A part of INTRASOFT INTERNATIONAL

2018

IFRS 9 – Report-31/12/2018

IFRS9 COMPLIANCE REPORT FOR THE PERIOD ENDING 31/12/20

Subject:	IFRS9 – 31/12/2018
Version:	Version number and Date
Status:	Confidential

Document Revision		
Version	Description	Date
1st	IFRS9 Report Calculation	02/02/2019
2nd	IFRS9 – Comments by Client	15/03/2019
FINAL	FINAL REPORT	31/03/2019

Contents

1. Introduction

1.1. Overview

In this document we discuss WEMETRIX®'s methodology for developing Credit Scoring Models for evaluating creditworthiness of counterparties of corporations. The coverage of the relevant evaluations extends from large listed companies to entrepreneurs. Although the methodology is unique, the selection of the models' criteria and the granularity of the results are mainly driven by the type of the counterparties and the availability of information about them.

The development and the maintenance of these models as well as the selection of the criteria used is an ongoing process and, therefore, they require a robust organization structure and standardization of the applied practices and techniques. For this purpose, **WEMETRIX®** employs a powerful modern software platform, the Models Factory.

The key factors for achieving the high-quality standards set by **WEMETRIX®** and supervisory authorities for the relevant evaluations are the:

- More than 15 years of experience in successfully collecting, cleansing and processing the required financial and qualitative data.
- Powerful modern techniques and practices of developing and validating the models and robust scoring methodology.
- High quality controls adopted by **WEMETRIX®** for ensuring that all procedures, activities and methodologies employed follow the most demanding modern quality standards.

The main characteristics of the process of developing our Credit Scoring Models are the following:

- The aim of a model is the more accurate estimation of the counterparties' probability of defaulting against its credit obligations.
- Before attempting to develop a model, the available raw data are comprehensively analysed both statistically and logically so that the appropriate criteria and their relationship to a possible default to be revealed.
- The development uses statistical as well as empirical techniques for both determining the most appropriate set of independent variables (criteria) and for calculating their weights (regressions).
- Multiple (several hundred of) different models are usually developed as candidates before the final choice is performed.
- The selection process of the final model between the candidates employs statistical as well as empirical methodologies and follows the modern standards.
- After the development, follows a comprehensive validation and calibration analysis in compliance with the requirements.

A detailed documentation accompanies each model development describing in detail (using graphs where necessary) the properties of data used, their predictive ability, the exact modelling, validation and calibration methodologies, the results of regressions and the metrics of the model's accuracy and power.

1.2. The New Standard, IFRS 9

Its implementation is mandatory as of January 1st 2018, and it introduces some major changes.

- new requirements for classification and measurement, impairment, and hedge accounting.
- requires provisions for credit losses both from performing and non-performing financial instruments. Even those segments that have never produced a delay in payments or a write off, require a loss provision. This means that even "safe customers" -as considered by your corporation- expose your assets and cash inflows to credit risk -the likelihood to incur credit losses in the future (Day One Loss).
- Regarding the Impairment, IFRS 9 requires at each reporting date, the reporting entity to measure the loss allowance (bad debt provision) for a financial instrument at an amount equal to the lifetime expected credit losses if the credit risk on that financial instrument has increased significantly since initial recognition.
- It requires a forward-looking approach, by considering macro-economic forecasts
- Needs to test the portfolio of financial instruments and measure the condition of their credit quality / condition.

While the Simplified Approach requires less calculations and data, the General Approach (calculating PDs and LGDs) can lead to more accurate results, and in certain cases to smaller P&L impact. Furthermore, the Simplified Approach cannot be accurate enough in cases of low / zero default portfolios, not to mention that in times of changing economic conditions, robust statistical methods are required to select the optimum historic reference period.

2. Data Analysis

2.1. General Comments on Data

Adequacy, Representativeness and Quality of data used for developing a statistical model are main keys of its performance. **WEMETRIX®** has a long successful history in collecting, cleansing and analyzing quantitative and qualitative data. Therefore, the processes and the database structures in place are mature, tested and accurate. In addition, the effort made for fully understanding the nature of the incoming information combined with an automated set of multiple filters and procedures ensures that the quality of the data is the higher possible and potentially bad data are appropriately isolated.

Even though the available information that is used for developing and operating Credit Scoring Models may differ significantly by model, it could be classified in the following categories:

- Financial data, which are publicly available in the form of annual or quarterly statements. Usually they are used for composing financial ratios which participate in the models as continuous independent variables.
- Non-financial quantitative data, which include the operation years of a business, the number of employees, the number of branches etc. They often combined with financial data in the form of ratios (e.g. number of employees over annual sales).
- Internal behavioural data gathered from the transactional and/or accounting information stored in the IT systems of the company that uses a model. Such data can be potential delays in paying, fluctuations in turnover and payments, usage of limits etc.
- Objective demographic and qualitative data, which include information on the industry sector(s), the premises location, the operations areas etc.
- Derogatory information, which is collected from external sources.
- Standardized qualitative information (usually in the form of check box or multiple choice) reflecting knowledge of relationship managers for their clients. Such type of information may be potential recent operational disorders that are not yet reflected in the financials, natural disasters, massive layoffs, petitions of bankruptcy etc.

In addition to the above sources, in some cases, macroeconomic indicators that reflect one or more economic environments are included as criteria in the models. Macroeconomic variables are usually used if Point-In-Time (PPT) measures are needed as direct output from the model or if the targeted population is split in several different macroeconomic sectors (such as countries) with low or negative default correlation between them. In other cases, the external environment is usually considered explicitly to the model.

2.2. Data Gathered

2.2.1. What the client needs to collect

We send a data template to our clients for them to fill. The template has 15 fields of which 7 are mandatory, 4 are “if available” and 4 are indicative.

Static data are data that do not change such a client VAT, portfolio ID etc. The following table illustrates the static data that were provided by Company XYZ:

Table 1 Sample Data

There are FOUR Data Rules

Sample Data

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Snapshot Date	Customer ID	Sales	Cash Payments	Other Credit	Balance	Past due 0-30	Past due 31-60	Past due 61-90	Past due 90+	Credit Limit	Customer's Rating	In Default	Industry	Portfolio
2018-03-31	12345	21530.50	14500	2300	12000	0	2000	0	2000	15000	3	0	Chemicals	Wholesale

1. Snapshot Date is the only field formatted as **Date**

2. The decimal separator is the **decimal point**

3. There is **NO** thousands separator. So, one thousand five hundred is written as 1500 and NOT as 1,500 or 1.500

4. All fields other than Snapshot Date must be formatted as **Text**

Legend

Mandatory
If available
Indicative

The table below explains in more detail what each of the field represents.

Table 2 Description of date fields

1. Snapshot Date	The date of the snapshot. Refers to END OF MONTH, and we need to have 24 months of data.
2. Customer ID	The customer id. You may use any kind of unique ID, this is the same key that we will be using in our report, so you can identify your clients.
3. Sales	Total invoices for the month, including VAT. The figure refers to **Monthly** and not cumulative/year to date sales.
4. Cash Payments	The total CASH/Bank Deposit payments reeived by the customer in the month, excluding received post-dated checks, if any.
5. Other Credit	Any other credit entries (could be post dated checks, discounts, write-offs, etc)
6. Balance	The total unpaid amount (both within the credit terms and overdue amounts), at the end of each snapshot period.
7. Past due 0-30	The value of **overdue** receivables, between 0-30 days.
8. Past due 31-60	The value of **overdue** receivables, between 31-60 days.
9. Past due 61-90	The value of **overdue** receivables, between 61-90 days.
10. Past due 90+	The value of **overdue** receivables exceeding 90 days.
11. Credit Limit	The Credit Limit granted to the client. Ideally, this amount is available for the entire 24 month period, in case that any changes have been made. This should be a monetary value, but in case you have payment terms (in days), rather than a Euro Credit Limit, you may of course use it.
12. Customer's Rating	You may provide either an internal rate (credit evaluation of your client, if possible referring to each single snapshot period), or an external rate that you obtain from Credit Bureaus. In any case, please use the comment / messaging feature of RISK 9 to provide as with the Rating Scale.
13. In Default	Usually this is a YES/NO, or TRUE/FALSE, or 0/1, or a similar description to show if a client is in Default or Not. By the term "Default", we usually refer to clients that have problems paying the invoiced amounts, and usually there are restrictions in doing new business with them. In case there are more than two categories (for example, Performing, Non-Performing, Cash Only, CEO Approval etc), please fill in accordingly. In any case, please use the comment / messaging feature of RISK 9 to provide as with any details.
14. Industry	Refers to your client's line of business, his industry or sector description. Can be any type of industry classification, but please use the comments area to explain the taxonomy used.
15. Portfolio	In case you have any **primary** groupings of your clients (for example, by product, business unit, service line, geography), please provide the relative information. In case of multi-dimensional groupings, preferably use the most important grouping, otherwise, you may add columns.

This is INDICATIVE. In these fields please provide values of overdue amounts, that is, amounts not paid on the the agreed payment date. You may provide a different breakdown (for example 1-90 days, 90+days or more detailed). In case you do not have this analysis available, you may give as the typical Ageing Analysis, ideally with agreed payment terms in days.

2.2.2. Example of Data received from client XYZ

We received from the financial department of company XYZ monthly data for the years 2016, 2017 and 2018 (January 2016 to June 2018), below one can see a monthly snapshot of the data.

