Central European Review of Economic Issues
EKONOMICKÁ REVUE

Modelling of Rating Downgrades Based on Multiple Failure-Time Data

Martina NOVOTNÁ*

Department of Finance, Faculty of Economics, VŠB – Technical University, Sokolská třída 33, 70200, Ostrava, Czech Republic

Abstract

This article aims to develop rating models based on survival analysis methods. The focus is on the use of the Cox proportional hazards model to analyse the time to an event defined as a rating downgrade and to examine the effect of selected financial variables on the rating. Two different approaches are used to estimate the models depending on whether we are considering one or multiple events for a subject. The results show that the probability of a rating downgrade is affected by annual changes in financial variables. Furthermore, the application indicates that the study of multiple failure-time data leads to a more suitable model based on the statistical significance of the estimated coefficients and the goodness of fit. Overall, the main findings suggest that it is more appropriate to use multiple failure-time analysis, which corresponds better to a given problem and allows the use of all the available data, for modelling rating downgrades.

Keywords

Cox model; credit risk; hazard function; rating; survival analysis

JEL Classification: C14, G24, G32

---

* martina.novotna@vsb.cz
This paper was supported by the Operational Programme for Competitiveness under Project CZ.1.07/2.3.00/20.0296 and Project SGS SP2019/132.

ISSN 1212-3951 (Print), 1805-9481 (Online)
ER-CEREI, Volume 22: 45–56 (2019).
doi: 10.7327/cerei.2019.06.02

Modelling of Rating Downgrades Based on Multiple Failure-Time Data

Martina NOVOTNÁ

1. Introduction

Rating analysis is a current topic as it is related to assessing the credibility of the debtor or issuer of debt securities. Some research on bond rating dates back to the 1950s, for example the studies by Hickman (1958) and Fisher (1959). Regression analysis became one of the most used methods to estimate ratings in this period. An alternative approach to predicting bond ratings is the multiple discriminant analysis introduced, for example, by Pinches and Mingo (1973), Ang and Patel (1975), Altman and Katz (1976) and Belkaoui (1980). Subsequent research compared particular statistical methods; for instance, Kaplan and Urwitz (1979) compare ordered probit analysis with ordinary least square regression and Wingler and Watts (1980) compare ordered probit analysis with multiple discriminant analysis. Recent studies come from the theoretical framework mentioned above and extend the statistical methods to new non-conservative approaches, such as neural networks (Dutta and Shekhar, 1988; Surkan and Singleton, 1990). Waagepetersen (2010) assesses the relationship between quantitative models and expert rating evaluation. In a more recent study, Altman, Sabato and Wilson (2010) focus on the importance of non-financial information within risk management.

In recent years, research on rating prediction has shifted mainly to the application of machine learning methods. Hsu, Chen and Chen (2018) propose a model based on the artificial bee colony approach and support the vector machine technique. The authors find that their bio-inspired computing mechanism provides improved prediction accuracy compared with other statistical methods. Golbayani, Florescu and Chatterjee (2020) compare neural networks, support vector machines and decision trees, finding that decision tree-based models achieve the best performance. In their research, they apply conventional accuracy measures and introduce the so-called notch distance approach, which is suitable for comparing the performances of various machine learning methods. As it turns out, all the methods mentioned above are appropriate for rating prediction. The individual models differ mainly in their methodology, the variables used and their ability to predict ratings. Therefore, research in this area is currently focused primarily on improving the predictive accuracy of the models. For example, Wang and Ku (2021) develop the parallel artificial neural networks model, which creates several independent artificial neural networks. As the authors suggest, their approach obtains competitive results compared with conventional artificial intelligence techniques.

The current research shows that conventional approaches and newer methods based on artificial intelligence are widely used to model credit ratings. The huge advantage of these models is their practical applicability and the possibility to use them for potential rating revisions. In addition, research studies show that models' predictive power is sufficient and comparable to other commonly used methods in valuing historical data based on averages or growth rates. For example, Jones, Johnstone and Wilson (2015) examine the predictive performance of binary classifiers using a large sample of international credit ratings. They apply conventional techniques (logit and probit regression and linear discriminant analysis) and fully nonlinear classifiers (neural networks, support vector machines, general boosting, AdaBoost and random forests). The authors conclude that, although the newer classifiers outperform the older ones, simpler classifiers can be viable alternatives to more sophisticated approaches, particularly if interpretability is an important objective of predictive models.

An alternative way to assess ratings, which is used in this article, is based on survival analysis. For example, Glennon and Nigro (2005) use a discrete-time hazard framework for measuring the default risk of small business loans. Roa, García and Bonilla (2009) propose a survival analysis methodology to analyse falling rating duration. They test macroeconomic variables to predict this event in selected countries and find differences between developed and emerging economies. Zhang and Thomas (2012) compare linear regression and survival analysis for modelling recovery rates in further research. The authors find that linear regression is better for recovery rate modelling; however, they suggest some adjustments and additional validation. Overall, survival analysis methods are typically used for modelling rating or credit transitions and time series rating patterns (e.g. Parnes, 2007; Figlewski, Frydman and Liang, 2012; Louis, Van Laere and Baesens, 2013;

Leow and Crook, 2014). Thus, we can model the rating behaviour over time and measure, for example, the probability of a certain change in the rating depending on time and other relevant variables. Therefore, survival analysis allows us to gain a better understanding of the data and their dynamics over time. In addition, it is a method used by rating agencies to estimate default rates, which are regularly published and used by analysts and researchers in the financial market.

