concluded that the

Gini coefficient is still of great interest for the international scientific community. After the long history, there is still an intensive scientific production on this coefficient in the last decades confirming the long wave of its interest. There are a number of studies introducing novel extensions, interpretations, and use of the Gini coefficient. In addition, various transformations to improve the properties of the coefficient have been presented (for example, Bowles & Carlin, 2020). However, there is still room for new analyses and approaches on the coefficient.

The Gini coefficient can also be called the accuracy ratio (AR), because it is the summary index of cumulative accuracy profile (CAP). In the scientific research on concentration measures, the connection of the Gini coefficient to the area under the ROC curve (AUC), has been intensively analyzed. Generally, the adjusted Gini coefficient is in a linear relationship to AUC as 2 AUC-1 (Adeodato & Melo, 2022; Schechtman & Schechtman, 2019). AUC is a very common metric also in medicine, demography, and similar sciences in addition to economics (for example, Kohl, 2012; Nahm, 2022). AUC has also been criticized as a measure of concentration. Lobo, Jiménez-Valverde and Real (2008) state that the good side of AUC is that in predictive models it avoids the supposed subjectivity in the threshold selection process, summarizing overall model performance over all possible thresholds. However, they present five reasons why they do not recommend using AUC (Lobo, Jiménez-Valverde and Real, 2008: p. 145). The Gini coefficient is also a special case of Somers' D (Newson, 2006). The common characteristic of these concentration measures is that they are ordinal (non-parametric) measures.

The Gini coefficient and AUC are very popular measures for evaluating the accuracy of predictive models in credit rating (Crook, Edelman, & Thomas, 2007; Kumar & Ravi, 2007; Langohr & Langohr, 2009; Lessmann, Baesens, Seow, & Thomas, 2015). In these approaches, the question is how the concentration metrics are able to assess the results produced by the predictive model, where "bad events" (desired outcome) and "good events" are placed in rank order. The goal is that bad events should concentrate at the top of the order as strongly as possible. Then, concentration measures are used to assess the distribution of bad events in the rank order. The most popular measures in this context are the Gini coefficient and AUC followed by Somers' D. This brief paper also considers a predictive model developed with logistic regression analysis as an example. The application area of this model is credit rating trying to separate bad events from good events as efficiently as possible. In addition to credit rating, the use of predictive models and concentration measures is also very extensive in many other areas. In economics, they have been applied, for example, to identifying different types of fraud and customer purchasing behavior (Joshi, Gupte, & Saravanan, 2018; Moon, Pu, & Ceglia, 2019).

The Gini coefficient and related other measures of concentration (ROC, AUC, D) have been evaluated in many studies by researchers in different fields. Therefore, it is challenging to develop novel extensions, interpretations, or use of the Gini coefficient as urged by Giorgi & Gigliarano (2017: p. 1144). In this paper, the traditional graphics associated with the Lorenz curve and the cumulative accuracy profile (CAP) are not explicitly used to derive and illustrate the Gini coefficient, as they are generally known. Instead, the starting point of the approach is a simple formulation of the Gini coefficient. Giorgi & Gigliarano (2017) presented a review summarizing the formulations of the coefficient. The formulation of the coefficient used in this study is closely related to the formulations presented by Sen (1973) and Cicchitelli (1976) (see Giorgi & Gigliarano, 2017). This formula is interpreted in the theoretical part of the paper, where the object of examination is above all the effect of individual events and their transitions in the rank order, on the Gini coefficient. In addition to this, a novel (as far as we know) measure E of concentration is presented. The advantage of E is that its interpretation and calculation is straightforward. In this paper, also numerical examples are presented to illustrate the results of the theoretical part.

The content of this paper is divided into four short sections. In the first section, the motivation and content of the paper was discussed briefly. The second section (framework or theoretical part) examines the properties of the mathematical formulation of the Gini coefficient and derives the new measure E. This analysis is descriptive and, as a restriction, any formal statistical inference is not used. The third section presents a set of numerical examples to illustrate the application of the approach. In this context, three separate examples are considered: two small sample and one big sample examples. However, the idea of the numerical cases is identical in each example: to show the contribution of events and their transitions along the ranking to the Gini coefficient. Finally, a short summary of the study is presented in the fourth section.

2. Framework of the Study

In 1912, Corrado Gini (1912) presented the original Gini coefficient or concentration index (G) for measuring concentration of income. The general formulation of G is as follows:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left|x_i - x_j\right|}{mn^2} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2mn^2} \tag{1}$$

In this formulation (1), n is the sample size, x_i is the value of event i of the variable under study, and m is the average value of the variable in the sample. Thus, the Gini coefficient G measures the average difference between the values of the variable in relation to the average of the values. The order in which the differences between the values are calculated does not matter: all orders give the same Gini coefficient. Thus, in this general formulation the rank order of events does not affect G.

In assessing the quality of a predictive model, the focus is set on the ranks. Therefore, the general formulation of G in Equation (1) is not sufficient for the analysis. Let us assume that in a sample of n items, we have n_1 observations of "bad events" (BE) and n_0 observations of "good events" (GE) so that $n_1 + n_0 = n$.

