Methodological Notes on Rating Development

Bilel Ammouri

August 31, 2023

Contents

1 Introduction

The ratings published by rating agencies are exclusively accessible to large companies. Many SMEs refrain from issuing publicly traded bonds; consequently, they remain unrated due to the rating agencies' focus on larger corporations. This document presents a methodology for "Credit rating" – a means of furnishing dependable information concerning credit quality. This can be conveyed as a numerical value or a combination of letters; regardless, its essence lies in communicating crucial information about quality. Much like a probability of default, it serves as an indicator of the creditworthiness of corporate bonds. Notably, it pertains to the attributes of a bond issue rather than those of a company.

2 Internal ratings based (IRB) approach

As observed, the internal ratings-based approaches of Basel II and III empower banks to employ proprietary methodologies for gauging the likelihood of default by a counterparty. This document zeroes in on the internal rating of corporate and retail clients, particularly in cases where external data is absent. Internal rating methodologies commonly lean on indicators such as profitability and balance sheet ratios. Given a bank's vested interest in the counterparty's capacity to meet its financial obligations, the liquidity of each client holds paramount significance (see figure 1.

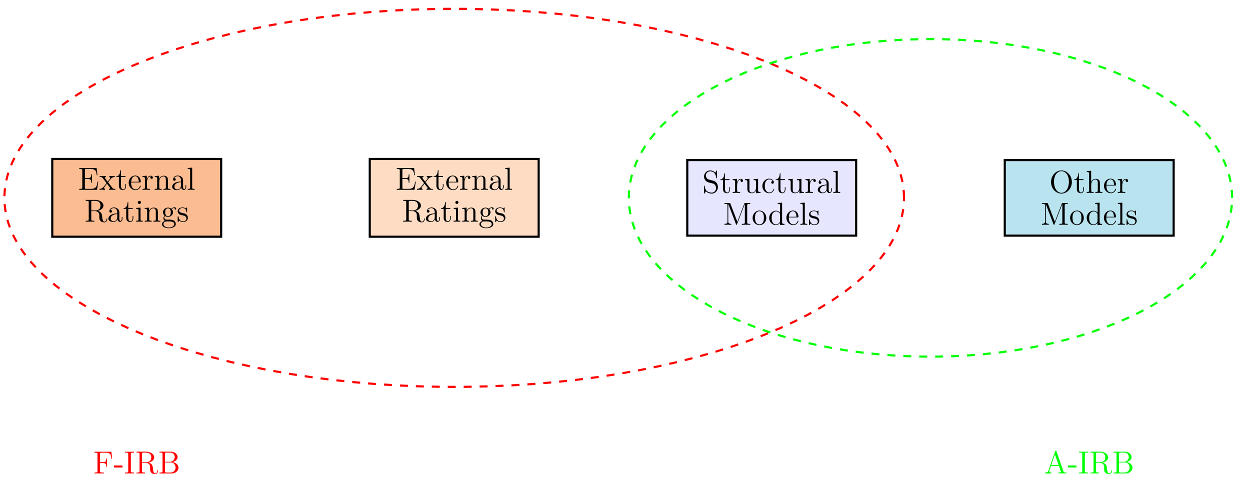

Figure 1: Internal Rating Based Approach

3 Internal ratings models

The internal rating models are divided into three categories: Causal models, Statistical models (which encompass various subtypes), and Hybrid forms (see Figure 2).

3.1 Causal models

Causal models establish direct analytical links to creditworthiness based on financial theory.

3.1.1 Option Pricing Model

The OPM (Option Pricing Model) facilitates the valuation of default risk for individual transactions, without the need for an extensive default history. However, this approach does necessitate data on the economic value of debt and equity, with a particular emphasis on volatilities.

Note

This approach is also applicable in situations where an adequate dataset of unfavorable cases is unavailable for the development of statistical models (DAM & LR).

3.1.2 Cash flow models (simulation)

Cash flow models are particularly effective for assessing credit in specialized lending transactions. In this context, creditworthiness hinges primarily on the future cash flow generated by the financed assets. Here, the evaluation centers on the transaction itself, rather than a specific borrower, resulting in what is known as a transaction rating.