Table 3 Graphical display of the data received

Date	CustomerID	Payments with VAT	Credit with VAT	Secession - Transition Amount	Sales with VAT	VAT of Sales
31/01/2016	1000	156.273.549,46	132.852.000,00		77.061.549,46	14.409.883,23
29/02/2016	1000	206.968.910,87	136.417.210,12		168.109.520,73	25.233.089,07
31/03/2016	1000	185.436.753,00	229.818.706,58		170.981.590,49	24.922.830,37
30/04/2016	1000	176.433.383,96	87.732.752,03		160.048.269,21	24.464.947,33
31/05/2016	1000	173.498.849,24	273.686.591,22		161.106.372,40	24.977.696,96
30/06/2016	1000	173.907.258,01	115.150.362,55		173.716.699,40	26.958.620,96
31/07/2016	1000	192.597.325,04	146.861.203,78		187.694.985,21	28.341.764,73
31/08/2016	1000	307.547.418,81	169.428.066,75		202.128.284,38	29.777.911,01
30/09/2016	1000	294.816.978,12	178.678.921,93		205.486.783,22	30.197.898,29
31/10/2016	1000	188.149.077,79	86.047.532,87		188.139.103,63	28.086.184,03
30/11/2016	1000	198.527.772,23	102.158.383,62		198.518.182,78	31.632.082,45
31/12/2016	1000	267.843.578,89	154.808.749,83		267.821.016,91	39.100.718,14
31/01/2017	1000	7.181.782,90	113.921.943,57		181.782,90	35.183,78
28/02/2017	1000	211.266.621,53	58.524.109,76		187.587.280,72	28.685.699,38
31/03/2017	1000	208.879.606,76	335.609.879,47		170.570.700,05	26.036.102,49
30/04/2017	1000	203.865.269,84	126.726.614,25		177.044.038,97	27.150.902,08
31/05/2017	1000	164.404.888,08	334.535.849,35		164.278.687,01	24.831.383,32
30/06/2017	1000	189.031.687,21	697.719.615,73		188.841.735,52	28.642.673,77
31/07/2017	1000	196.147.067,27	185.618.470,12		196.141.766,89	28.520.684,25
31/08/2017	1000	218.988.187,37	352.552.405,79		218.732.100,46	31.277.241,02
30/09/2017	1000	215.212.875,32	190.522.325,40		215.212.169,89	31.094.555,70
31/10/2017	1000	206.425.039,91	71.542.905,72		202.157.537,46	30.214.346,42
30/11/2017	1000	167.316.183,23	318.723.901,36		167.313.587,27	25.184.025,84
31/12/2017	1000	178.493.987,39	54.599.527,90		170.873.020,59	27.069.126,93
31/01/2018	1000	199.221.646,54	251.802.823,66		199.221.646,54	28.636.535,43
28/02/2018	1000	221.241.167,61	48.677.709,48		221.230.283,90	32.015.117,69
31/03/2018	1000	187.950.640,58	153.619.079,97		187.840.737,66	26.911.205,90
30/04/2018	1000	771.252.723,90	906.976.180,92	507.022.702,97	127.662.399,43	20.812.278,25
31/05/2018	1000	182.740.091,91	55.120.113,08		182.565.559,62	26.260.659,22
30/06/2018	1000	249.957.344,69	143.014.712,15		152.287.567,33	25.073.961,51
31/01/2016	2100	66.770,72	33.676,34		33.385,36	6.242,79
29/02/2016	2100	180.579,81	180.579,81		90.289,91	16.883,48
31/03/2016	2100	1.522,44	66.434,96		761,22	142,34

2.2.3. What **WEMETRIX®** has to do with the Data

One of the first thing we notice is that the portfolio we are considering is a low default portfolio. Therefore, we define the breach as follows:

- The payment is overdue for more than 90 days.
- And this delayed payment accounts for at least 10% of the total monthly balance.

We want to clarify that these definitions are up for discussion and stem from our experience and international practice, nevertheless should it be deemed necessary alternative default definitions can be easily used.

With this definition of default, we get the following statistics

- Total number of observations: 261,151
- Total number of unique customers: 141,829
- Total number of observations with defaults: 13,326
- Total number of non-defaulted observations: 247,825
- Average default rate (in percentage): 5.10%

From the above information is evident that a large percentage of customers have very few observations. It was decided, therefore, to exclude all customers with less than 12 observations (i.e. one year) and all customers which have a negative balance. After we have applied the above "filtering" to our initial data we are left with the following information:

- Total number of observations: 17.090
- Total number of unique customers: 766

2.2.4. Economic and Financial Data

We have enriched our models with macroeconomic and financial indices. To do this we have gathered economic data from January 2016 to June 2018. The first index is the spread of the 10-year Greek bond in relation to the corresponding German bond. From this index we calculate the monthly percentage change using the following formula $\left(\frac{Spread_t}{Spread_{t-1}} - 1\right)$.

The second index we use is the general index of the Greek Stock Exchange (ASE). From this index we calculate the monthly percentage change using the following formula $\left(\frac{ASE_t}{ASE_{t-1}} - 1\right)$.

We want to stress here that model can easily incorporate any kind of macroeconomic of financial series we may like. In this example we chose the spread and the return of the Greek market since they were proven to have high explanatory power. Nevertheless, as the Greek economy changes with time so it might be the case that in the future other indices may have higher explanatory power.

3. Definitions

The definitions of the main concepts used in this document (and other relevant **WEMETRIX®**'s documents) are described below.

3.1. Default

Default is defined as the inability of a counterparty to meet their contractual obligations. In case of an obligor or a customer it is their inability to promptly pay against a commitment deriving from credit granted to them.

Apart from the banks which must define default in compliance with Basel II/III requirements, the exact events that evidence default may vary by case. Some of the rules that are usually used for determining an appropriate default definition are the following:

- The default is an irreversible event. In other words, as potential triggers of default we seek for events that lead to permanent cease of Credit Lines. A small number of potential exceptions to the rule are acceptable.
- Multiple triggers of default can be combined.
- A past-due of a material amount for several days is a usual default trigger. Both the materiality of the amount and the past-due period depend on the credit rules set by the creditor and thus, are examined for each case independently.

3.2. Time Horizon

The time horizon of a model is the period that the credit scores refer to. **WEMETRIX®** usually set time horizon to one year, i.e. the credit scores represent the Probability of Default (PD) of the assessed counterparty in the 12 months period following the assignment date.

However, it is worth mentioning that often in practice shorter or longer periods should be used by several applications. In such cases, even if the horizon used for the models' development remains annual, Default Probabilities are separately calculated for each grade under the required horizon.

Generally, speaking one can identify three stages, stage 1 (initial recognition), stage 2 (significant increase in credit risk) and stage 3 (asset becomes credit impaired). For most of the cases the assets will fall at stage 1, but if there is an increase in the credit risk or the asset becomes credit impaired than the assets needs to be moved in stage 2 or 3. For stage 2 and stage 3 one needs to calculate a life time expected credit loss, for stage 1 one needs only to calculate twelve month expected credit losses and hence a 12 month PD.

3.3. Model's Observation Period

Because drivers for predicting the future are facts known at the time of assessment they refer to the present and the past. Therefore, a model's criteria should be sought in a window in the past. The time interval that is used for collecting information for estimating the values of a model's criteria is the observation period of the model. It is determined according to the type of the model and the statistically measured significance of the data describing facts in the past. Some rules of thumb followed by **WEMETRIX®** are the following:

- For financial information, the financial statements of the two most recent fiscal years, if available, are usually collected and taken into account. If more recent information is available (e.g. quarterly statements) then it can be considered as well.
- For externally collected derogatory data, there is strong evidence that the data of the three years prior to the assessment date are actually exploitable as default predictors.
- The observation period for internally collected behavioral data, that is data related to past defaults, frequency of past-dues, limits coverage etc., is estimated statistically. Usually, the optimum period is between 3 and 12 months and most often 6 months.
- For non-financial quantitative and objective qualitative data, the last available most updated information as of the time of the assessment date is used.

4. Data Transformation before modelling

4.1. Univariate Analysis

WEMETRIX® uses the total sample (development and test) for the univariate analysis. The step consists of selecting the candidates for the model's independent variables. As a rule of thumb, variables with high missing values frequency are excluded. In addition, the quantitative data are often combined in the form of ratios to be better exploitable as model inputs. The final selection of the independent variables is the most important part of modelling Credit Risk. **WEMETRIX®** honours that rule by particularly analysing all possible candidates for their default predictive power.

Since the models' inputs may be both qualitative and quantitative data, the following types of independent variables are supported:

- Continuous variables (ratios, years etc.),
- Discrete variables, which usually correspond to qualitative information and
- Binary variables (true, false).

The tools used in univariate analysis consist of the following:

- Graph of the relationship between the values of a variable and the corresponding observed default frequencies.
- Graph of the observed default frequencies per quantile of the population (sample) sorted by the variable's values.
- Calculation of the Accuracy Ratio (AR or Gini coefficient) of the univariate model after appropriately treating statistical noise and non-linearity.
- In cases the above tools reveal any anomaly in the behavior of a variable, additional tools are employed (ROC curves, information entropy analysis etc.).

Relying on the above tools **WEMETRIX®** examines several factors to determine which variables should be candidates. Those factors, among others, include:

- The monotonicity in the relationship between the variable and the default frequencies. The experience of the analyst is necessary for determining if a potential anomaly is due to statistical noise or data deficit or if it simply demonstrates the normal behavior of the variable. In the latter case, the variable may be used as a candidate only if its non-monotonic behavior is fully explainable and accepted (e.g. sales growth).
- The portfolio areas in which the variable is more powerful. Typically, these areas are characterized by a steep slope of the curve of the default frequencies per quantile. This information could be also useful for deciding if it is safe for two high correlated independent variables to co-exist in a model. More specifically, the more distant are their areas of high discriminatory power, the less probable is for the model to overfit because of their coexistence.
- The total discriminatory power of the variable, which is determined by the value of Accuracy Ratio.