It is also assumed that we have a predictive model based on a set of relevant independent variables, which estimates for each event a conditional probability or some other indicator ("core", event risk) to belong to the group of BEs. Then, the events are ranked on the basis of the event risk in descending order. In this context, it essential to see how BEs are concentrated in the sample in relation to GEs. The quality of the predictive model is considered the higher, the stronger BEs are concentrated at the top of the rank order. The question is how to measure the strength of this concentration.

The explicit analysis of concentration is focused here on the ranks of BEs. However, also GEs play an important role in concentration, since they collectively calibrate the level of concentration of BEs by affecting their positions in the rank order and the sample size as $n = n_1 + n_0$. The more the sample contains GEs, the lower the relative number of BEs, the default rate, and the concentration of risky events in the sample. Because the Gini coefficient G is an ordinal measure, it is independent of the event risk produced by the model, but fully determined by the ranks based on this risk. In typical credit rating samples, BEs refer to failed (bankrupt, default) companies while GEs refer to non-failed (acting, non-default) companies. The number of BEs n_1 is usually small in comparison to the number of GEs n_0 and the share of BEs of n sample companies describes the default rate in the sample.

There exist numerous mathematical formulations of the Gini coefficient, derived through various inference approaches (Giorgi & Gigliarano, 2017). Most of these formulations are mathematically complex. However, when the sample firms are arranged in descending order according to their estimated event risk, the expression of the Gini coefficient in Equation (1) can be substantially simplified. In this case, the Gini coefficient can be represented as a ratio of two linear combinations of order statistics, as discussed by Sen (1973), Cicchitelli (1976), and Giorgi & Gigliarano (2017). When solving for the sum of ordinal series in this ratio, the following simplified result is obtained:

$$G = \frac{\sum_{i=1}^{n} (n+1-2i) x_i}{mn^2} \tag{2}$$

This formulation (2) clearly shows that G is consisted of a weighted sum of the variable x_i where the weight will diminish as a function of the position i of the events on the rank order scale.

The situation is further simplified assuming that x_i is a binary variable so that $x_i = 0$ when the event is GE and, correspondingly, $x_i = 1$ when the event is BE. In this situation, using the sum of ranks for BEs the formulation (2) can be transformed into the following formulation:

$$G = \frac{n_1(n+1) - 2 \cdot S}{n_1 n} = 1 + \frac{1}{n} - \frac{2 \cdot S}{mn^2} = 1 + \frac{1}{n} - \frac{2 \cdot S}{n_1 n} \tag{3}$$

In this Equation (3) S is the sum of the ranks for BEs. The formulation (3) clearly shows that G is a simple function of ranks of GEs, default rate, and the size of the sample. If the sum of ranks increases by one, from S to $S+1$, G will decrease by $2/(n_1 n)$.

The separate effect of a single event i on the concentration of the entire sample (the Gini coefficient) is thus the following:

$$G(i)=\frac{(n+1-2i)x_i}{mn^2} \tag{4}$$

In the Equation (4), $G(i)$ refers to the effect of event i on the coefficient G. G is the sum of $G(i)$ over the events in the sample ($i = 1, 2, \ldots, n$). Since $x_i = 0$ for the GEs, also $G(i) = 0$ for them. Thus, in this sense GEs do not seem to have a direct effect on G. However, for BEs with $x_i = 1$, the position i of the event on the rank order determines its effect on G. Equation (4) shows that if the position of BE is $i = (n+1)/2$, the effect $G(i)$ is zero. Therefore, a BE in this position i neither increases nor decreases G. For BEs in ranks lower than this middle point, the effects on G turn to be negative.

For the interpretation of G, it is important to understand how G will change if BEs move forward or backward on the rank order. Equation (4) indicates that G will change in the following way, when a BE moves from position i to position j:

$$G(j)-G(i)=\frac{2(i-j)}{mn^2}=\frac{2(i-j)}{n_1 n} \tag{5}$$

Equation (5) shows logically that the absolute change in G is directly proportional to how many positions BE on position i moves to position j. However, this change is independent of the positions of i and j in the rank order and only the difference between i and j is important. Thus, the effect is equal for each equal difference i-j in the sample. The effect of the move on G is the lower, the larger are the sample size and the default rate in the sample. In addition, GEs affect the ranking of events by the same weight as BEs do: the effect of a transition in ranking on G is equal for both GEs and BEs. Therefore, GEs factually have a significant impact on the change in G.

The ratio of the separate effects of BEs which are located on positions i and j can through (3) presented in the following formulation:

$$\frac{G(i)}{G(j)}=\frac{1}{1-\frac{2(j-i)}{n+1-2i}} \tag{6}$$

Equation (6) shows the effect of sample size n on the ratio of the effects. The ratio of the effects is small when j-i is small and n large. This ratio is also dependent on the rank i in addition to the difference j-i indicating that the ratio of the effect on G is related to the rank. Note that when $i = (n+1)/2$ (middle point), the ratio in (6) is undefined. Then, after that point, the sign of the ratio will change.