Cash flow can be defined from various angles. For capital market company valuation purposes, free cash flow is derived by subtracting investments from EBITDA. Typically, the average free cash flow over the past five years provides the starting point for calculating a company's value. Dividing this average by the weighted costs of equity and debt

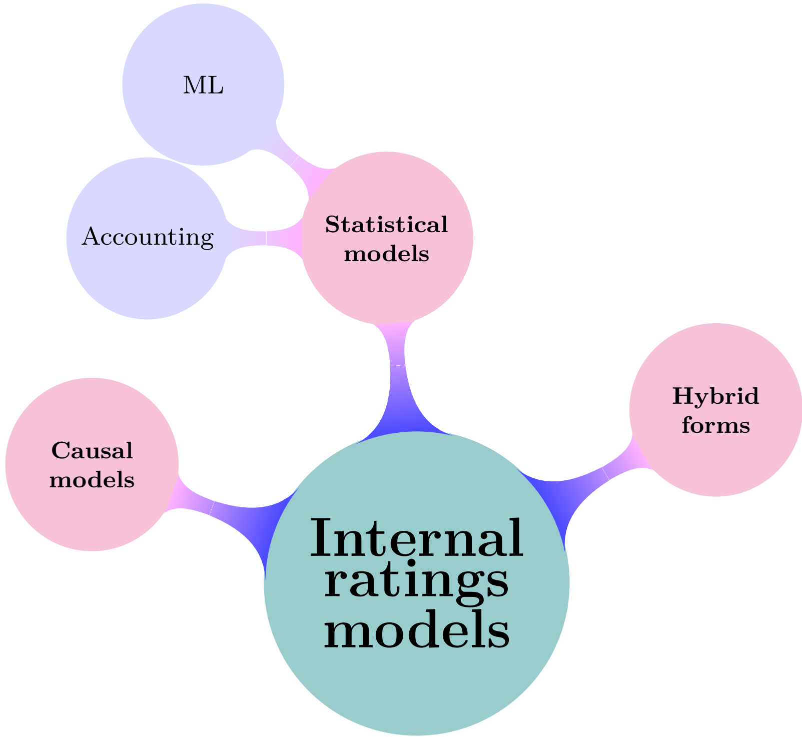

Figure 2: Internal ratings models

financing yields the company value, which can then serve as an input for the option pricing model.

The volatility of this value can be computed through:

- Regression models, which minimize observed deviations,
- Stochastic time series models: simulation methods generate and weigh potential future cash flow realizations based on historical data by developing macroeconomic models.

3.2 Statistical models

Most rating models are classified as statistical models. The statistical models commonly used in practice include:

i **Discriminant analyses**: Developed by Beaver (1966), Altman (1968), and Mare et al. (2017).

ii **Binary response models**: Introduced by Ohlson (1980), Zmijewski (1984), Fareman (2003), Campbell et al. (2008), Kukuk and Rönnberg (2013), and Aretz et al. (2018).

iii **Hazard models**: Created by Shumway (2001), Chava and Jarrow (2004), Nam et al. (2008), Bonfim (2009), Dakovic et al. (2010), Duan et al. (2012), Figlewski et al. (2012), Tian et al. (2015), and Traczynski (2017).

3.2.1 Altman's Z-score

This model was introduced in 1968 by Edward Altman. The renowned Z-score serves as a prototype for internal rating methods. In straightforward terms, it functions as a financial distress index and is considered a fundamental analysis essential.

Formula

$$Z - score = 1.2 \times R_1 + 1.4 \times R_2 + 3.3 \times R_3 + 0.6 \times R_4 + 0.999 \times R_5 \tag{1}$$

Where,
R_1: Working capital/Total assets
R_2: retained earnings/Total assets
R_3: EBIT/Total assets
R_4: Market value of equity/Book value of total liabilities
R_5: Sales/ Total assets.

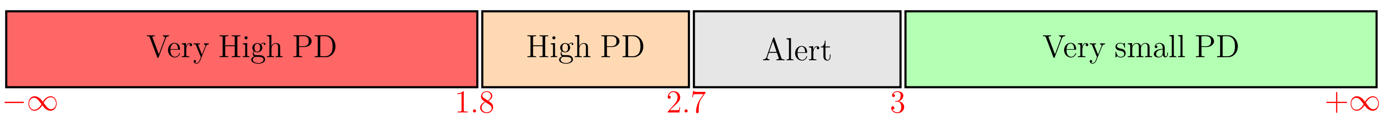

Figure 3: Z-score's scale

3.2.2 O-score

In 1980, Ohlson proposed an updated version of the Z-score, incorporating nine variables. This revised version includes a combination of financial ratios and specific dummies, aimed at enhancing the model's predictability.