It is evident that the current research deals with scoring and rating models and that many studies on this topic have been published. Still, academic research pays more attention to scoring than rating models. Scoring models are typically used in predicting credit default or corporate bankruptcy. Various modifications of these models have a wide application in corporate finance when assessing a company's financial health, especially in the banking sector, in which the probability of loan repayment default is analysed. Rating models work similarly as their purpose is to assess the rating of a debt instrument or issuer. These models are also important in corporate finance, especially in analysing debt securities and the investment process. Their application is vital when a certified rating is not available. These models can also play a key role in valuing bonds that are not traded in the public market. In these cases, it is especially necessary to consider the risk of default carefully, and this is when the models can be used.

The motivation for our research is the insufficient attention paid to this issue and the effort to understand the rating behaviour in selected countries over time. The aim of this paper is to model the probability of a rating downgrade using the method of survival analysis. The focus is on assessing the impact of financial variables on negative rating changes. In this study, we apply the Cox proportional hazards model to identify the economic variables with great potential to signal a deterioration in credit ratings. The survival models are estimated using two approaches. First, we examine the time to the first rating downgrade, ignoring additional events. Next, we make use of all the available data and account for multiple failure-type data. This procedure allows us to compare the two approaches and make recommendations. The estimated models are designed to evaluate the individual rating of an entity, security or other debt instrument. The practical application consists mainly of the valuation of debt securities and the related estimation of the risk premium and the cost of capital.

The structure of this article is as follows. First, the main concepts and methods of survival analysis are described in Chapter 2. Attention is paid especially to the Cox proportional hazards model and the study of the multiple-time data used in the application. Then, in Chapter 3, the data and variables entering the model are described. Subsequently, the models are estimated and the main results are interpreted. Finally, the overall findings are summarized in Chapter 4.

2. Description of the Methodology

Survival analysis is a statistical method used to analyse the probability of an event occurring as a function of time. This method is primarily used in the natural sciences, in which, for example, the likelihood of patient survival from the moment of diagnosis, initiation of treatment and so on is examined. In economic and related fields, this method is applied mainly to the analysis of bankruptcy or default. In our study, survival analysis is used for rating modelling. This chapter is devoted to explaining the methodology used, including the terminology and main principles.

Survival analysis should be used to analyse data in which the time until the event is of interest. The response variable, the time until that event, is typically called the failure time, survival or event time (Harrell, 2010). Survival analysis allows the response to be incompletely determined for some subjects; for example, we cannot follow all the observations in the dataset. This method is based on the mechanism of censoring when censored and uncensored observations are defined. For example, Hosmer et al. (2008, p. 18) describe a censored observation as an incomplete value due to random factors for each subject. In the following text, we focus on the fundamental theoretical background, which can be supplemented by a variety of relevant literature, for example Gourieroux and Jasiak (2007), Tabachnik and Fidell (2007), Hosmer et al. (2008), Cleves et al. (2010), Harrell (2010), Royston and Lambert (2011) and Klein et al. (2014).

2.1 Main Concepts of Survival Analysis

A key issue in survival analysis is to estimate the likelihood that subjects will survive a certain length of time. The likelihood of survival to a certain point in time is conditioned by the fact that the subject has survived the previous length of time. The total probability of survival is then given by the product of these individual probabilities, for example

$$S(t) = p_1 p_2 p_3 ... p_t, \qquad (1)$$

where $p_1 p_2 p_3 ... p_t$, is the conditional probability of surviving time t after having survived time $t-1$. The p_t can be expressed as

$$p_t = \frac{(n_t - d_t)}{n_t} = 1 - \frac{d_t}{n_t}, \qquad (2)$$

where n_t is the number of subjects alive at the start of the interval ending at time $t+1$ and d_t is the number of subjects failing in the short time interval just after t. Thus, equation (1) can be rewritten as

$$S(t) = \prod_t \left(1 - \frac{d_t}{n_t}\right). \tag{3}$$

The successive overall survival probabilities, $S(1), S(2), ..., S(t)$, are referred to as the Kaplan–Meier (K-M) or product-limit survival estimates. A graphical representation of $S(t)$ as a function of time *t* is the K-M estimate of the survival curve, which is plotted as a step function. The K-M estimate describes the time to a given event based on all the available data.

Klein et al. (2014) define the distribution function of the survival time, commonly called the failure function, as the probability of failure up to time *t*,

$$F(t) = \Pr(T \leq t), \tag{4}$$

where *T* is a non-negative random variable denoting the time to a failure event and *F(t)* refers to the cumulative distribution. For practical reasons, it is often more appropriate to use a complementary function in survival analysis, the survival function *S(t)*, which is the probability of surviving beyond time *t*,

$$S(t) = 1 - F(t) = \Pr(T > t). \tag{5}$$

Using the survival function, we can estimate the probability of no failure event occurring prior to *t*. The density function *f(t)* can be obtained both from *S(t)* and from *F(t)*:

$$f(t) = \frac{dF(t)}{dt} = \frac{d}{dt}\{1 - S(t)\} = -S'(t). \tag{6}$$

To assess how the risk of a particular outcome varies with time, we use the hazard rate $h(t)$. According to Harrell (2010), the hazard at time t is related to the probability that the event will occur in a small interval around t, given that the event has not occurred before time t. Cleves et al. (2010) explain the hazard rate as the conditional failure rate or the intensity function. As they emphasize, the hazard rate represents the instantaneous rate of failure with $1/t$ units:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t \mid T > t}{\Delta t} = \frac{f(t)}{S(t)}. \tag{7}$$

The hazard function can range from zero (no risk) to infinity (the certainty of failure at that instant) and can be decreasing, increasing or constant; alternatively, it can even take on other shapes.