As the result of the above analyses, the variables with a fully explainable behaviour and the highest default predictive power are chosen as candidates for participating in the model.

4.1.1. Treating noise and non-monotonicity

Since Default Probability depends on multiple other factors apart from the currently examined variable, a high amount of statistical noise is unavoidably involved in the problem. Therefore, the exact shape of the relationship between the variable and the default probabilities cannot be uniquely and clearly derived from the graphs of univariate analysis. Relying on the strict statistical relationship shapes for modelling Default Risk could result in an overfitted model, which exhibits high in-sample accuracy ratios, but it is actually weak in real world operation.

To avoid overfitting and to obtain the best approximation of the actual relationship between a variable and Default Risk, **WEMETRIX®** analyses the variables as follows:

- The monotonicity of the relationship and its direction is determined, based on the empirical expectations for each variable. In some cases, U (smile) shaped variables may be acceptable, if their shape is explainable and consistent with other similar studies.
- The continuous variables are discretized by using an advanced methodology of optimal binning, which employs identifying and repairing concavities (applying convex hull) in the variable's ROC curve. For an independent variable i, the binning process results in a number n_i of bins. The bin $j \in \{1,2,\dots,n_i\}$ maps a continuous area of possible values $(x_{i.j.b}, x_{i,j,e}]$ of the variable to the mean default frequency $p_{i,j}$ of the observations belonging to this area. The n_i bins fully cover all the possible values of the variable i. In other words, $x_{i,j,e} = x_{i,j+1,b}$ for each bin $j \in \{1,2,\dots,n_i - 1\}$ and $x_{i.1.b} = -\infty$ and $x_{i.n_i.e} = \infty$.

  This technique isolates the majority of statistical noise while ensuring monotonicity and improving the AUROC accuracy measure.
- For the discrete variables, if necessary, monotonicity is achieved by unifying subsequent values into bins.

4.1.2. Transforming Independent Variables

The model methodology adopted by **WEMETRIX®** assumes that the Default Probability is almost linear in each explanatory variable. However, as mentioned above, that is not the case in the majority of the variables. Some of them may be non-numerical (discrete variables) while even for the numerical ones linear (or close to linear) relationship with default is rather rare. In addition, some meaningless extreme values may appear here and there, thus complicating the problem.

WEMETRIX® uses a non-parametric technique of transforming the independent variables for treating the above issues. The transformations rely on the discretization process (optimal binning) described above. More precisely, the transformed values directly derive from the mean default frequencies of the bins. The main characteristics of the transformed values are the following:

They ensure linearity with default frequencies (by definition).

They are highly normalized, i.e. their values areas are the same (0% to 100%).

Potential extreme values are automatically treated without requiring any assumptions.

Since the mean default frequencies of the bins are in fact default rates, the appropriate way of entering them in a logistic model is through the logit function:

$$T(x_i) = \Lambda^{-1}(p_{i,j}) = ln\left(\frac{p_{i,j}}{1 - p_{i,j}}\right)$$

where $p_{i,j}$ is the mean observed default frequency in the bin j of the variable i and $T(x_i)$ the transformed value that should be entered in the model.

4.2. Correlation matrix

Correlations between independent variables are unavoidable. However, in some cases, too high correlations may significantly worsen out-of-sample performance of a model by introducing overfitting. In general, correlations between variables are not linear and, therefore, linear measures (e.g. Pearson's statistic) are inappropriate. Moreover, high correlations between two variables may be located in specific areas of their possible values and this fact, although can be significant, may remain hidden if the correlation is measured as linear. As the transformed variables are already linearized, **WEMETRIX®** measures the correlations of the transformed variables, instead of the initial ones, to deal with non-linearity.

The correlations between the transformed values of all the candidate variables between them are collected in a relevant matrix. Potential correlations that go against common sense or analysts' expectations may hide raw data deficiencies and are further explored until they are fully explained. Otherwise, the problematic variable is rejected as a model's criterion candidate.

It is worth noting that a high correlation between two independent variables does not lead per se to excluding one of them from the model. It is however a strong sign for the developer to be extremely careful in the case two highly correlated criteria coexist. Typically, only one of the two candidates (usually that with the highest AR) joins initially the model, while the other is tested last, after the inclusion of all the rest selected variables. Its participation or not depends mainly on (a) the right signs of the corresponding coefficients, (b) the comparability of their absolute values both between them and with the other coefficients and (c) the improvement of the model's performance in the out-of-sample tests by including both into the model.

5. Modelling

5.1. PD Modelling

The mathematical form of **WEMETRIX®**'s statistical default models is the "linear logistic function". More specifically, that form generates the (theoretical) default probability from the following formula:

$$P(Y = 1|X) = F\left(w_0 + w_i \cdot \sum_{i=1}^{n} T_i(x_i)\right)$$

where

$$F(u) = \frac{1}{1+e^u}$$

and X the vector of the *n* values of the independent variables $(x_1, x_2, \dots, x_n)$, $T(X)$ the vector of the corresponding transformed values and w_i the model's coefficients.

The methodology that **WEMETRIX®** applies for determining the model's coefficients is the logistic regression by maximizing the likelihood function (Maximum Likelihood Estimation or MLE). Under this approach the model's coefficients w_i are determined so that the value of the following function is maximized:

$$\ln L = \sum_{i=1}^{n} \{ y_i \cdot \ln\big(p(x_i)\big) + (1-y_i) \cdot \ln\big(1-p(x_i)\big) \}$$

where *n* is the number of the observations used for training (in-sample), y_i is the dependent variable and has 1 or 0 as its values (defaulter or non-defaulter) and $p(x_i)$ is the Default Probability.

Each regression, apart from the coefficients, returns the following set of statistics:

- The Pseudo-R^2 of the model, which can be used for comparing regression attempts (its absolute value is of limited importance as the dependent variable is binomially distributed),
- The p-value of the LR test of the model and
- The standard errors SE, the t ratios and the p-values of coefficients w_i.

5.1.1. Independent Variable Selection

The method that **WEMETRIX®** employs for optimizing the independent variables selection is the "forward selection" process. Under that approach, the model is initially built with variables exhibiting the highest relationship to default and then those with lower relationship are entered until no additional significance is achieved. The above process is extremely flexible but hardly controllable without relevant experience. **WEMETRIX®**, in order to facilitate and to standardize the process as much as possible, instituted a bundle of guidelines:

- Multiple attempts should be performed, each starting with a different set of initial variables.
- Steep monotone variables with high AR are considered the best candidates and should be introduced to the model first.
- Adding regressors usually increases fit, but also increases the variance of the predicted variable. Therefore, one should be very careful in incorporating too many variables into the model. As a rule of thumb, more than 8-9 variables increase the probability of overfitting.
- The participation of variables highly correlated to each other should be avoided. However, if too few variables are available and if it is obvious that, despite their strong relationship, part of the information carried by the two variables is different and complimentary, then their coexistence in the model could be attempted. In this case, as already mentioned, the more powerful member should be added first, while the other is added very last, regardless its univariate correlation to default.

- After each regression attempt the signs and the values of coefficients are checked for their rationale. Potential deviations are more probably evidence of overfitting.
- Also, the p-values of the coefficients are checked. However, a potential moderately high p-value does not necessarily lead to rejecting the model. High p-values may be observed if one and unique value of the corresponding independent variable (criterion) appears in the bulk of the sample. Nonetheless, if strong empirical evidence exists that the cases of a different value decisively influence creditworthiness, the model could be accepted. A typical, although extreme, example of such a criterion is the flag of "bankruptcy petition", which, while rare, is crucial for an assessment. In similar situations the key factor of accepting or rejecting a model is analyst's experience.
- After adding a new variable, the attempted model should be tested both in-sample and out-of-sample. Potential significantly improved in-sample performance that is not followed by proportionate out-of-sample results (or, even worst, performance deterioration in the out-of-sample tests) is a strong evidence of overfitting. In such cases, the last added variable should normally be removed from the model.

It is noted that the above guidelines are incorporated into the Model Factory so that they warn the analyst for their violation so that they cannot be ignored or overruled by accident.

5.1.2. Multivariate Analysis

The objective of the multivariate analysis is the development of several models (candidates) and it takes place after analysing each independent variable for its default predictive power. The issues involved and the steps of performing the multivariate analysis are discussed in the following sections.

5.1.3. Validation

After developing a model, a full validation procedure takes place, based solely on the test part of the sample. The concept pf validation is mixture of art and science. Nevertheless, broadly speaking the following benchmark values apply when validating a model.

- The maximum population concentration in a Score is set to 25%.
- For the Accuracy Ratio, the minimum acceptable value is set to 40%.
- For the K-S statistic, the minimum acceptable value is set to 28%.

For the calibration check, there should not be a deviation from the respective ranges in more than two categories. In addition, a deviation in the same category and to the same direction should not be observed for two subsequent years.

However, it should be mentioned that potential violations of the above benchmark values do not automatically constitute evidence of inappropriateness for a model. Yet, in such cases the analysts shall conduct further thorough investigations. Some basic guidelines on the additional checks that should be performed if some of the above rules have been violated are provided below:

- Potential variations in the synthesis of the relevant population are checked. In cases of significant differentiations (i.e. values of Population Stability Index larger than 0.25 on certain variables), the analysts shall suggest the review of the respective model or the development of a new, independent model for the population section that is responsible for this variation.