Formula

$$\begin{aligned} O-score = &-1.32 - 0.407 \times Size + 6.03 \times TLTA - 1.43 \times WCTA + 0.0757 \times CLCA - 2.37 \times NITA \\ &- 1.83 \times FULT + 0.285 \times INTWO - 1.72 \times ENEG - 0.521 \times CHIN \end{aligned} \tag{2}$$

Where,
$SIZE$: is the log of the ratio of total assets to the GNP price level index. The index assumes a base value of 100 for 1985.
$TLTA$: Total liabilities/Total assets
$WCTA$: Working capital/Total assets
$CLCA$: Current liabilities/Current assets
$NITA$: Net income/Total assets
$FUTL$: Cash flows from operations/Total liabilities
$INTWO$: One if net income was negative for the last two years, zero otherwise.
$OENEG$: One if total liabilities are greater than total assets, zero otherwise.
$CHIN$: (NIt-NIt-1)/(|NIt|+|NIt-1|), where NI is Net Income.

3.2.3 Zmijewski's

In 1984, Zmijewski proposed a model based on probit analysis, formulated according to the following equation:

Formula

$$X = -4.3 - 4.5 \times X_1 + 5.7 \times X_2 + 0.004 \times X_3 \tag{3}$$

Where,
X_1: Net income/Total assets
X_2: Total liabilities/Total assets
X_3: Current assets/Current liabilities.

Note

- Accounting models have faced criticism due to their reliance on historical information as input and their failure to consider a firm's asset volatility when estimating default risk.
- Contemporary credit risk models in financial literature leverage data from capital markets, where a company's shares or bonds are traded.
- In theory, market prices embody investor expectations regarding a firm's future performance. These prices encompass forward-looking information that is ideally suited for calculating the probability of a firm defaulting in the future.

3.2.4 Multivariate Discriminant Analysis

The primary aim of Multivariate Discriminant Analysis (MDA) is to effectively differentiate between solvent and insolvent borrowers using a function that incorporates multiple independent creditworthiness criteria. In Linear Multivariate Discriminant Analysis, a weighted linear combination of indicators is developed to achieve optimal classification of both positive and negative cases, aiming to maximize the discriminatory power based on the computed outcome.

The advantage of utilizing MDA in contrast to other classification procedures lies in the direct interpretability of the linear function and the individual coefficients in economic terms.

Formula

$$D = \alpha_0 + \alpha_1 k_1 + \alpha_2 k_2 + \ldots + \alpha_n k_n \tag{4}$$

Note

- MDA operates effectively with quantitative data; in the case of qualitative data, rescaling becomes necessary. Lancaster scaling can be employed for this purpose.
- MDA requires a normal distribution of variables.

3.2.5 Binary response models (Logistic regression)

A binary response model categorizes a corporation's state as either normal (denoted by a "0" characteristic) or in default (denoted by a "1" characteristic). It calculates the likelihood of default through explanatory variables and

typically employs a logistic or probit function, commonly referred to as an O-score.
A binary response model offers several advantages:

- A BRM doesn't necessitate assumptions regarding the probability of default or the distributions of predictor variables.
- It permits testing the significance of individual independent variables and enables the computation of the probability of default in the subsequent period.

Note

- The outcome of Logistic Regression (LR) can be construed as the likelihood of belonging to a particular group.
- Relative to MDA, Logistic Regression is marked by more robust and accurate outcomes.
- In banking, one practical application of Logistic Regression is evident in the BVR-II rating model utilized by the Federal Association of German Cooperative Banks to assess SMEs.
- Logistic Regression does not necessitate the input indicators to follow a normal distribution.

Formula

$$P = \frac{1}{1 + \exp\{-(\alpha_0 + \alpha_1 k_1 + \alpha_2 k_2 + \ldots + \alpha_n k_n)\}} \tag{5}$$

3.2.6 Hazard models

Shumway employs a duration analysis through a hazard model, demonstrating its superior predictive capacity for corporate defaults compared to traditional single-period models. The hazard model, also known as survival analysis, is instrumental in estimating the likelihood of a corporate default occurring over a span of time (utilizing techniques such as Cox's hazard regression).

3.2.7 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) leverage information technology to mimic the information processing mechanisms of the human brain. NNs have the capability to handle both quantitative and qualitative data directly, rendering them particularly well-suited for modeling complex rating systems that need to encompass diverse information categories.

Note

- ANNs are often regarded as non-transparent to users, functioning as black box models.
- Ratings models employing ANNs typically achieve high to very high levels of discriminatory power.

Note

- The advantages of Statistical and Causal models lie in their objectivity and typically superior classification performance in contrast to other models.
- Statistical and causal models are constrained in their ability to handle only a limited number of creditworthiness factors.