Depending on the assumptions about the failure-time distribution, there are various methods of survival analysis. Cleves et al. (2010) specify three approaches and relevant models as follows:

- Nonparametric models (Kaplan–Meier and Nelson–Aalen);
- semiparametric models (Cox proportional hazards model); and
- parametric models (e.g. exponential, Weibull, lognormal, log-logistic, gamma and Gompertz).

While parametric models require assumptions about the distribution of failure times, semiparametric models are parametric in the sense that the effect of the covariates is assumed to take a certain form. In this case, no parametric form of the survival function is specified, yet the effects of covariates are parametrized to modify the baseline survivor function. Thus, compared with the previous approaches, nonparametric models do not require any assumptions about the distribution of failure times.

Nonparametric methods estimate the probability of survival past a certain time or compare survival experiences for different groups (Cleves et al., 2010). The common characteristic of nonparametric models is that they do not make any assumptions about the distribution of failure times or the way in which covariates change the survival experience. The Kaplan–Meier estimator of the survivorship function at time *t* can take the form of the following equation:

$$\hat{S}(t) = \prod_{t_{(i)} \leq t} \frac{n_i - d_i}{n_i}, \tag{8}$$

where n_i is the number at risk of dying (company failure) at $t_{(i)}$, d_i refers to the observed number of failures and $\hat{S}(t) = 1$ if $t < t_{(i)}$. If we assume that the time variable is absolutely continuous, then the survival function may be expressed as

$$S(t) = e^{-H(t)}, \tag{9}$$

where *H(t)*, the cumulative hazard function, can be written as

$$H(t) = -\ln(S(t)). \tag{10}$$

Aalen, Nelson and Altshuler propose the indicator *H(t)*, which is referred to as the Nelson–Aalen estimator (Hosmer et al., 2008). The Nelson–Aalen estimator of *H(t)* is given by

$$\hat{H}(t) = \sum_{t_{(i)} \leq t} \frac{d_i}{n_i}. \tag{11}$$

Semiparametric models do not assume any parametric form of the survival functions. The effects of the covariates are parametrized to modify the baseline survival function. According to Hosmer et al. (2008), the regression model for the hazard function can be expressed as

$$h(t,x,\beta)=h_0(t)r(x,\beta), \tag{12}$$

where $h_0(t)$ characterizes the change in the hazard function as a function of the survival time and $r(x,\beta)$ describes how the hazard function changes as a function of the subject covariates. The model makes no assumptions about the shape of the hazard over time. However, the general shape of the hazard is the same for everyone. One subject's hazard is a multiplicative replica of another's, which is constant. The quantities estimated from the model are hazard ratios, which measure the extent to which a covariate increases or decreases the rate of a particular event.

Parametric models are used when the distribution of the survival time has a known parametric form. They generally provide smooth estimates of the hazard and survival functions for any combination of covariate values. According to Hosmer et al. (2008), using these models may have the following advantages. Full maximum likelihood may be used to estimate the parameters, the estimated coefficients or their transformations can provide clinically meaningful estimates of effects, fitted values from the model can provide estimates of survival time and residuals can be computed as differences between observed and predicted values of the time.

2.2 Cox Proportional Hazard Model

The Cox proportional hazard model is a semiparametric model of survival analysis. The effect of the covariates is assumed to take a certain form compared with the nonparametric approach. In this case, no parametric form of the survival function is specified, yet the effects of the covariates are parametrized to modify the baseline survivor function. In general, the baseline survival function is the function for which all the covariates are equal to zero in a certain way. The hazard function (9) is the product of two parts, which characterizes how the hazard function changes as a function of the survival time, and the function $r=(x,\beta)$ describes how the hazard function changes as a function of the subject covariates.

It follows from the model that:

- The functions must be chosen such that $h(t,x,\beta)>0$,
- $h_0(t)$ is the hazard function when $r(x,\beta)=1$,
- $h_0(t)$ is referred to as the baseline hazard function when the function $r(x,\beta)$ is parametrized such that $r(x=0,\beta)=1$.

Thus, the baseline hazard function can be seen as a generalization of the intercept or constant term found in parametric regression models. We do not make any assumptions about $h_0(t)$, however, at the cost of a loss of efficiency. Although the model makes no assumptions about the shape of the hazard over time, the general shape is assumed to be the same for everyone.

The ratio of the hazard functions for two subjects with covariate values denoted x_0 and x_1 is:

$$\begin{aligned} HR(t,x_1,x_0)&=\frac{h(t,x_1,\beta)}{h(t,x_0,\beta)},\text{or}\\ HR(t,x_1,x_0)&=\frac{h_0(t)r(x_1,\beta)}{h_0(t)r(x_0,\beta)}=\frac{r(x_1,\beta)}{r(x_0,\beta)}. \end{aligned} \tag{13}$$

As we can see in (13), the hazard ratio (HR) depends only on the function $r(x,\beta)$.

This model was originally proposed in 1972 by Cox, who suggested using $r(x,\beta)=\exp(x\beta)$ for practical reasons. Then, the hazard function can be expressed as:

$$h(t,x,\beta)=h_0(t)e^{x\beta}, \tag{14}$$

and the hazard ratio is

$$HR(t,x_1,x_0)=e^{x\beta(x_1-x_0)}. \tag{15}$$

This model is the most-used semiparametric model, called the Cox model, the Cox proportional hazards model or the proportional hazards model. The term proportional hazards (PHs) refers to the fact that the hazard functions are multiplicatively related; thus, their HR is constant over time (Hosmer et al., 2008, p. 70). In other words, we assume that the covariates multiplicatively shift the baseline hazard function. Then, one subject's hazard is a multiplicative replica of another's (Cleves et al., 2010). Besides the assumption of proportional hazards, other parametrizations can be used, for example additive models. These parametrization approaches are described in the relevant literature (Hosmer et al., 2008; Klein et al., 2014).