- Potential changes in the behaviour of the model's criteria are checked. In cases of significant differentiations (i.e. values of Population Stability Index larger than 0.25), the causes are investigated (e.g. modifications in the data entry procedures, malfunction of the information systems, changes in the data structure or in the external environment). Depending on the results of the above investigations, the analysts either ensure the restoration of the data consistency or suggest the review/re-development of the model.
- The historical accuracy measures are checked. If the tendency is steadily deteriorating, then a re-development shall be scheduled.
- In case that, during the first validation of a model after its initial development, significantly lower accuracy measurements are observed, while on the same time no significant variation of the population synthesis or the economic environment has taken place, then most likely the development sample was improper (not representative of the current population)[1]. In such cases the review or the re-development of the model shall be suggested.

5.1.4. Scoring Scale

As described in the previous chapter, the model's output is a probability of default given a vector of criteria (X) i.e. $P(Y=1|X)$. However, (a) the sample is not always representative of the actual population (e.g. often is intentionally biased for enriching it in default observations), (b) the exact default definitions of the sample may differ from the current (it is a common practice in cases of a recently redefined default definition) and (c) the PDs with which Risk Management is performed are usually considered as through-the-cycle (TTC) measures while it is not always guaranteed that the sample is representative of a full cycle period.

Moreover, a continuous scale of PDs is not usually practical, so credit scores are usually exhibited as n grades of credit quality in a scale instead of rough PDs. However, still each grade $i \in \{1,2,\dots,n\}$ is associated to a Probability of Default[2] $PD_i \in (0,1)$ so that it can participate in calculations of expected loss estimation. This grade-wide PD_i can be considered as the mean value of a range $\left(PD_{i,b}, PD_{i,e}\right]$ assigned to the grade i. Note that if $PD_{i,b} = PD_{i-1,e}$, $PD_{0,b} = 0$ and $PD_{n,e} = 1$, then the continuous range of all probable PDs can be fully distributed in grades. The above way of defining a rating scale is necessary for binding one or more models to it (see the next chapter). **WEMETRIX®** does not enforce a predefined scale; instead, a custom scale may be used per installation.

The granularity of the scale (number of grades) is very important since more grades may better separate a credit portfolio according to the counterparties' creditworthiness. However, the maximum possible number of grades while they retain significant separation ability is restricted by the model's

---

[1] This is possible if the initial development was performed despite the fact that the available sample was inappropriate. This is often the case when installing a model to a new company (where external sample are usually used) or to a company with no historical data.

[2] Those PDs are usually Through-The-Cycle (TTC) measures so that the scoring system achieves increased stability through time. However, in some cases a Point-In-Time (PIT) PD is considered more appropriate.

accuracy. For models mainly based on reliable financials or behavioural data a maximum of 9 to 12 grades is often achievable while for combined models (financials and behavioural data) a number of 14 to 15 grades is not unusual. Of course, in cases that a credit policy should be bound to the scale then the grades could be limited accordingly. Also, the model's output can be mapped to more than one scales simultaneously, one for each use.

It is worth noting that it is not unusual for different models' outputs to be mapped on the same scoring scale. This is often the case when the available information differs between the counterparties of a unique credit portfolio. As an example, for newcomers there may be available financials but not behavioural data while for the mature counterparties of the portfolio behaviour is available and no financials are collected. In this situation, two different models (one with pure financial criteria and one with behavioural) could be mapped on a unique scale so that the portfolio can be easily managed.

5.1.5. Calibration

A model generated using the above described methodology is typically well calibrated to the sample's default frequencies. The calibration process maps original model's outputs to rating scales.

The steps followed by **WEMETRIX®** to map a model's output to a given rating scale follow:

- A past sampling period is selected. The observations should refer to specific times in the past, usually 1, 2, 3 (or even more) model's horizons (often years) back.
- The data collected during the sampling process are the values of the model's criteria as they were known at the time of observation. Alternatively, if the observations of the sample are already scored by the model, the model's output is collected instead of the criteria. If through the cycle (TTC) grades are the target, an attempt is made so that the sample is as much representative to a full business cycle as possible.
- In case the above sample contains criteria and not the model's outputs then the model is retrospectively applied to the observations of the sample (back-scoring) and its outputs are stored.
- The sample is sorted in descending order according to the model's outputs and is split in *n* equal parts with orders $i \in \{1,2,\dots,n\}$. The number n of parts depends primarily on the total number of observations in the sample and secondarily on the model's discriminatory power (AR). It is usually between 40 and 100.
- For each part the mean value of the model's output is calculated as well as the mean observed default frequency (the number of defaults over the number of observations).
- The number of defaulters and non-defaulters and the conditional probabilities $Pr[I < i|N]$ and $Pr[I \leq i|N]$ are calculated for non-defaulters.
- According to the above split of the sample in parts, the sample's unconditional Probability of Default and Accuracy Ratio (AR) are calculated based on the observed defaults.
- A raw calibration curve $DF_i = c(s_i)$ is derived from the above *n* separate parts, where s_i is the mean model's output and DF_i the observed default frequency of the part i.
- The above function is often susceptible of severe fluctuation, especially if the number of observations is small. Therefore, a quasi-moment matching method is performed for smoothing the curve by assuming that:

$$PD_i \approx \frac{1}{1 + exp\left(\alpha + \beta \Phi^{-1}\left(\tilde{F}(s_i)\right)\right)}$$

and

$$\tilde{F}(s_i) = \frac{Pr[I < i|N] + Pr[I \leq i|N]}{2}$$

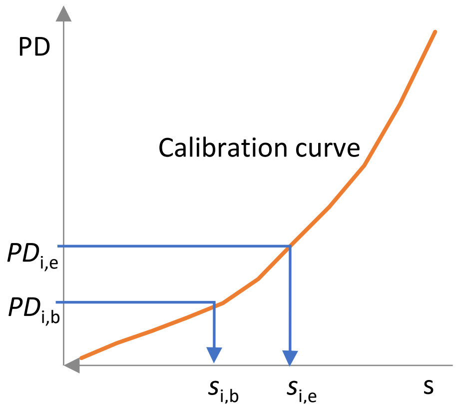
where I denotes the part in which a borrower belongs at the beginning of the observation period and α, β are estimated so that the value of the unconditional Probability of Default and the Accuracy Ratio (AR) of the original curve are retained by the new one. The term $\Phi^{-1}\left(\tilde{F}(s_i)\right)$ transforms the distribution of i conditional on non-default status (N) into an approximately normal distribution.

According to the above methodology the smoothed calibration curve is described by the above function which associates a number of model's outputs (s_i) with their PDs. Although the function is actually discrete, the known points are enough for achieving an excellent approximation of the continuous calibration curve by using interpolation:

$$PD(s) = f(s)$$

where $s \in (0,1)$ the model's continuous output.

- Finally, the assignment of the model's outputs to the scale grades is done using the above smoothed calibration curve. More precisely, the continuous space of s is split in n areas (where n is the number of grades) so that the low and high bounds of each area ($s_{i,b}$ and $s_{i,e}$ respectively) satisfy the equations $PD(s_{i,b}) = PD_{i,b}$ and $PD(s_{i,e}) = PD_{i,e}$ for every grade i of the scale.



5.2. LGD Modelling

In the literature, loss magnitude is usually expressed as a percentage loss rate: the loss given default (LGD). If a lender has a claim of 100 but receives only 40, the LGD would be (100 40)/100= 60%. Alternatively, we can capture the same information through recovery rates. The recovery rate is obtained by (1 – LGD). To complete the terminology, note that a lender's claim is usually called exposure at default (EAD).

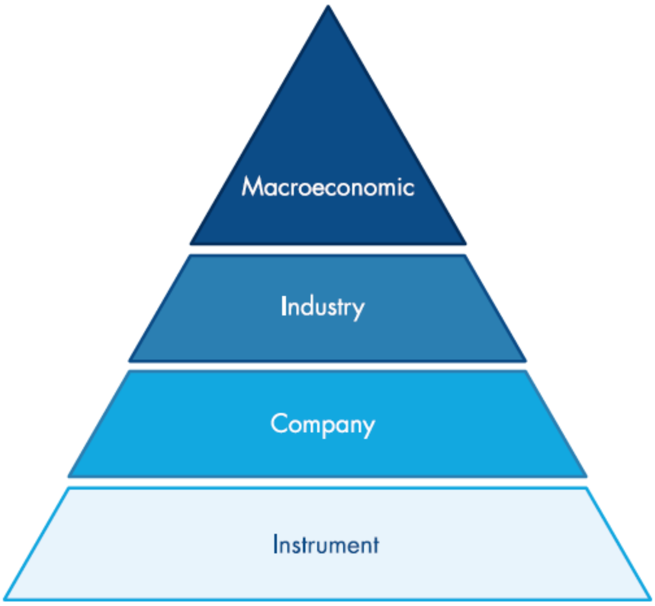
Fundamentally, LGD is associated with many factors, including:

- Instrument-specific factors:
- Loan collateral
- Debt seniority
- Company-specific factors:
- The company's financial ratio

- Industry-specific factors
- Macro-level factors

In fact, we can define LGD as 1 – recovery rate. Scorecard models, regression models, and statistical learning techniques, in general, can be used to predict recovery rates. These models can include some of the factors mentioned above (credit scores, financial ratios, macro variables) as predictors.

Figure 1 LGD and associated factors



Factors associated with loss given default (LGD).

Usually the LGD is modelled with a beta distribution. However, since most of the times data for modelling the LGD are tither scarce or nonexistent a fixed number is given which ranges between 25% to 75% percent.