In the formulation (2), the Gini coefficient cannot attain its theoretical maximum value of 1, but it remains always strictly less than unity. If in the extreme case there is only one BE located at the top of ranking, then the concentration is at its maximum value. Formulation (2) shows that in that case G is $(n-1)/n$ for any n when the rank $i=1$, $x_1=1$, and $n_1=1$. In this maximum concentration case, $x_i=0$ for $i>1$ making the maximum value independent of n. Thus, G cannot exceed unity for any n. However, the coefficient G can be adjusted to make its maximum value equal to 1. In the extreme case above, this can be done by multiplying G by its reciprocal $n/(n-1)$, i.e. by dividing G by the factor $(n-1)/n$, which is the proportion of GEs in the sample. This result generally applies to the Gini coefficient, that can be adjusted in this way to reach 1. The adjustment can thus be made dividing G by the proportion of GEs in the sample, i.e. by (1-default rate). This adjustment does not make G to exceed unity in any value of n. The minimum value of the Gini coefficient is zero, which is in this framework reached, when all $x_i=0$.

Thus, the adjusted Gini coefficient G^* can be presented in the following formulation:

$$G^* = \frac{G}{1-\frac{n_1}{n}} = \frac{G}{1-m} = D = 2\cdot \mathrm{AUC}-1 \tag{7}$$

In this formula (7) the original (Corrado) Gini coefficient G has been adjusted as G^* to reach the maximum value of unity. This equation also presents the well-known result that the adjusted Gini coefficient is equal to Somers' D statistic. G^* has also a simple linear relationship to AUC (Area under ROC curve), since AUC is simply $(G^*+1)/2$. Thus, the three popular measures of concentration are closely related to each other.

The Gini coefficient G, G^*, and AUC are ordinal measures of concentration being completely based on the positions (ranks) of BEs in the sample. This feature of concentration metrics can be illustrated by three sums of ranks, which are as follows:

$$S = \sum\nolimits_{i=1}^{n} x_i \cdot i \tag{8a}$$

$$L = \sum\nolimits_{i=1}^{n_1} i = \frac{n_1\left(n_1+1\right)}{2} \tag{8b}$$

$$H = \sum\nolimits_{i=n-n_1+1}^{n} i = \frac{n_1\left(n_1+1\right)}{2} - n_1^2 + n\cdot n_1 \tag{8c}$$

The first of these rank sums (S) indicates the actual sum of the ranks of the BE events in the sample. The second sum (L) is the lowest possible sum of ranks for n_1 BEs, which is realized when all BEs are ranked on the n_1 first positions in the ranking: in this case, concentration is the highest possible. However, the third sum (H) is the highest possible sum of BE ranks, which is realized when all BEs are ranked on the n_1 last positions on the ranking: concentration is then the lowest possible.

Using the presented three sum formulas of ranks S, L, and H, the following measure E for concentration can be presented:

$$E = 1 - \frac{S-L}{H-L} = 1 - \frac{\sum_{i=1}^{n} x_i \cdot i - \frac{n_1(n_1+1)}{2}}{n_1 \cdot n - n_1^2} = \text{AUC} \tag{9}$$

The interpretation of the indicator E is logical in measuring concentration. If BEs are ranked at the top of the ranking order and ranks get their lowest value, S-L = 0 and E = 1 referring to the highest rate of concentration. If the BEs are ranked on the last positions of ranking scale, then $S = H$ and S-L = H-L, in which case E = 0 which refers to the lowest rate of concentration. When comparing the concentration measures E in (9) and AUC, it is found that they are equal, i.e. E = AUC. Equation (9) thus gives a usable formulation and interpretation for AUC that is closely related to the adjusted Gini coefficient G^* as is shown in (7).

The concentration measure E also offers a simple method to assess the effects of changes in ranks of BEs. If the sum of actual ranks S is reduced by one position (rank), it will cause the following change in E:

$$E(S-1) - E(S) = \frac{1}{n_1(n-n_1)} = \frac{1}{n_1 \cdot n_0} \tag{10}$$

This effect is half of a corresponding effect on G in (3), since E refers to AUC. The effect is independent of the position where the change in ranks is made. Thus, the effect of a change, say, in the first positions is identical with that in the last positions. The result (10) again shows that the effect on G is related to the number of BEs and GEs (all events) in the sample being consequently small in large samples.

3. Numerical Experiments

3.1. Small Sample

This experiment illustrates the calculation and interpretation of the Gini coefficient using a simple numerical example. When the above formulas of G are applied, there are no restrictions for the size of the sample: the sample size (n) and the number of BEs in the sample (n_1) can be any integer. However, to illustrate the concepts in a simple context, a small sample of only 15 events is firstly analyzed. Table 1 shows the data of a sample of 15 events. The events are set in descending order of event risk estimate (core) from position i = 1 to i = 15 on the basis of the estimates given by a predictive model. These 15 units include 5 BEs (n_1), so that the default rate is n_1/n = 5/15. BEs are placed in the rank order of risk given by the model to positions 1, 2, 4, 5 and 8. Since $(n + 1)/2$ = 16/2 = 8, the effect of the BE ranked as i = 8 on the Gini coefficient G is 0. If its ranking were worse than 8, the effect of BE on G would be negative. All BEs with a better ranking than i = 8, make a positive effect on G (increasing concentration).