3.3 Hybrid forms

Hybrid (heuristic) models involve the integration of one of the two other model types, namely statistical models or causal models. This approach is regarded favorably due to its ability to effectively combine and leverage the strengths of different methodologies. This integration facilitates the incorporation of credit experts' knowledge, a capability that statistical models inherently lack in handling qualitative information directly. The heuristic component of this approach engenders a higher level of involvement from credit experts in the rating process compared to automated credit assessment using a solely statistical or causal model.

3.3.1 Horizontal Linking of Model types

Statistical and causal models are employed for analyzing annual financial statements or assessing a borrower's financial position. Qualitative data is evaluated through a heuristic module within the model. Subsequently, the outputs generated by these two modules can be combined to produce a comprehensive credit assessment (see figure 4).

Note

A practical example of this model can be observed in the Deutsche Bundesbank's:

- Annual financial statements undergo analysis through statistical discriminant analysis
- This qualitative creditworthiness analysis is complemented by additional qualitative criteria, evaluated utilizing a fuzzy logic system.

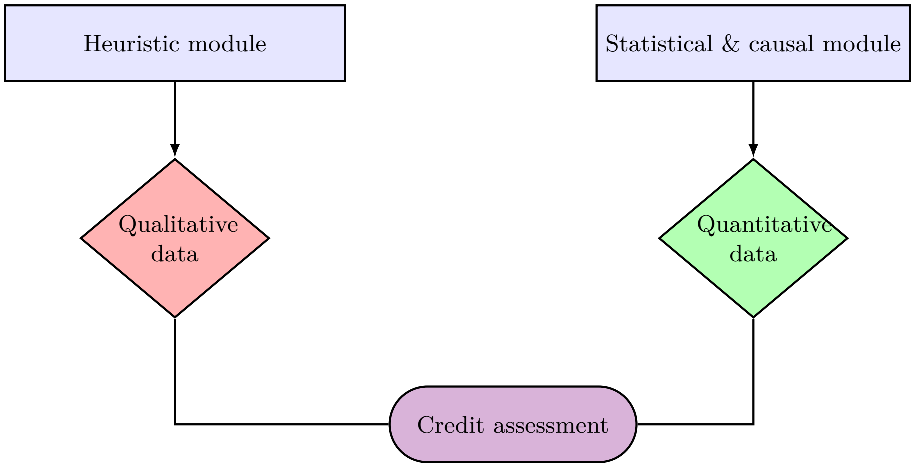

Figure 4: Horizontal Linking of Rating Models

3.3.2 Vertical linking of model types using overrides

This approach establishes a connection between partial ratings to formulate an initial classification, subject to potential modifications by credit experts to arrive at the ultimate credit rating. In this process, both quantitative and qualitative creditworthiness attributes are initially evaluated using a statistical or causal model, culminating in a proposed classification. This suggested classification is then open to adjustments by credit analysts, leveraging their specialized knowledge.

The heuristic component plays a pivotal role in incorporating creditworthiness factors that are exclusively known to the credit analyst and might not be encompassed by the preceding module. While this module is meticulously designed, overrides should ideally be necessary only in specific instances. An excessive use of overrides might suggest a potential lack of user acceptance or an inadequate comprehension of the rating model, warranting a thorough review during the validation process (see figure 5).

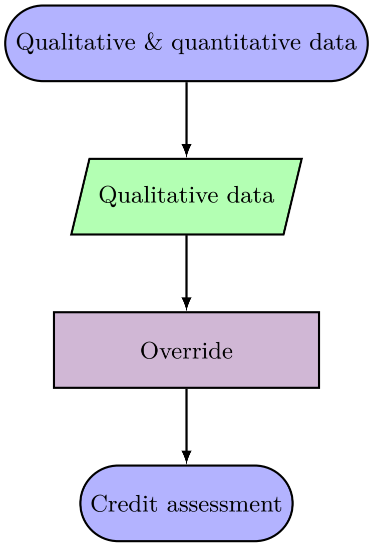

Figure 5: Vertical Linking of Rating Models using Overrides

3.3.3 Upstream inclusion of heuristic Knock-Out criteria

Preceding this module are Knock-out criteria, which are established based on the practical insights of credit experts and the specific strategy of the bank. Should a potential borrower satisfy a Knock-out criterion, the subsequent credit assessment process does not progress to the statistical module.

4 Rating WORKFLOW

The project steps will be classified as follows (see Methodological Notes on Scorecard at origination Development[1]):

1. Raw Data Collection
2. Data Preprocessing

[1]DOI: 10.13140/RG.2.2.33901.56801

3. Data Transformation
4. Sampling
5. Training model
6. Cross Validation
7. Final model