2.3 Multiple Failure-Time Data

Cleves (2000) describes multiple failure-time data as data in which any of two or more events (failures) occur for the same subject or from identical events occurring to related subjects. The typical feature is that the failure times are correlated within a cluster (subject or group), violating the independence of failure times assumption required in traditional survival analysis. As the author points out, failure events should be classified according to whether they have a natural order and recurrences of the same type of events. The events are supposed to be ordered when the second event cannot occur before the first event. On the contrary, unordered events can happen in any sequence.

There are more approaches to examining multiple failure-time data. The first method considers the time to the first event, ignoring additional failures. However, this means that we do not make use of all the available data. The second method is based on analysing the available data while accounting for the lack of independence of the failure times. Cleves (2000) suggests corresponding procedures for estimating these models using the Cox proportional hazard model. Under the proportional hazard assumption, the hazard function (14) of the *i*th cluster for the *k*th failure type is as follows:

$$h_k(t, Z_{ki}) = h_0(t)e^{Z_i, \beta}, \tag{16}$$

where Z_{ki} is a *p*-vector of possibly time-dependent covariates for the *i*th cluster to the *k*th failure type. While we presume in equation (16) that the baseline hazard function is equal for every failure type, the baseline hazard function is allowed to differ by failure type in the following formula:

$$h_k(t, Z_{ki}) = h_{0k}(t)e^{Z_i, \beta}. \tag{17}$$

As Cleves (2000) suggests, the maximum likelihood estimates for models (16) and (17) are obtained from Cox's partial likelihood function $L(\beta)$, assuming independence of failure times.

Concerning the analysis of multiple failure-time data, Cleves (2000) emphasizes the need to determine whether they are ordered or unordered data and select a suitable method for estimating models accordingly. In the case of unordered times, which is the case for rating analysis, it is first necessary to determine whether the events are of the same or different types. Similarly, a decision is required on whether the baseline hazard is the same or different for all event failures. In any case, it is necessary to implement the methods in which the data are correctly structured, including identifying individual failure events.

3. Modelling of Rating Downgrades

The aim of this study is to examine the relationship between time and corporate rating downgrades, including the influence of the variables used. In this study, the event, or failure in terms of survival analysis, is defined as a rating downgrade. As the rating can be downgraded more than once during the period under our observation, multiple failure-time analysis approaches should be used. Thus, survival analysis models are estimated using the Cox proportional hazard model for multiple failure-time data. Furthermore, we aim to model the rating downgrade depending on the time and the annual changes in financial variables. Therefore, we use yearly changes in the rating as the dependent variable and yearly changes in the financial covariates as the independent variables in our model. Specifically, the dependent variable is a rating downgrade. This choice is quite natural because it is crucial to detect a potential deterioration in investment quality, which increases the credit risk.

This study is focused on the analysis of corporate credit ratings from eight countries in Central and Eastern Europe (CEE): the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia. As global rating agencies do not rate many companies in these countries, the models are estimated based on the Multi-Objective Rating Evaluation (MORE).[1] The MORE methodology, as part of the issuance of credit ratings, complies with Regulation (EC) N. 1060/2009 of the European Parliament and the Council of 16 September 2009 (the Credit Rating Agencies Regulation). Therefore, with effect from 10 July 2015, it is registered as a credit rating agency in accordance with this regulation.[2] The data sample contains records of 1249 companies (2002–2007).

We follow rating assessments and various financial variables on an annual basis. Thus, we have a total of 7494 observations. Without specifications of related subjects based on id, the overall data consist of 6245 observations, 705 (single) and 870 (multiple) events defined as rating downgrades and 18,735 (single) and 15,358 (multiple) of the total analysis time. In both cases, we use 2,436 observations. Since we assume that the observations for each company may be correlated, we adjust for this by clustering as follows. The companies are interpreted as the same sampling units by using the id() identifier, allowing us to specify related subjects. Thus, we analyse the data based on 737 clusters.

[1] MORE ratings classify companies similarly as rating agencies (Bureau van Dijk Electronic Publishing, 2008). MORE ratings are calculated using a unique model that references the company's financial data to create an indication of the company's financial risk level.

[2] MORE rating by the first Fintech Rating Agency – S-Peek, online access: https://www.s-peek.com/en/more-rating (17 June 2019).

For each company in our data sample, we follow the rating assessment and various financial variables annually (Table 1-1). The independent covariates in the model are the annual percentage changes in the relevant financial indicators, referred to as covariates in the table. We consider all rating downgrades as the same event type and do not distinguish the downgrade size. Nevertheless, the rating changes by more than one degree in a marginal number of cases, so we do not take this fact into account.

Table 1–1 Description of financial variables

Financial indicator	Covariate	Mean
Total assets	tag	22.25
Return on assets	roag	86.70
Return on equity	roeg	66.24
EBITDA to total debt	ebitdarg	27.57
Equity to total assets	eqtag	12.33
Cash flow	cfg	55.62
Interest coverage	intcovg	149.13

In this study, we estimate survival models using two approaches. First, we examine the time to the first event, ignoring additional downgrades. Next, we make use of all the available data and account for multiple failure-type data. This procedure allows us to compare the two approaches and make recommendations.