Table 1 shows that the direct effects, contributions, of GEs on the Gini coefficient G are zero due to the assumption that x_i = 0 for them. The effects of BEs on the Gini coefficient decrease when you move down the ranks. If you move one position lower, the effect decreases according to Equation (5) by $2/(n_1 n)$ = 2/75 = 0.027. When all the effects are summed up, a value of 0.533 for G is got. If it is

Table 1. Experiment with the data of 15 observations.

Rank i	Events	Value of event x_i	Weight of event $(n+1-2i)x_i$	Contribution to Gini coefficients	
				Gini	Adjusted Gini
1	BE	1	14	0.1867	0.2800
2	BE	1	12	0.1600	0.2400
3	GE	0	0		
4	BE	1	8	0.1067	0.1600
5	BE	1	6	0.0800	0.1200
6	GE	0	0		
7	GE	0	0		
8	BE	1	0		
9	GE	0	0		
10	GE	0	0		
11	GE	0	0		
12	GE	0	0		
13	GE	0	0		
14	GE	0	0		
15	GE	0	0		
	Total	5.0000	40.0000	0.5333	0.8000

Legend: BE = Bad event; GE = Good event.

divided by the coefficient (1-default rate) = (1 − 5/15), the adjusted Gini coefficient is got as G^* = 0,800 being equal to Somers' D. Correspondingly, the AUC of the sample is simply $(G^* + 1)/2$ = 0.900 following (7). If you want the coefficient G^* = 0,800 to rise to its maximum value of 1.000, it requires (1.000 − 0.800)/0.040 = 5 positions move up the scale, because the contribution of one position upwards is $2/(n_1(n-n_1))$ = 0.040 on G^* taking account of adjustment for (5). This means that BEs $i = 4$ and $i = 5$ move to $i = 3$ and $i = 4$. In addition, BE $i = 8$ moves to $i = 5$ (so, BEs occupy the first five positions). The effects of BEs clearly illustrate how the Gini coefficients G and G^* are cumulatively constructed in a sample of BEs and GEs.

The concentration measure E can easily be calculated and interpreted using the three sums of ranks presented in (8a), (8b), and (8c). In this case, the actual ranks of BEs are 1, 2, 4, 5, 8 which together make a sum S of 20. The minimum sum L of ranks from five BEs is based on ranks 1, 2, 3, 4, 5 making a sum of 15. In the same way, the maximum sum of ranks (when the BEs have the last positions in ranking) is calculated from 11, 12, 13, 14, 15 leading to the sum H of 65. Thus, the concentration measure E can be calculated as the ratio $1-(S-L)/(H-L)$ resulting in 0.900 (that is equal to AUC).

Table 2 is similar to Table 1 with one exception: BE that was previously in position $i = 8$ has moved up two positions in the rank order, i.e. to position $i = 6$. The predictive model is thus supposed to perform better than the previous results documented in Table 1 indicate. Consequently, concentration should be higher in this case. According to result (5), one position improvement in performance up the rank order means that the increase in G is 0.0267. Thus, two positions up the ranking had increased the Gini coefficient by 2·0.0267 = 0.053. Because it is the only difference in the results, G is now 0.533 + 0.053 = 0.5866. Since the sample in Table 2 has the same default rate (5/15) as in the previous sample, the Gini coefficient can be divided by the coefficient (1 - 5/15), resulting in the adjusted Gini coefficient $G^* = 0.880$ being equal to D. Accordingly, AUC is got as (0.880 + 1)/2 = 0.940. This same result can be obtained through E. In this case, L and H are the same as in the previous case, but S has declined two ranks being now 18. Thus, AUC = $E = 1 - (18 - 15)/(65 - 15) = 0.940$. When the actual ranks are declined by two, the corresponding increase in AUC (and E) is $2/(5(15 - 5)) = 0.040$.

Table 2. Experiment when event (BE) in position $i = 8$ moves to position $i = 6$.

Rank i	Events	Value of event x_i	Weight of event $(n+1-2i)x_i$	Contribution to Gini coefficients	
				Gini	Adjusted Gini
1	BE	1	14	0.1867	0.2800
2	BE	1	12	0.1600	0.2400
3	GE	0	0		
4	BE	1	8	0.1067	0.1600
5	BE	1	6	0.0800	0.1200
6	BE	1	4	0.0533	0.0800
7	GE	0	0		
8	GE	0	0		
9	GE	0	0		
10	GE	0	0		
11	GE	0	0		
12	GE	0	0		
13	GE	0	0		
14	GE	0	0		
15	GE	0	0		
	Total	5.0000	44.0000	0.5867	0.8800

Legend: BE = Bad event; GE = Good event.