“Nevertheless, it is worthwhile examining the case in which we have enough data to model LGD. The idea is to transform the dependent variable LGD so that we can expect it to be normally distributed; then we run a regression on the transformed variable and derive predictions. In a final step, these predictions are re-transformed such that they again conform to the actual LGD distribution.”[3]

“We start by identifying a probability distribution that is able to describe the observed empirical distribution of LGDs. For this purpose, the beta distribution is often used. In its standard form, it is a two-parameter distribution bounded between 0 and 1. The distribution is fully specified once we have determined its mean and standard deviation. The mean and variance of a beta-distributed variable Y are given by”

[3] (Gunther & Peter, 2011) page 122.

$$E(Y) = \frac{a}{a+b}$$

$$Var(Y) = \frac{ab}{(a+b)^2(a+b+1)}$$

where a and b are the two parameters of the distribution. Having determined estimates for the mean and the variance of observed LGDs, we can solve the above two equations to calibrate the parameters a and b:

$$a = \frac{E(Y)}{Var(Y)}(E(Y)(1 - E(Y) - Var(Y))$$

$$b = \frac{1 - E(Y)}{Var(Y)}(E(Y)(1 - E(Y) - Var(Y))$$

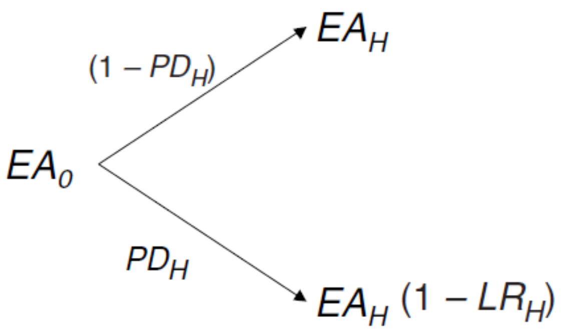
If the calibrated beta distribution provides a good description of the empirical LGD distribution, transforming the LGDs with the cumulative beta distribution function should result in a variable that is uniformly distributed on the unit interval.

5.3. Exposure at Default

Exposure at default (EAD) is the amount of exposure at the time of default. In fact, EAD is closely related to LGD because the multiplication of these two parameters will give us the amount of loss at the time of default.

5.4. Expected Loss (EL) Modelling

Expected Loss (EL) is just the product of PD, LGD and EAD (EL= PD*LGD*EAD). Since the default event D is a Bernoulli variable, that is, D equals 1 in the event of default and 0 otherwise, we can define the expected amount lost (EL) in the event of a default as follows (Ong, 1999) and (Schroeck, 2002):



Hence

$$\begin{aligned} \boldsymbol{EL_H} &= \boldsymbol{EA_H} - \mathbf{E}(\boldsymbol{EA_H}) \\ &= \boldsymbol{EA_H} - \left[(\mathbf{1} - \boldsymbol{PD_H}) * \boldsymbol{AE_H} + \boldsymbol{PD_H} * \left(\boldsymbol{EA_H} * (\mathbf{1} - \boldsymbol{LR_H})\right)\right] \\ &= \boldsymbol{PD_H} * \boldsymbol{EA_H} * \boldsymbol{LR_H} \end{aligned}$$

Where

$$PD_H = Probability\ of\ default\ up\ to\ time\ H\ (horizon)$$

$$EA_H = Exposure\ amount\ at\ time\ H$$

$$LR_H = Loss\ rate\ experienced\ at\ time\ H$$

$$E(.) = Expected\ Value\ of\ (.)$$

Therefore, EL is the product of its three determining components (EAD, PD and LGD).

6. Results

We run a series of logistic regression models that we validate with appropriate sampling statistics. We enrich these models with economic times series data.

Specifically, we have enriched our regression models with the following two economic indices:

1) The monthly percentage change of the 10-year Greek bond spread against the corresponding German bond.

2) The monthly percentage change of the Greek Stock Exchange (ASE).

In addition to the economic indices, we also used the following two time series we for each customer.

1) Total payments including VAT.

2) The total sum of all the delayed payments.

Here are the first 10 lines of data we used:

Table 4 First 10 lines of the data used

Date	Default	Payments With VAT	Sum Of All Past Due	GR 10Y Spread	Athens Composite
20160131	0	156,27	1,70	0,24	-0,12
20160229	0	206,97	1,70	0,14	-0,07
20160331	0	185,44	1,68	-0,19	0,12
20160430	0	176,43	1,68	-0,02	0,01
20160531	0	173,50	1,68	-0,14	0,11
20160630	0	173,91	1,68	0,18	-0,16
20160731	0	192,60	1,68	0,00	0,05
20160831	0	307,55	1,68	-0,01	0,01
20160930	0	294,82	1,68	0,03	-0,02
20161031	0	188,15	1,90	-0,03	0,05

The following table illustrates the final model used. We can see that the model is well defined since all the variables are statistically significant and the explanatory power of the model is relatively high for these kind of data (as indicated by the different R-squared measures). In general, an absolute number of statistics t greater than 1.96 is considered statistically significant.

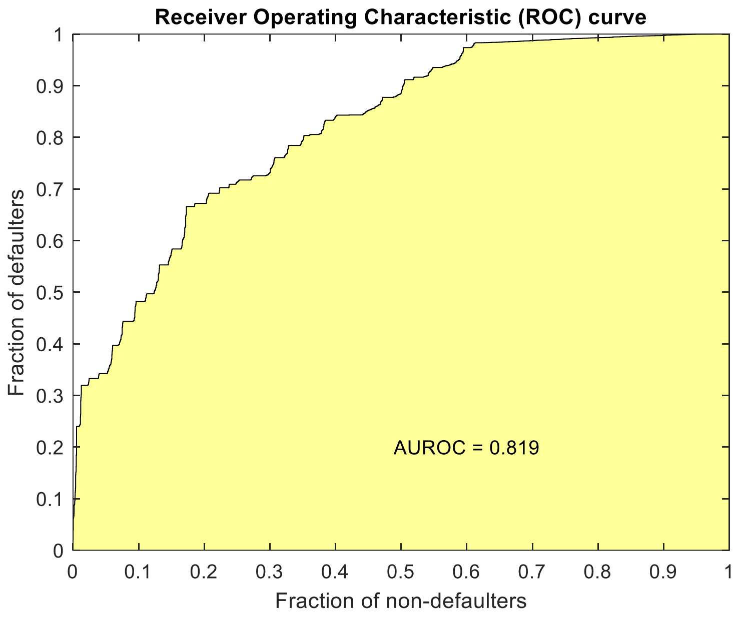
The ROC curve indicates also that the model is well-define with a value of around 81%. Typical ROC curves used in practice lie between 50% and 90%, so a score of 81% can be thought as a good indication that the model is well behaved and is able to distinguish between default and non-defaulted entities.

Table 5 Logit Model

Logit Estimates		
Dependent Variable		
McFadden R-squared	10,26%	
Estrella R-squared	11,44%	
Ordinary R-squared	16,26%	
Adjusted R-squared	16,25%	
LLR	0,11	
LR-ratio, 2*(Lu-Lr)	1968,49	
Log-Likelihood	-8611,20	
Nobs, Nvars	17090	5
# of 0's, # of 1's	4259	12831

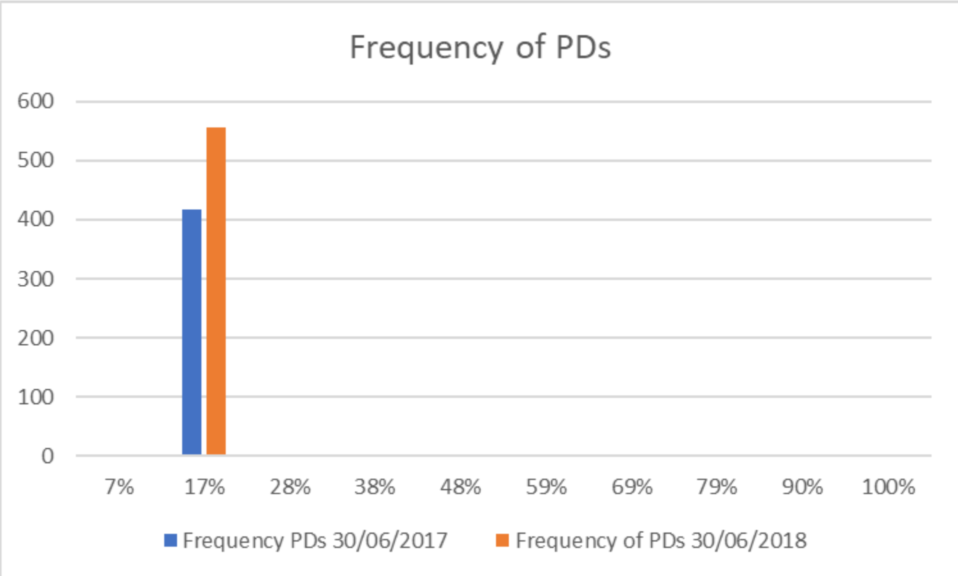
Variable	Coefficient	t-statistic	t-probability
CONST	1,15	62,62	0,00
Payments With VAT	0,00	-11,31	0,00
Sum Of All Past Due	0,00	2,80	0,01
GR10YSpread	-0,50	-2,40	0,02
Athens Composite	-0,92	-2,21	0,03

Figure 2 ROC Curve


Receiver Operating Characteristic (ROC) curve
Fraction of defaulters
Fraction of non-defaulters
AUROC = 0.819
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

The following graphs depicts the frequency of the PD as it was on the 30/06/2018 and one year before on the30/06/2017. We observe that except for a few clients with a large PD (close to 100%) most of the clients have a PD which ranges from 5% to 15% and that the distribution remains almost constant from one year to the next.