3.1 Estimation of Rating Survival Models

The models will be developed based on the above-described approaches. Since we consider single and multiple failure-time data, we estimate two models, referred to as single and multiple. First, we perform a simple survival analysis, in which we consider only the first rating downgrade, ignoring additional ones for each company. Thus, the data consist of 256 defined events (downgrades) and 8253 total analysis time data. The time is measured in years after the time of origin, set as the year 2002.

The graph of the Kaplan–Meier estimate of the survival function is shown in Figure 1-1. Next, we estimate the survival model based on multiple failure-time event data. Using this procedure, we will consider any rating downgrades for the entity. Therefore, since numerous events can occur for each subject, it is a multiple failure-time analysis of the same type. The time until the event is measured as the time since the last event for each subject. The data consist of 331 downgrades and 6732 of the total analysis time data. The Kaplan–Meir estimate of the survival function for multiple failure-time data is shown in Figure 2-1. In both cases, the survival curve has a descending, stepped shape. By comparing Figures 1-1 and 2-1, we can see that the probability of survival, that is, a stable or improved rating, is higher in the case of single data. It is a consequence of the assumptions used in the survival analysis as, in this case, only one event is allowed for each subject.

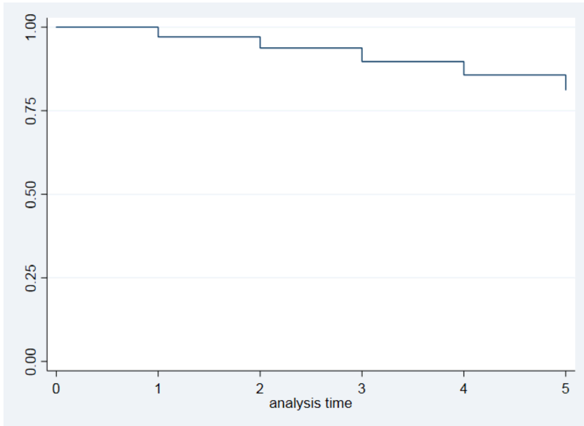


Figure 1–1 Kaplan–Meier survival estimate (single)

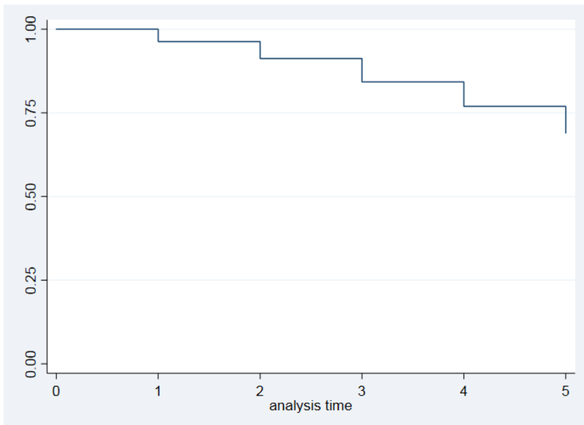


Figure 2–1 Kaplan–Meier survival estimate (multiple)

The resulting models are summarized in Table 2-1. We can see the estimated coefficients of the independent variables (coeff.), including the standard error, in the brackets. For completeness, the hazard ratios (HRs) are also listed in the table. If we compare the two models, we can see that the estimated coefficients do not differ significantly. Thus, the effect of the variables used on the probability of survival, that is, a stable or upgraded rating, is similar. All the coefficients are statistically significant at the 0.05 level in the multiple model.

Table 2–1 Estimated coefficients of the Cox models

Indep. variable	Coeff. Single	HR Single	Coeff. Multiple	HR Multiple
tag	0.0083* (0.002)	1.0083	0.0084* (0.002)	1.0084
roag	-0.0027* (0.001)	0.9973	-0.0034* (0.001)	0.9966
roeg	0.0030 (0.002)	1.0030	0.0038* (0.001)	1.0038
ebitdarg	-0.0079* (0.004)	0.9921	-0.0083* (0.003)	0.9918
eqtag	-0.0153* (0.005)	0.9848	-0.0184* (0.004)	0.9817
cfg	-0.0117* (0.003)	0.9884	-0.0142* (0.003)	0.9859
intcovg	-0.0062* (0.002)	0.9940	-.0063* (0.002)	0.9938

* Significant at the 0.05 level; standard errors adjusted for 737 clusters.

The estimated coefficients can be used to interpret the effect of individual variables, but it is more appropriate to use the hazard rates. For example, an increase in tag by one unit (the annual change in total assets by 1%) increases the hazard of a rating downgrade by 0.84%. Conversely, if the covariate eqtag increases by one unit, the hazard decreases by 1.83%. From the overall results, we can see that an annual percentage change of one unit in the variables roa, ebitda, eqta, cf and intcov reduces the hazard of the rating downgrade. In contrast, the hazard is increased by changes in the variables ta and roe.

To assess the influence of the variables on the hazard, we determine the hazard for the so-called average company (H1); that is, the values of the variables are equal to their mean values. Subsequently, we compare this hazard with the baseline hazard, when the values of all the variables are equal to zero (H0). The graphical representation is presented in Figures 3-1 and 4-1. In both graphs, we can see that H1 lies below the baseline hazard, H0. It follows from the fact that the baseline hazard corresponds to a situation in which all the covariates are zero. However, a zero annual change in the financial indicators means that the risk of a rating downgrade must be higher than in the average company's yearly changes.