3.2. Large Sample

The third case, which is discussed in this context, concerns a large sample with a total of $n = 23.533$ limited companies from Finland. The empirical data were obtained from the ORBIS database maintained by Bureau van Dijk (BvD) (https://www.bvdinfo.com/en-gb/). The selection criteria required that a firm be Finnish, industrial, and incorporated as a limited company, with a status of either active or bankrupt, and with financial statements available for 2022 or 2023. No restrictions were imposed on firm size. Initially, 25.618 firms met these criteria. However, due to missing values in the extracted variables, the number of usable observations was reduced to 23.533 firms, as all firms with incomplete or missing data were excluded from the sample. Of these, 23.386 were active (non-bankrupt) firms, and 147 were bankrupt firms, the latter typically having failed within one year after publishing their last financial statements. The most recent financial statements for the firms in the final dataset were from 2022-2023, with the vast majority originating from 2023.

Thus, the sample includes 23.386 (n_0) acting companies (GE) and only 147 (n_1) bankrupt companies (BE): the sample is therefore heavily imbalanced. In this sample, the default rate is very low, $m = n_1/n = 147/23.533 = 0.0062$. Using this sample of limited companies, a predictive model was developed in which the binary dependent variable (0 = GE and 1 = BE) is explained with key financial ratios calculated for bankrupt companies from the last financial statements. For acting companies, financial ratios are extracted from the same years (2022 or 2023) as for bankrupt firms. The predictive model had three explanatory variables, which are key financial ratios: solvency or equity ratio (solvency), liquidity or quick ratio (liquidity) and return on assets ratio (profitability). The span of prediction is short, mostly less than a year. The hypothesis for the results was that the coefficients of all indicators in the model are negative: the worse the solvency, liquidity, or profitability, the higher is the conditional bankruptcy risk.

The logistic regression analysis (LRA) was used to develop the predictive model. Table 3 shows the results of the estimated logistic regression model. Logistic regression analysis produced the following linear logit for assessing bankruptcy risk: Logit = −3.957 − 0.025·Solvency ratio − 0.483·Liquidity ratio − 0.023·Return on assets ratio. This linear logit can be transformed using the logistic function into the conditional probability of bankruptcy P, where $P = 1/(1 + \exp(-\text{Logit}))$. Using the conditional probability (bankruptcy risk) as a classifier, the companies in the sample were sorted in descending order. Note that the order would be the same, if the logit were used as the classifier: the probability is an increasing function of the logit. Thus, both classifiers would lead to the same rate of concentration. Figure 1 presents the ROC curve for the sample, which intuitively indicates a good quality of the model.

Table 4 shows the first 20 observations (events) from the sorted material. To calculate the Gini coefficient, adjusted Gini coefficient, Somers' D and AUC, only this sorted material with rank numbers is needed. Typically for an ordinal analysis, the probability of bankruptcy itself is not needed in calculations. The (first) 20 events shown in the table include 4 BEs and 16 GEs. Because there are so many

Table 3. Estimated logistic regression model of three financial ratios.

Variable	*b*	Standard error	Wald test	*p*-value
ROA	−0.023	0.004	41.645	<0.001
LRA	−0.483	0.168	8.278	0.004
SRA	−0.025	0.003	90.448	<0.001
Constant	−3.957	0.163	586.312	<0.001