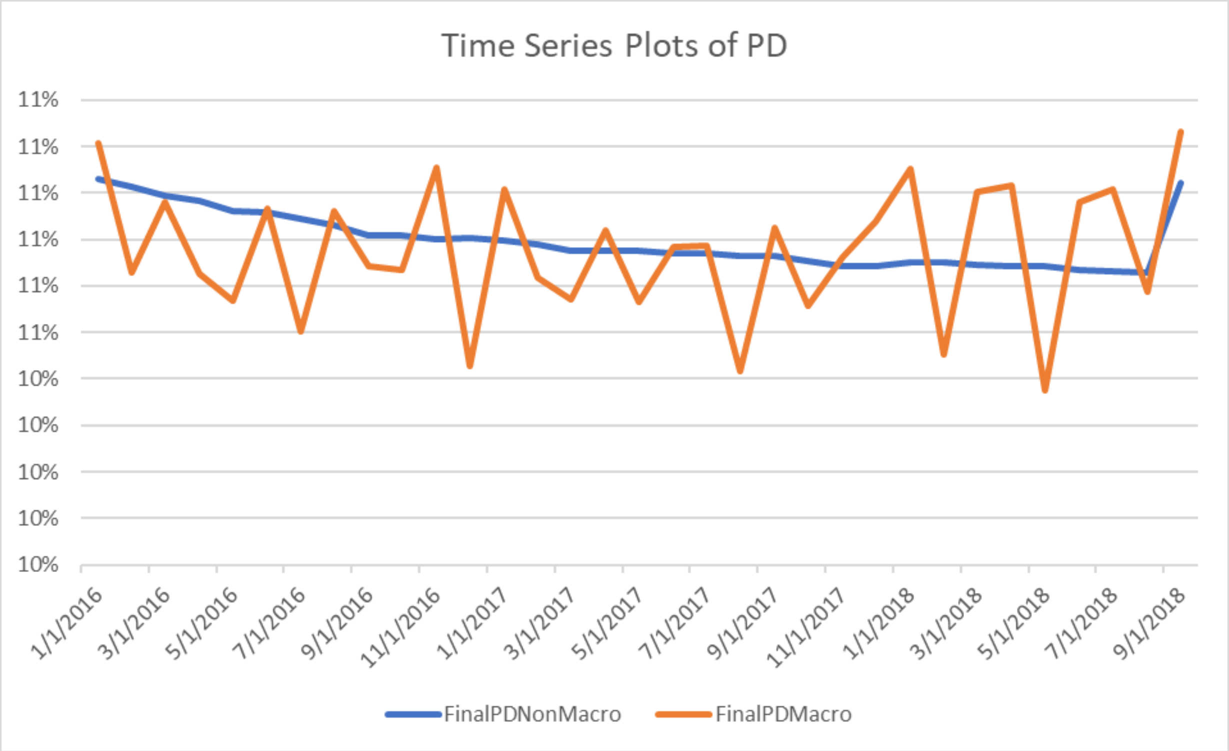
Figure 3 Frequency of PDs



The above graph is effectively a cross sectional graph with a time series component since it depicts the frequency of the PDs within a specific year and compares it to the previous year. What would be also interesting to view a time series graph of the PDs over time.

The graph below depicts the PDs from January 2016 to June 2018. The red model is the enhanced macroeconomic/financial model and the blue line is the model without the macroeconomic and financial time series. It is evident from the graph that the enhanced macroeconomic/financial model has more variability since it has a higher explanatory power than the simple model and is also able to capture the macroeconomic cyclicality of the economy.

Figure 4 Time series plots of PDs



7. Results for the Expected Losses (EL)

The expected loss (EL), sometimes also referred to as expected credit loss (ECL) measures the total amount, at the level of the receivables, that is estimated not to be collected in the next 12 months. Consequently, the provisions for bad debt that the company needs to account for is equal to the expected loss, minus the already established provisions (previous years), minus the current value of any collateral.

The expected loss arises from the product of PD * LGD * EAD, where

PD = Probability of Default, the probability of default in the next 12 months

LGD = Loss Given Default, the percentage of claims that will not be received in the event of a credit event (default)

EAD = Exposure at Default, the amount which at the reporting date is unsecured (open balance)

Below is a sample of the EXCEL sheet we provide per customer

Table 6 Expected Loss (EL) per client

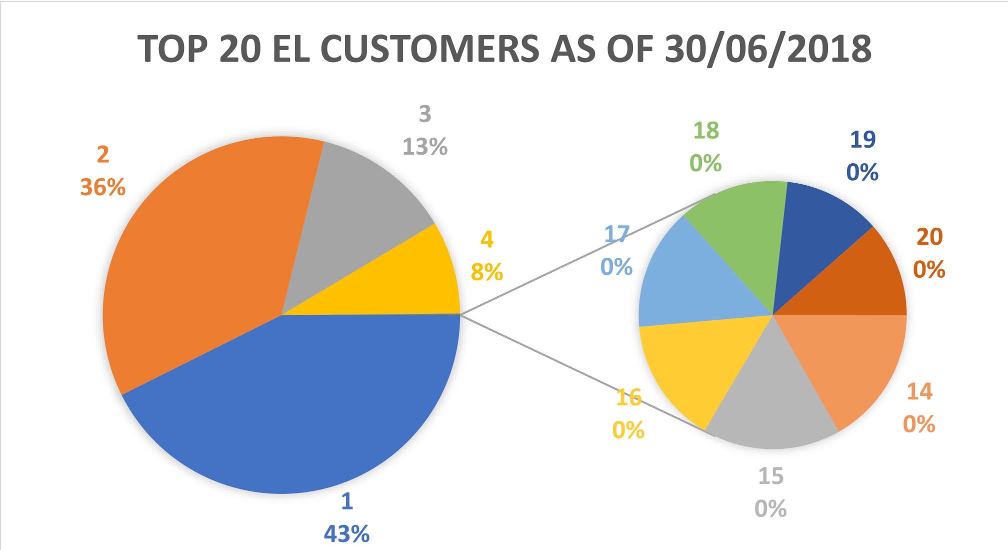
Date	CustomerID	Balance	NetBalance	Default	Past Due Over 90 Days	Payments Rolling 12M	PD	LGD	EL
20180630	1001	603.186.966,52	285.927.779,21	-	4.896.471,02	249.578.912,9767	7,0%	25%	10.555.771,91
20180630	1002	53.483.980,14	53.483.980,14	1,00	53.483.980,14	904.212,3242	100,0%	100%	53.483.980,14
20180630	1003	45.465.916,02	45.465.916,02	1,00	45.465.916,02	885.652,8483	100,0%	100%	45.465.916,02
20180630	1004	15.790.048,44	15.790.048,44	1,00	15.790.048,44	86.617,0408	100,0%	100%	15.790.048,44
20180630	1005	8.406.239,75	8.406.239,75	-	25.836,46	6.423.192,8108	10,0%	25%	211.188,94
20180630	1006	7.222.063,11	7.222.063,11	-	10.772,43	6.686.255,3108	10,0%	25%	181.439,02
20180630	1007	6.911.305,20	4.627.366,28	-	-	5.289.539,4208	10,0%	25%	173.631,88
20180630	1008	5.694.743,76	5.694.610,21	-	2.709,41	1.682.376,5675	10,0%	25%	143.068,36
20180630	1009	4.269.958,91	4.275.625,46	-	131.634,06	664.204,5225	10,0%	25%	107.273,66
20180630	1010	3.855.684,40	3.855.684,40	-	-	2.227.312,7950	10,0%	25%	96.865,89

The table below shows that the estimated damage is in the range of 119 to 125 million Euro (depending on whether you use the balance or the net balance to calculate the expected loss) for 06/30/2018. However, there are forecasts of around 117 million euros. As a result, the additional provision requirement is only 1.4 million euros. Note that the 1.4 million EUR is at the client level and that is why one cannot simply deduct the 117 provisions from the 125 or the 119 million EUR.

Table 7 Descriptive Statistics

Descriptive Statistics	Balance	NetBalance	Default	PD	LGD	EL	EL on Net Balance	Provisions	Mitigant	Additional provision requirement
Sum	769.350.960,21	449.090.462,62	519,00	63,97	154,75	126.587.636,67	120.960.199,34	117.995.612,18	52.484.023,38	1.469.851,55
Mean	1.261.231,08	736.213,87	0,85	0,10	0,25	207.520,72	198.295,41	193.435,43	86.039,38	2.409,59
Median	267,85	267,85	1,00	0,10	0,25	6,73	6,73	-	-	4,27
Min	0,01	0,01	-	0,07	0,25	0,00	0,00	-	-	-
Max	603.186.966,52	285.927.779,21	1,00	1,00	1,00	53.483.980,14	53.483.980,14	53.483.980,14	35.000.000,00	1.324.013,57
Std	24.591.846,63	11.940.479,62	0,36	0,06	0,05	2.939.546,41	2.915.956,37	2.889.120,46	1.443.778,16	53.623,95
Count	610,00	610,00	610,00	610,00	610,00	610,00	610,00	610,00	610,00	610,00
Expected Loss (EL) as Percentage to Balance						16,45%	26,93%			

The top 20 customers account for almost 100% of the expected loss. In these first 20 clients we see that the lion's share is held by the first customer with 43% followed by the second with 36% and so on.



The table below lists the top 20 customers in detail.