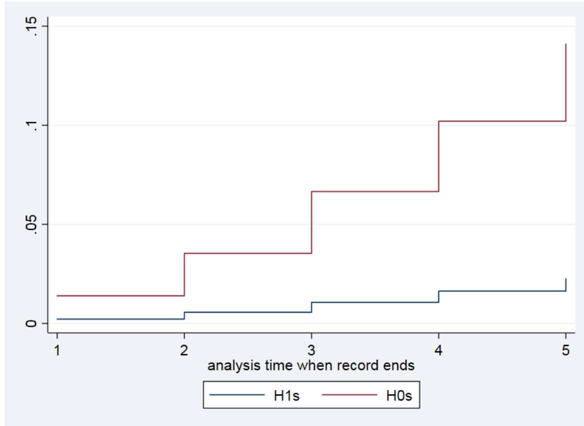


Figure 3–1 Baseline and average hazard (single)

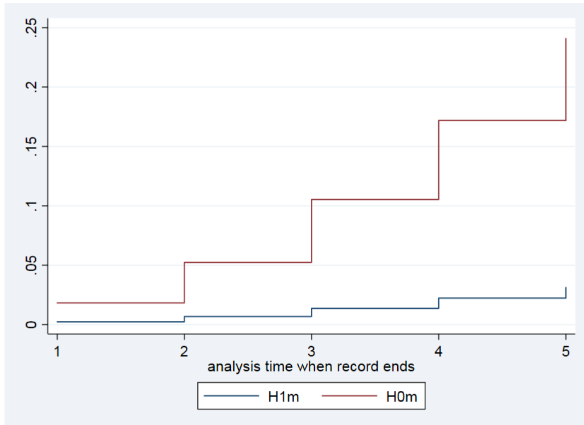


Figure 4–1 Baseline and average hazard (multiple)

3.2 Results and Interpretation

When using the Cox model, it is advisable to verify that the hazard functions are multiplicatively related. Therefore, we use the plot of the estimated hazard on a log scale using the kernel smoother to test the proportional hazards assumption. Since the lines in Figure 1 (Appendix) seem to be parallel, we conclude that the proportionality assumption in both models is not violated.

The test of the proportional hazards specification is based on the Schoenfeld residuals after fitting the model. It is used to test the independence between residuals and time. The test results in Table 3-1 suggest that the hazard assumption is not proportional for three variables: roag, roeg and intcovg. However, based on the global test, we find no evidence that our specification violates the proportional hazards assumption in both models.

Table 3–1 Test of PH assumptions

Indep. variable	**Single rho (Chi2)**	**Multiple rho (Chi2)**
tag	-0.0862 (1.53)	-0.0724 (1.58)
roag	-0.0935 (6.44) *	-0.0800 (4.95) *
roeg	0.0954 (6.48) *	0.0836 (5.46) *
ebitdarg	-0.0148 (0.13)	-0.0087 (0.05)
eqtag	0.0873 (2.65)	0.0556 (1.46)
cfg	0.0817 (2.57)	0.0803 (3.34)
intcovg	-0.0704 (7.23) *	-0.0683 (7.16) *
Global	(11.64)	(10.59)

* Significant at the 0.05 level; standard errors adjusted for 737 clusters.

The explained variation in the estimated models is measured using the adjusted index of determination R^2 (Royston, 2006). The value of R^2 equals 0.4689 (SE = 0.104) for the single model and 0.4973 (SE = 0.0231) for the multiple model. The greatest contribution to the explained variation is made by the covariates intcovg, cfg, eqtag and tag based on a comparison of the models with different numbers of predictors. Detailed results are presented in the Appendix (Table 1), which contains the data for models with a decreasing number of covariates. The last row includes a model containing only one variable, intcovg.

The overall model fit is evaluated using Cox–Snell residuals. Figures 5-1 and 6-1 show the Nelson–Aalen cumulative hazard estimator plots for the Cox–Snell residuals for both models. We can see some variability around the 45° line, particularly in the right-hand tail. Cleves et al. (2010) argue that some variability is expected due to the reduced effective sample caused by prior failures and censoring. Comparing Figures 7-1 and 8-1, we can conclude that the multiple model fits the data better than the single model.

Overall, the multiple failure-time data analysis leads to a more suitable model based on the statistical significance of the estimated coefficients and goodness of fit. On the other hand, it should be noted that the two survival models are very similar based on the estimated coefficients and explained variation. However, with respect to the validation and interpretation of both models, we should use a multiple failure-time analysis.

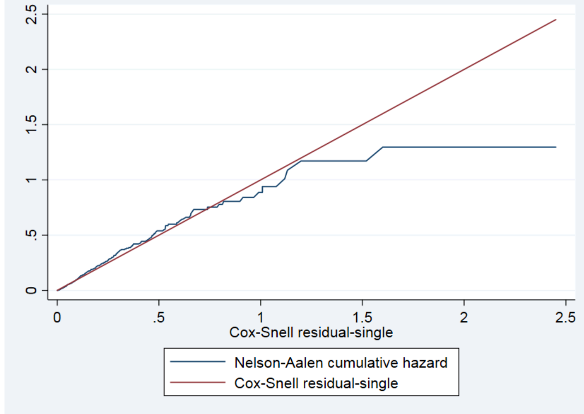


Figure 5-1 Cumulative hazard of Cox–Snell residuals (single)