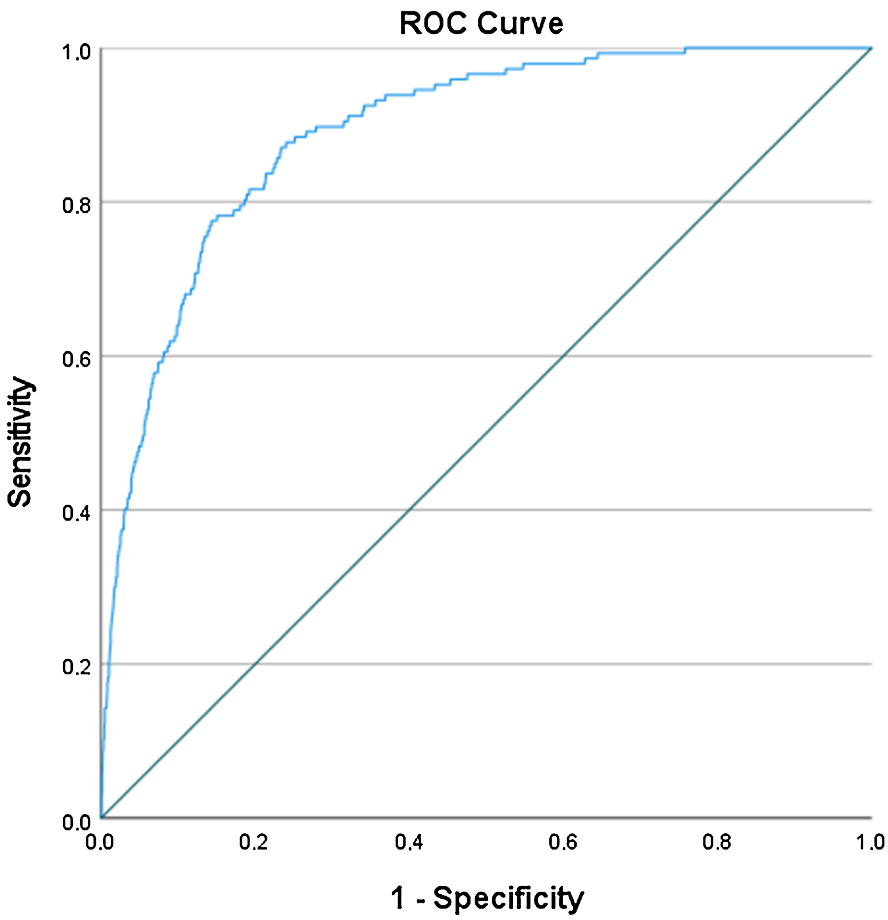

Figure 1. The ROC curve for the three-variable logistic regression model.

Table 4. Calculation of Gini coefficients in a sample of 23.533 active and 147 bankrupt companies.

Rank i	Risk estimate	Events	Value of event x_i	Weight of event $(n+1-2i)x_i$	Contribution to Gini coefficients	
					Gini	Adjusted Gini
1	0.6498	GE	0			
2	0.6034	GE	0			
3	0.5744	GE	0			
4	0.5724	GE	0			

Continued

5	0.5346	BE	1	23,524	0.0068	0.006843
6	0.5100	GE	0			
7	0.4664	GE	0			
8	0.4467	GE	0			
9	0.4464	GE	0			
10	0.4380	GE	0			
11	0.4302	GE	0			
12	0.4152	BE	1	23,510	0.006796	0.006839
13	0.4047	GE	0			
14	0.3959	GE	0			
15	0.3893	GE	0			
16	0.3731	GE	0			
17	0.3705	GE	0			
18	0.3687	BE	1	23,498	0.006793	0.006835
19	0.3619	BE	1	23,496	0.006792	0.006835
20	0.3577	GE	0			
….						
23,533	0.0000	GE	0	0	0	0
		Total	147	2,687,742	0.77695	0.781834

Legend: BE = Bad event; GE = Good event.

observations, the effects of individual BEs on the Gini coefficient are small. Their effect on the coefficient of adjusted Gini is somewhat larger, as these are divided by the coefficient (1 − default rate) = (1 – 0.0062). When all the effects of BEs in the sample are added together, the Gini coefficient G is obtained as 0.7769 and the adjusted coefficient G^* is got as 0.7818. Accordingly, the AUC of the sample = (G^* + 1)/2 = 0.8909 which holds for the ROC curve in **Figure 1**.

Figure 2 shows the effects of events (multiplied by 100) on the adjusted Gini coefficient G^*. The effects of GEs are all zero and are thus locating on the horizontal axis. However, the effects of BEs decrease linearly and the effect is zero in the middle of the rank order for i = (23.533 + 1)/2 = 11.767. Note that after this middle point, there are only 5 BEs among the rest of the events. Most of the BEs (120 or 86.63%) are ranked among the first 5.000 events, which indicates a strong concentration. Consequently, the adjusted Gini coefficient G^* = 0.7818 indicates a very good predictive model. This is also shown by AUC = 0.8909, which is almost at the level of excellent (0.900). AUC can easily be calculated also through E. In this case, S = 385,878, L = 10,878, and H = 3,437,742, which give 0.8909 as E and AUC. Thus, this example shows that the predictive model works well and bad events