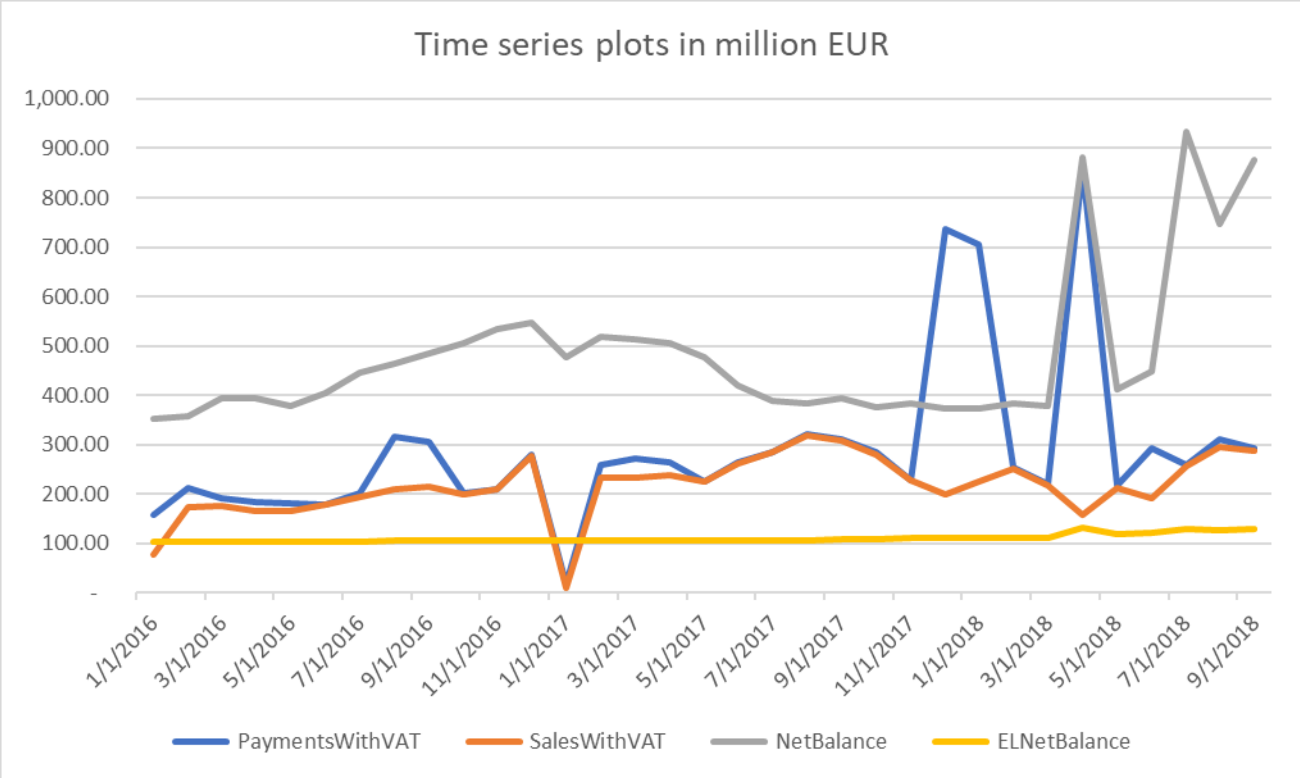
Table 8 Top 20 EL customers

TOP20EL

Position	Date	CustomerID	Balance	NetBalance	Default	PD	LGD	EL	EL on Net Balance	Provisions	Mitigant	Additional provision requirement
1	20180630	3000000318	53.483.980,14	53.483.980,14	100%	100%	100%	53.483.980,14	53.483.980,14	53.483.980,14	-	-
2	20180630	3000000224	45.465.916,02	45.465.916,02	100%	100%	100%	45.465.916,02	45.465.916,02	44.141.902,45	-	1.324.013,57
3	20180630	3000000544	15.790.048,44	15.790.048,44	100%	100%	100%	15.790.048,44	15.790.048,44	17.114.062,01	-	-
4	20180630	1000	603.186.966,52	285.927.779,21	0%	7%	25%	10.555.771,91	5.003.736,14	-	35.000.000,00	-
5	20180630	3000000523	8.406.239,75	8.406.239,75	0%	10%	25%	211.188,94	211.188,94	-	4.184.911,09	-
6	20180630	3000108959	7.222.063,11	7.222.063,11	0%	10%	25%	181.439,02	181.439,02	-	3.141.554,75	-
7	20180630	3000207828	6.911.305,20	4.627.366,28	0%	10%	25%	173.631,88	116.252,76	-	2.196.656,28	-
8	20180630	3000000522	5.694.743,76	5.694.610,21	0%	10%	25%	143.068,36	143.065,00	-	1.227.587,37	-
9	20180630	3000055276	4.269.958,91	4.275.625,46	0%	10%	25%	107.273,66	107.416,02	-	676.679,37	-
10	20180630	3000179514	3.855.684,40	3.855.684,40	0%	10%	25%	96.865,89	96.865,89	-	3.506.489,31	-
11	20180630	3000000297	3.125.036,31	2.558.799,81	0%	10%	25%	78.509,91	64.284,42	-	1.408.436,19	-
12	20180630	3000000102	1.886.592,36	1.886.592,36	100%	10%	25%	47.396,63	47.396,63	1.831.496,48	-	-
13	20180630	3000026448	1.706.841,06	1.706.841,06	0%	10%	25%	42.880,76	42.880,76	-	596.052,45	-
14	20180630	3000000521	1.148.182,39	1.000.113,95	0%	10%	25%	28.845,65	25.125,74	-	-	25.125,74
15	20180630	3000177500	921.905,29	921.905,29	0%	10%	25%	23.160,91	23.160,91	-	-	23.160,91
16	20180630	3000000270	912.176,46	912.176,46	100%	10%	25%	22.916,50	22.916,50	842.768,72	-	-
17	20180630	3000198232	747.024,73	738.425,31	0%	10%	25%	18.767,41	18.551,37	-	27.557,11	-
18	20180630	3000108906	560.481,97	560.481,97	100%	10%	25%	14.080,92	14.080,92	-	-	14.080,92
19	20180630	3000002222	488.035,63	488.035,63	0%	10%	25%	12.260,86	12.260,86	-	-	12.260,86
20	20180630	3000000172	459.921,77	459.921,77	100%	10%	25%	11.554,56	11.554,56	-	-	11.554,56

It is also interesting to form time series of the following data: Payments, Sales, Balance and Expected Loss. We can see from the graph below that except a fluctuation in May 2018, the results are relatively stable over the period we are considering. Graphs like the one below can help companies to understand their risk and monitor it over time.

Figure 5 Time Series in millions of EUR



8. Liability

WEMETRIX® will be available to provide clarifications and explanations regarding the methodology and the technical and scientific procedures it applies in order to calculate the provisions for receivables under IFRS9 and to provide written or oral relevant documentation as requested by the COMPANY, to which the statutory auditors, the shareholders, the board of directors and, in general, any person or authority may be concerned.

For this purpose, **WEMETRIX®** will make available to the COMPANY members of the project team or other executives with knowledge as to how the project was implemented for meetings, telephone conversations, teleconference or participation in presentations.

The employment of **WEMETRIX®** executives will be done in a manner that will not disrupt the work of these executives but aims to respond within 10 working days from the date of submission of the request by the COMPANY for support, except in cases of force majeure.

The commitment of **WEMETRIX®** will have a maximum duration of one (1) year from the date of publication of the Company's financial statements, which will include the results of the Company's work.

9. Bibliography

Basel Committee on Banking Supervision. (2004). International Convergence of Capital Measurement and Capital Standards, A Revised Framework Comprehensive Version. Switzerland.

Basel Committee on Banking Supervision. (2006). International Convergence of Capital Measurement and Capital Standards, A Revised Framework Comprehensive Version. Switzerland.

Christoffersen, P. (2012). *Elemenst of Financial Risk Management* (2 ed.). Elsevier Science.

De Servigny, A., & Renault, O. (2004). *Measuring and Managing Credit Risk.* McGraw-Hill.

Dougherty, C. (2016). *Introduction to Econometrics* (5 ed.). Oxford University Press.

Greene, W. (2017). *Econometric Analysis* (8 ed.). Pearson.

Gunther, L., & Peter, P. (2011). Credit Risk Modeling using Excel and VBA with DVD. Wiley.

KPMG. (2017, July). Demystifying Expected Credit Loss (ECL).

Ong, M. (1999). Internal Credit Risk Models : Capital Allocation and Performance Measurement. Risk Publications.

Schroeck, G. (2002). Risk Management and Value Creation in Financial Institutions. Wiley.

10. Glossary

Cases

EL: Expected Loss 20
LGD: Loss Given Default 18
PD: Probability of Default 10
ROC: Receiver Operating Characteristic 21

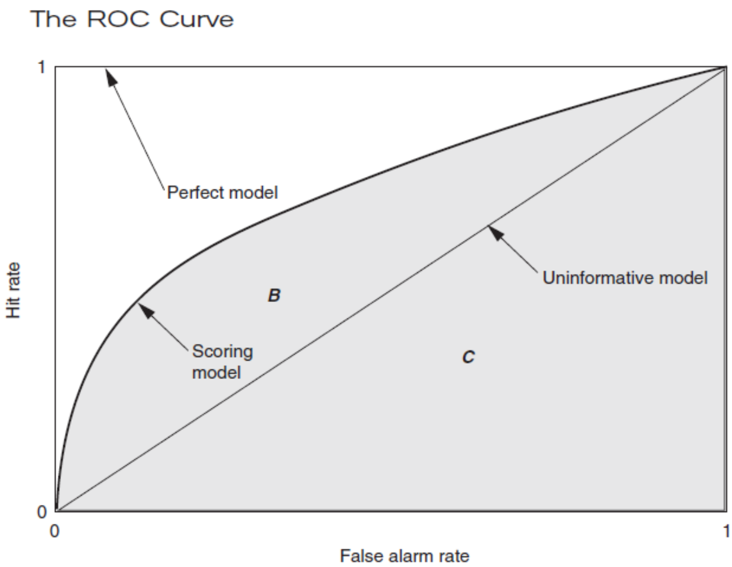
11. Appendix

11.1. Measuring Predictive Power

11.1.2. Measuring Classification: The ROC Approach

"The ROC receiver operating characteristic measure and the Gini coefficient, described later, are probably the most commonly used approaches to measuring credit scoring performance. The figure below illustrates how to calculate the ROC coefficient. We assume that we have the output of a scoring model calibrated on a population with D defaults out of N firms.