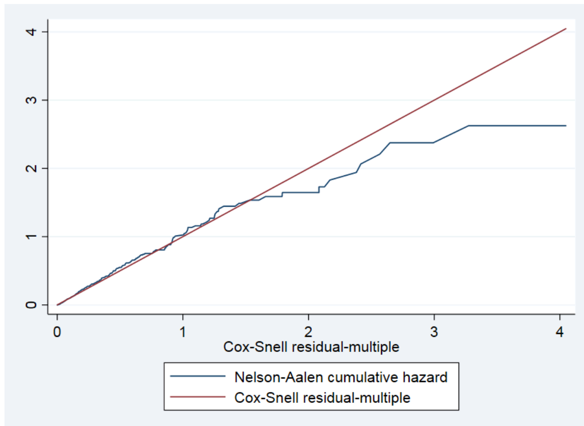


Figure 6-1 Cumulative hazard of Cox–Snell residuals (multiple)

4. Conclusion

This study aimed to develop rating models using survival analysis methods. Specifically, we applied the Cox proportional hazards model to analyse the survival time until the event. In our case, we focused on using survival analysis to model the time to a rating downgrade. As part of the analysis, we examined the effect of financial variables on the probability of a negative annual rating change.

Two different approaches were used to estimate the models, depending on whether we were considering only one or more events, defined as rating downgrades, for one company. The single model was derived on the assumption that the event can occur only once for each subject. The multiple model, on the other hand, assumed that the event can occur repeatedly. Due to these different assumptions, the input data and their structure also had to be adjusted.

The resulting models were presented based on the estimated coefficients for the variables used in the analysis. Both models were statistically significant, as were the estimated coefficients of the individual variables in the multiple model. Based on the overall results, we concluded that the annual changes in the financial variables used affect the probability of the rating downgrade. We used the baseline hazard and the hazard of the so-called average company to interpret the models based on the mean values of the variables. The fit of both models was assessed using Cox residuals. Based on the main findings of this study, we concluded that the multiple model approach is more suitable for events that might occur repeatedly. The simple model should be used to examine the survival time until the first event for any reason. In other cases, we should use multiple failure-time data analysis, which makes better use of the available data.

The overall findings of this work show that survival analysis is, in addition to typical financial problems, such as the analysis of survival to bankruptcy or default, also suitable for other types of tasks. However, when applying it, it is necessary to consider the specific data structure, especially whether the event can occur repeatedly for one subject or whether more events can occur for a given subject. In these cases, it is appropriate to use multiple failure-time analysis, which corresponds better to the problem.

References

ALTMAN, E.I., KATZ, S. (1976). Statistical bond rating classification using financial and accounting data. M. Schiff and G. Sorter, eds.: *Topical Research in Accounting*, NYU Press, New York.

ALTMAN E.I., SABATO G., WILSON N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *The Journal of Credit Risk*, 6: 95–127. https://doi.org/10.21314/JCR.2010.110

ANG J. S., PATEL K. (1975). Bond rating methods: Comparison and validation. *The Journal of Finance*, 30: 631–640. https://doi.org/10.1111/j.1540-6261.1975.tb01836.x

BELKAOUI, A. (1980). Industrial bond ratings: A discriminant analysis approach. *Working paper (University of Ottawa)*. https://doi.org/10.2307/3664892

CLEVES, M. (2000). Analysis of Multiple Failure-time Data with Stata. *Stata Technical Bulletin*.

CLEVES, M. et al. (2010). *An Introduction to Survival Analysis Using Stata*. College Station: Stata Press.

HSU, F. J., CHEN, M. Y., CHEN, Y. C. (2018). The human-like intelligence with bio-inspired computing approach for credit ratings prediction. *Neurocomputing*, 279: 11–18. https://doi.org/10.1016/j.neucom.2016.11.102

DUTTA, S., SHEKHAR, S. (1988). Bond ratings: A non-conservative application of neural network. *IEEE International Conference on Networks*, 2, San Diego, California, 443–450. https://doi.org/10.1109/ICNN.1988.23958

FISHER, L. (1959). Determinants of Risk Premiums on Corporate Bonds. *Journal of Political Economy*, 67: 217–237. https://doi.org/10.1086/258172

FIGLEWSKI, S., FRYDMAN, H., LIANG, W. (2012). Modelling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics & Finance*, 21: 87–105. https://doi.org/10.1016/j.iref.2011.05.004

GLENNON, D., NIGRO, P. (2005). Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach. *Journal of Money, Credit and Banking*, 37: 923–947. https://doi.org/10.1353/mcb.2005.0051

GOLBAYANI, P. FLORESCU, I., CHATTERJEE, R. (2020). A comparative study of forecasting corporate credit ratings using neural networks, support machines, and decision trees. *North American Journal of Economics and Finance*, 54: 1–16. https://doi.org/10.1016/j.najef.2020.101251

GOURIEROUX, CH., JASIAK, J. (2007). *Econometrics of Individual Risk: Credit, Insurance and Marketing*. Princeton: Princeton University Press. https://doi.org/10.1515/9781400829415

HARRELL, F. E. (2010). *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis.* New York: Springer.

HICKMAN, W. B. (1958). *Corporate bond quality and investor experience*. Princeton University Press: Princeton.

HOSMER, D. W., LEMESHOW, S., MAY, S. (2008). *Applied survival analysis: regression modeling of time-to-event data.* Hoboken: Wiley-Interscience. https://doi.org/10.1002/9780470258019

JONES, S., JOHNSTONE, D., WILSON, R. (2015). An empirical evaluation of the performance of binary classifiers in the prediction of credit rating changes. *Journal of Banking & Finance*, 56: 72–85. https://doi.org/10.1016/j.jbankfin.2015.02.006

KAPLAN, R., URWITZ, G. (1979). Statistical models of bond ratings: A methodological inquiry. *Journal of Business*, 52: 231–261. https://doi.org/10.1086/296045

KLEIN, J. P. et al. eds. (2014). *Handbook of survival analysis.* Boca Raton: CRC Press, Taylor & Francis Group.