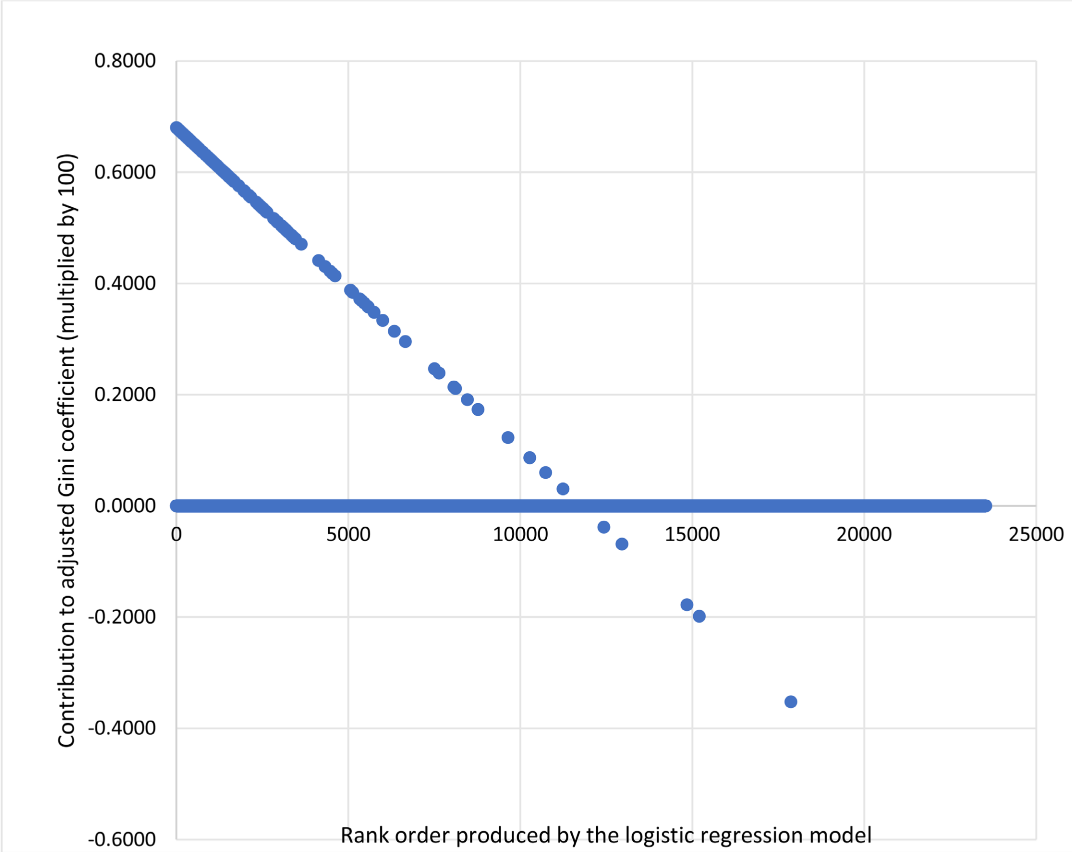

Figure 2. Contributions of BEs on the adjusted Gini coefficient (multiplied by 100).

(BEs) are strongly concentrated at the top of the rank order.

However, there is a feature of the Gini coefficient G, which weakens its usability. Let us assume that BEs move forward in the ranking say by a total of 10.000 positions. This means that each BE (n_1 = 147) moves on average 68.0303 positions forward, which on the basis of (5) makes the Gini coefficient G increase by only 0.00578 leading to G = 0.78273 instead of 0.77695. This means that in large and imbalanced samples, the Gini coefficient is not sensitive to the differences produced by the models in the rankings of events. The situation would be different if in this case the sample size for example were n = 1.000 (instead of 23.533). Then, the effect on the Gini coefficient G would be 0.13605 with an improvement of 10.000 positions, which is already intuitively very significant. This leads to G = 0.91300 instead of 0.77695. Since the sample is imbalanced and the number of BEs is relatively small, their positions should improve very significantly to make the effect on the Gini coefficient G intuitively relevant. Of course, this same feature also deals with AUC. The effect of a drop in 10.000 ranks means that AUC will rise only by 10.000/(147 (23,533 − 147)) = 0.0029. In this case, a multivariate model of three financial ratios is used as the predictive model. If we use only a univariate model of the solvency ratio, AUC of 0.8750 will resulted. This means that AUC is decreased by 0.0159 (from 0.8909) referring to an increase of 54717.75 in actual ranks S.

4. Summary of the Study

In order to measure the performance of predictive models, a large number of different metrics have been developed, presented and analyzed in the scientific literature. This paper focused on analyzing mainly three measures, all of which are traditional measures of concentration: (Corrado) Gini coefficient, adjusted Gini coefficient, and AUC. The more bad events (BEs) are concentrated on the front end of the ranking produced by the predictive model, the better the performance of the model is considered by these metrics. Since the performance of the models in this case is based only on the rank order of the events, the metrics are ordinal non-parametric indicators. Here, in a way, is their goodness and weakness. These measures only need the order of the events, and the magnitude of the risk, for example, is irrelevant. This turns into a weakness of the metrics if the differences in the predicted risk vary strongly between events: regardless of the difference in the predicted risks, only the rank order is paid attention to.

The objective of this paper was to analyse a formulation of the Gini coefficient, and to present numeric examples of measuring concentration and to analyze how strongly the differences in the rank order of events affect the measures of concentration. Since all evaluated metrics belong to the family of ordinal metrics, they produce very similar results that are already known in the literature. The adjusted Gini coefficient G^* can be calculated from the (Gorrado) Gini G with a simple transformation. In the same way, AUC is obtained as a linear transformation of the adjusted Gini coefficient. It can also be shown that the adjusted Gini coefficient is equal to the fourth ordinal measure of concentration, Somers' D. Since the measures are simple transformations of each other, they tend to have the same advantages and disadvantages.

All these measures have the property that moving between the positions of the rank order contributes to the concentration measure to the same (absolute) extent. The contribution of events to the metric decreases relatively quickly along with the ranking, because the weight of the positions in these ranking declines strongly. Despite this, a similar move between the positions of the ranking always causes a change in the metrics of the same size, regardless of where in the ranking the move takes place: if, for example, an event second in the ranking moves to first, the change in metrics is the same as if the event last in the ranking moved one position from last to last. In large samples, the absolute effect of single moves is very small, so the measures do not effectively detect changes or differences in orders. The features of the meters are completely based on the fact that they pay attention only to the ranks of the events. This dependence was illustrated in the paper by developing the index E, which is explicitly based on three sums of ranks for BEs: acting sum, lowest possible sum and highest possible sum of ranks. The metric E gives the same result as AUC but is easier to interpret and to calculate. Thus, it may provide a useful tool for analyzers of concentration. This study indicates that it is crucial to focus on the effects of ranks, when assessing the classification performance of models.