Figure 6 ROC Curve



A score s and a probability of default p_i are assigned to each firm $i = 1, \ldots\ldots N$ and the analyst chooses a cut off level T such that the firm is considered bad if $p_i > T$ and good if $p_i \leq T$. For each firm, four cases are possible:

1. It defaults, and the model had classified it as bad (appropriate classification).
2. It defaults, and the model had classified it as good (Type I error).
3. It does not default, and the model had classified it as bad (Type II error, false alarm).
4. It does not default, and the model had classified it as good (appropriate classification).

We use as C_T and F_T to denote, respectively, the number of firms correctly and wrongly classified as bad (note that they depend on the cut-off level T). Then the hit rate H and false alarm rate F are:

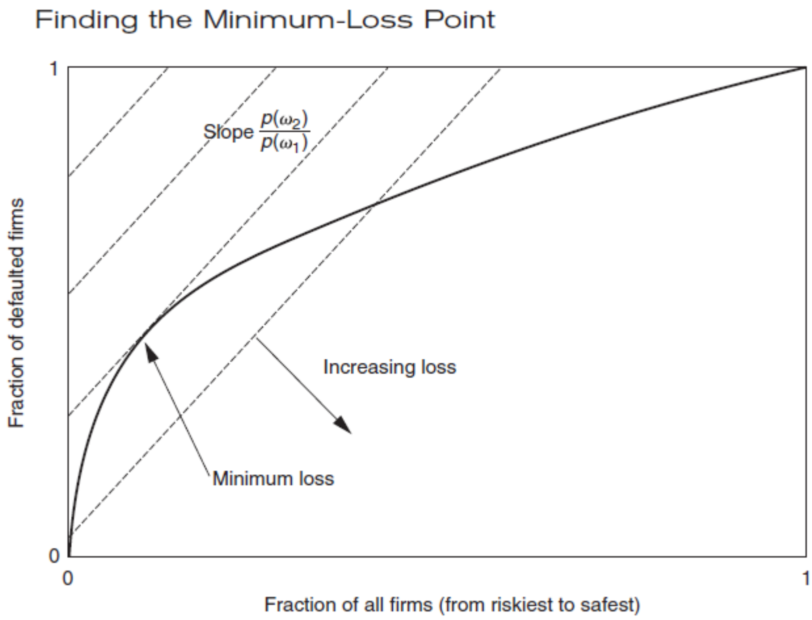
$$H_T = \frac{C_T}{D}, \quad F_T = \frac{W_T}{N-D}$$

The ROC curve is a plot of H_T against F_T. The steeper the ROC curve, the better, as it implies that there are few false alarms compared with correctly detected bad firms. On the figure the perfect model is a vertical line going from (0,0) to (0,1) and then a vertical line linking (0,1) to (1,1). An uninformative model would on average have as many false alarms as correct detections and would result in a diagonal (0,0) to (1,1) ROC curve. Credit scoring models will produce intermediate ROC curves.

The ROC curve can also be seen as a trade-off between Type I (E_I) and Type II (E_{II}) errors. $H_T$indeed, corresponds to 1 - E_I and E_{II} to F_T.

The ROC curve is also related to the minimum loss or minimum cost as defined above. One starts by defining isoloss lines whose slope is $p(\omega_2)/p(\omega_1)$ the ratio of ex ante probabilities. By definition, any point on a given line corresponds to the same loss. The minimum loss point will be at the tangency between the ROC curve and the lowest isoloss line.

Figure 7 Minimum Loss Point



The area below the ROC curve (either B or B + C) is widely used as a measure of performance of a scoring model. The ROC measure, however, suffers from several pitfalls. First, ROC is focused on rank ordering and thus only deals with relative classification. In credit terms, as long as the model produces a correct ranking of firms in terms of probabilities of default, it will have a good ROC coefficient, irrespective of whether all firms are assigned much lower (or higher) probabilities than their "true" values. Therefore, one may have a model that underestimates risk substantially but still has a satisfactory ROC coefficient. Second, ROC is an acceptable measure as long as the class distribution is not skewed. This is the case with credit, where the no defaulting population is much larger than the defaulting one. ROC curves may not be the most adequate measure under such circumstances".[4]

11.1.3. Measuring Classification: The Gini/CAP Approach

"Another commonly used measure of classification performance is the Gini curve or cumulative accuracy profile (CAP). This curve assesses the consistency of the predictions of a scoring model (in terms of the ranking of firms by order of default probability) to the ranking of observed defaults. Firms

[4] (De Servigny & Renault, 2004) page 95-96.

are first sorted in descending order of default probability as produced by the scoring model (the horizontal axis in the figure). The vertical axis displays the fraction of firms that have actually defaulted.

A perfect model would have assigned the D highest default probabilities to the D firms that have actually defaulted out of a sample of N. The perfect model would therefore be a straight line from the point (0,0) to point $(D/N, 1)$ and then a horizontal line from $(D/N, 1)$ to (1,1). Conversely, an uninformative model would randomly assign the probabilities of defaults to high-risk and low-risk firms. The resulting CAP curve is the diagonal from (0,0) to (1,1).

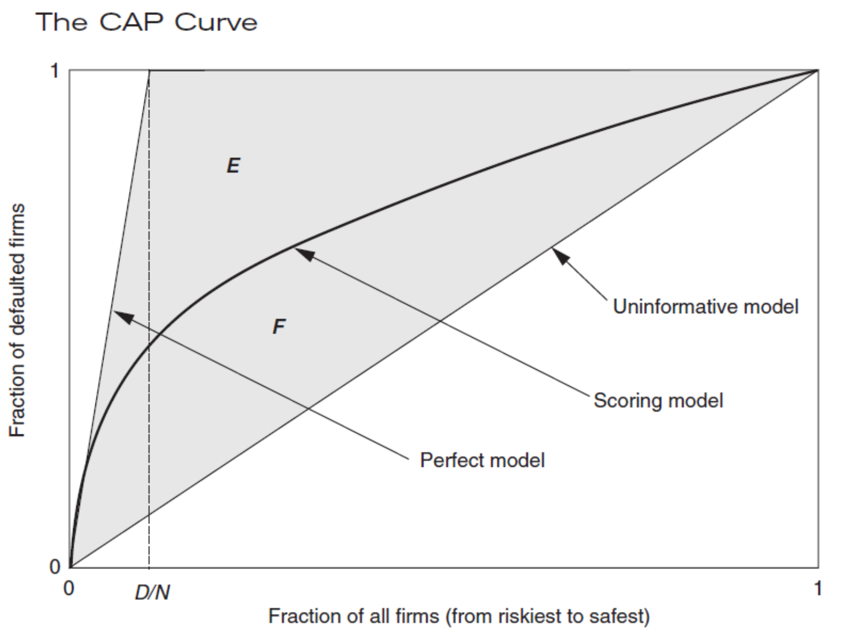
Any real scoring model will have a CAP curve somewhere in between. The Gini ratio (or accuracy ratio), which measures the performance of the scoring model for rank ordering, is defined as $G = F/(E + F)$. This ratio lies between 0 and 1; the higher this ratio, the better the performance of the model.

The CAP approach provides a rank-ordering performance measure of a model and is highly dependent on the sample on which the model is calibrated. For example, any model that is calibrated on a sample with no observed default and that predicts zero default will have a 100 percent Gini coefficient. However, this result will not be very informative about the true performance of the underlying models. For instance, the same model can exhibit an accuracy ratio under 50 percent or close to 80 percent, according to the characteristic of the underlying sample. Comparing different models on the basis of their accuracy ratio, calculated with different samples, is therefore totally nonsensical.

When the costs of misclassification are the same for Type I and Type II errors (corresponding to the minimum-error Bayesian rule), the summary statistics of the ROC and the CAP are directly related: If $A = B + C$ is the value of the area under the ROC curve and G is the Gini coefficient or accuracy ratio calculated on the CAP curve, then $G = 2(A - 0.5)$.

In this case the ROC curves and CAP curves convey exactly the same information. When a specific structure of costs of misclassification is introduced in the calculation of ROC, the link between the two curves is lost. ROC can probably be considered as more general than CAP because it allows for differing costs to be selected by the user.

Figure 8 CAP Curve



Overall:

- The CAP curve provides valuable information if the user considers that misclassification costs are equal and provided the size of the defaulting subsample is somehow comparable with the non-defaulting one. If this is not the case, then another type of measure should complement the comparison of the performance of different models.
- Since bankers and investors are usually risk averse and would tend to avoid Type I errors more than Type II errors, CAP curves or Gini coefficients are not best suited to assess the performance of credit scoring models.
- The ROC measure is broader than the CAP measure because it enables the users of the model to incorporate misclassification costs or their utility function. If the objective is to assess the ability if the model to classify firms, then ROC, and in particular ROCC is an attractive measure.
- A significant weakness of both the ROC and CAP approaches is that they are limited to rank ordering. This is a weak measure of performance in the area of credit where not only the relative riskiness but also the level of risk is crucial."[5]

11.2.

[5] (De Servigny & Renault, 2004) pages 96-99.

12. Contact Information

Mr. Chris Giannakopoulos, Director Head of Operations, Email: chrisg@WEMETRIX®.com, tel: +30 210 6679852.

Mr. Apostolos Ypsilantis, Project Manager, Email: apostolosy@WEMETRIX®.com, tel: +30 210 6679491.

Dr. Panagiotis Ballis-Papanastasiou, Manager, Email: panosbpy@WEMETRIX®.com, tel: +30 210 6679851.