LEOW, M., CROOK, J. (2014). Intensity models and transition probabilities for credit card loan delinquencies. *European Journal of Operational Research*, 236: 685–694. https://doi.org/10.1016/j.ejor.2013.12.026

LOUIS, P., VAN LAERE, E., BAESENS, B. (2013). Understanding and predicting bank rating transitions using optimal survival analysis models. *Economics Letters*, 119: 280–283. https://doi.org/10.1016/j.econlet.2013.02.033

PARNES, D. (2007). Time series patterns in credit ratings. *Finance Research Letters*, 4: 217–226. https://doi.org/10.1016/j.frl.2007.09.001

PINCHES, G., MINGO, K. A. (1973). A multivariate analysis of industrial bond ratings. *The Journal of Finance*, 28: 1–18. https://doi.org/10.1111/j.1540-6261.1973.tb01341.x

ROA, M., GARCÍA, A. F., BONILLA, L. (2009). Country Risk Ratings and Financial Crisis 1995 – 2001: A Survival Analysis. *Review of Business*, 30: 33–45.

ROYSTON, P. (2006). Explained variation for survival models. *The Stata Journal* 6(1), pp. 83-96. https://doi.org/10.1177/1536867X0600600105

ROYSTON, P., LAMBERT, P. C. (2011). *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*. College Station: Stata Press.

SURKAN, A. J., SINGLETON, J. C. (1990). Neural networks for bond rating improved by multiple hidden layers. *Proceedings of the International Joint Conference on Neural Networks*, 2, San Diego, California, 157–162. https://doi.org/10.1109/IJCNN.1990.137709

TABACHNIK, B. G., FIDELL, L. S. (2007). *Using Multivariate Statistics*. Boston: Pearson Allyn & Bacon.

WAAGEPETERSEN, R. (2010). A statistical modelling approach to building an expert credit rating system. *The Journal of Credit Risk*, 6: 81–94. https://doi.org/10.21314/JCR.2010.107

WANG, M., KU, H. (2021). Utilizing historical data for corporate credit rating assessment. *Expert systems with applications*. 165: 1–12. https://doi.org/10.1016/j.eswa.2020.113925

WINGLER, T. R., WATTS, J. M. (1980). An analysis of bond rating changes in the electric utility industry. *The Financial Review*, 15: 1–89. https://doi.org/10.1111/j.1540-6288.1980.tb01601.x

ZHANG, J., THOMAS, L. C. (2012). Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling LGD. *International Journal of Forecasting*, 28: 204–215. https://doi.org/10.1016/j.ijforecast.2010.06.002

Appendix

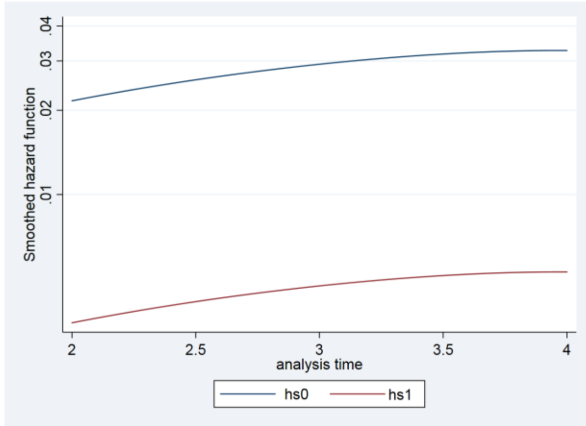
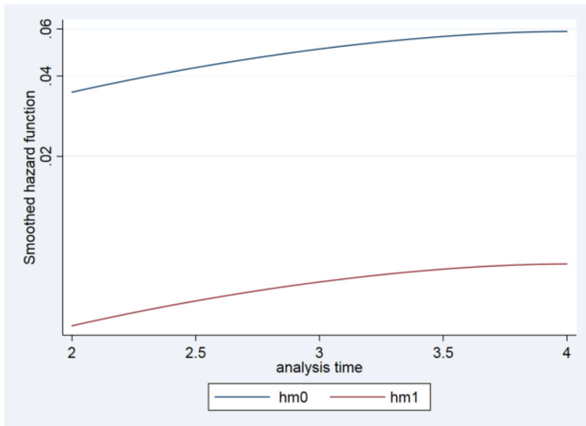


Figure 1: Smoothed hazard functions

(a) Single model



(b) Multiple model

Table 1: Explained variation

(a) Single model

Variables	R^2	St. error	Confidence interval	
All	0.4690	0.0270	0.4136	0.5193
-tag	0.4485	0.0275	0.3923	0.5000
-roag	0.4483	0.0275	0.3922	0.4999
-roeg	0.4418	0.0280	0.3846	0.3922
-ebitdarg	0.4471	0.0227	0.4010	0.4900
-eqtag	0.4200	0.0231	0.3734	0.4638
-cfg	0.3864	0.0242	0.3377	0.4327

(b) Multiple model

Variables	R^2	St. error	Confidence interval	
All	0.4973	0.0231	0.4500	0.5406
-tag	0.4814	0.0235	0.4333	0.5256
-roag	0.4814	0.0235	0.4334	0.5255
-roeg	0.4664	0.0240	0.4174	0.5115
-ebitdarg	0.4530	0.0202	0.4122	0.4912
-eqtag	0.4241	0.0206	0.3825	0.4633
-cfg	0.3953	0.0219	0.3514	0.4370