Finally, it is important to refer to the problems in highly imbalanced (skewed) samples found in this study. In these kinds of samples, ROC, AUC, and the Gini coefficient may give an overly optimistic view of the performance of the model. These measures are largely insensitive to the proportion of BEs and GEs, as they give equal weight to both classes. Consequently, a model that performs well on the abundant GE class can still achieve a high AUC or Gini coefficient, even its ability to identify true BEs is limited. In these kinds of situations, this limitation can be addressed by imbalance-tolerant metrics such as the Precision-Recall Area Under the Curve (PR-AUC). This approach focuses specifically on the performance with respect to the BE class (Boyd, Eng, & Page, 2013). Precision measures the proportion of correctly identified BEs among all instances predicted as positive, while recall (or sensitivity) measures the proportion of actual BEs correctly identified. The PR-AUC summarizes the trade-off between these two measures across different classification thresholds. PR-AUC decreases sharply when false BEs increase, reflecting the real-world cost of incorrect BE predictions.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Adeodato, P., & Melo, S. (2022). A Geometric Proof of the Equivalence between AUC_ROC and Gini Index Area Metrics for Binary Classifier Performance Assessment. In *2022 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-6). IEEE. https://doi.org/10.1109/ijcnn55064.2022.9892048

Bowles, S., & Carlin, W. (2020). Inequality as Experienced Difference: A Reformulation of the Gini Coefficient. *Economics Letters, 186,* Article 108789. https://doi.org/10.1016/j.econlet.2019.108789

Boyd, K., Eng, K. H., & Page, C. D. (2013). Area under the Precision-Recall Curve: Point Estimates and Confidence Intervals. In: H. Blockeel, K. Kersting, S. Nijssen, & F. Železný (Eds.), *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2013.* Springer.

Cicchitelli, G. (1976). The Sampling Distribution of the Ratio of Two Linear Combinations of Order Statistics. *Metron, 34,* 219-228.

Crook, J. N., Edelman, D. B., & Thomas, L. C. (2007). Recent Developments in Consumer Credit Risk Assessment. *European Journal of Operational Research, 183,* 1447-1465. https://doi.org/10.1016/j.ejor.2006.09.100

Gini, C. (1912). *Variabilita e mutabilita: Contributo allo studio delle distribuzioni e relazioni statistiche* (Vol. III (Part II)). Cuppini.

Giorgi, G. M., & Gigliarano, C. (2017). The Gini Concentration Index: A Review of The Inference Literature. *Journal of Economic Surveys, 31,* 1130-1148. https://doi.org/10.1111/joes.12185

Joshi, R., Gupte, R., & Saravanan, P. (2018). A Random Forest Approach for Predicting Online Buying Behavior of Indian Customers. *Theoretical Economics Letters, 8,* 448-475. https://doi.org/10.4236/tel.2018.83032

Kohl, M. (2012). Performance Measures in Binary Classification. *International Journal of*

Statistics in Medical Research, 1, 79-81. https://doi.org/10.6000/1929-6029.2012.01.01.08

Kumar, P. R., & Ravi, V. (2007). Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques—A Review. *European Journal of Operational Research, 180,* 1-28. https://doi.org/10.1016/j.ejor.2006.08.043

Langohr, H., & Langohr, P. (2009). *The Rating Agencies and Their Credit Ratings: What They Are, How They Work, and Why They Are Relevant.* John Wiley & Sons, Incorporated.

Lessmann, S., Baesens, B., Seow, H., & Thomas, L. C. (2015). Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research. *European Journal of Operational Research, 247,* 124-136. https://doi.org/10.1016/j.ejor.2015.05.030

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A Misleading Measure of the Performance of Predictive Distribution Models. *Global Ecology and Biogeography, 17,* 145-151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Moon, H., Pu, Y., & Ceglia, C. (2019). A Predictive Modeling for Detecting Fraudulent Automobile Insurance Claims. *Theoretical Economics Letters, 9,* 1886-1900. https://doi.org/10.4236/tel.2019.96120

Nahm, F. S. (2022). Receiver Operating Characteristic Curve: Overview and Practical Use for Clinicians. *Korean Journal of Anesthesiology, 75,* 25-36. https://doi.org/10.4097/kja.21209

Newson, R. (2006). Confidence Intervals for Rank Statistics: Somers' D and Extensions. *The Stata Journal: Promoting Communications on Statistics and Stata, 6,* 309-334. https://doi.org/10.1177/1536867x0600600302

Schechtman, E., & Schechtman, G. (2019). The Relationship between Gini Terminology and the ROC Curve. *Metron, 77,* 171-178. https://doi.org/10.1007/s40300-019-00160-7

Sen, A. (1973). *On Economic Inequality.* Clarendon Press.