Advanced Mathematical Models & Applications
Vol.9, No.1, 2024, pp.14-26
https://doi.org/10.62476/amma9114

SURVIVAL ANALYSIS MODELING ON CREDIT RISK WITH LATENT VARIABLE INDICATOR MODELS

Adfi Bio Fimba*, **Solimun**, **Adji Achmad Rinaldo Fernandes**

Departement of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University, Malang, Indonesia

Abstract. Survival analysis has several advantages, including the ability to handle incomplete data. In general, the explanatory variables used in survival analysis are manifest variables. In other fields such as social sciences, variables such as perception, attitudes, and psychology are often involved in statistical analysis. This research aims to model cox proportional hazard regression on data with latent variable types. Research was conducted in the banking sector using Likert scale questionnaire data. It will examine how the assessment of the 5C variable in the form of a latent variable relates to the time of delay in credit payments. Confirmatory Factor Analysis (CFA) was used to form a reflectively latent variable indicator model which was then formed into a cox proportional hazards regression model. The results show that all 5C variable assessments influence the speed of credit payments. The novelty in this research lies in the use of indicator analysis models to form latent variables before undergoing survival. The resulting interpretation and conclusions are expected to provide more in-depth information because they include the relationship between survival time and indicators.

Keywords: Credit colletability, factor analysis, indicator models latent variables, measurement models, survival analysis.

AMS Subject Classification: 62H25, 62N02, 62P25.

Corresponding author: Adfi Bio, Fimba, Departement of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University, Veteran Street, Malang 65145, Indonesia, e-mail: *adfibiofimba@student.ub.ac.id*

Received: 13 November 2023; Revised: 11 December 2023; Accepted: 10 January 2024;
Published: 22 April 2024.

1 Introduction

Survival analysis is a statistical technique used to examine the time required for a specific event to occur. Survival analysis has several advantages, including its ability to handle incomplete data. It can integrate censored data into calculations and survival time estimations. To handle informative censoring, sensitivity analyses, such as best-case and worst-case scenarios, can be used to try to quantify the effect that informative censoring has on the analysis (Liu, 2017; Campigotto & Weller, 2014). It is important to check for informative censoring and use appropriate statistical methods to handle it (Lin et al., 2023). The elements of randomness in survival models are not expressed through random variables as in linear regression models (Hougaard, 1995; Šiaulys & Puišys, 2022). The random element in the Cox Proportional Hazard model lies in censored survival data, which lack complete information about an individual. Additionally, the element of randomness exists in the method of parameter estimation since the estimated

How to cite (APA): Fimba, A.B., Solimun & Fernandes, A.A.R. (2024). Survival analysis modeling on credit risk with latent variable indicator models. *Advanced Mathematical Models & Applications*, *9*(1), 14-26 https://doi.org/10.62476/amma9114

parameter values are random and depend on the given data.

The field of health is one of the primary applications of survival analysis. The analysis plays a crucial role in medical and epidemiological research because it helps researchers understand the factors influencing survival time, the risk of specific events, and the effectiveness of preventive actions or medical interventions (Ferreira & Patino, 2016). Covariate variables inherent to patients, such as age, gender, smoking status, disease severity, Body Mass Index, and Treatment Status, are often studied to analyze their relationship with survival time (Zhang et al., 2018; Austin et al., 2020). Therefore, in general, the explanatory variables used in survival analysis are manifest variables.

In other fields such as the social sciences, variables like perceptions, attitudes, and psychology are often involved in statistical analysis. These latent variables are relevant and considered important because they reflect the understanding, views, preferences, and perceptions of individuals or groups on various issues and topics that affect social and economic contexts. In psychology research, it is often necessary to measure attitudes, opinions, or perceptions, which are examples of latent variables (Perron & Gillespie, 2015; Lesnoff et al., 2021). In order to measure these variables, researchers use a variety of methods, including surveys, interviews, and experiments. Latent variables often represent abstract or complex concepts that cannot be directly measured. Latent variables are unobservable and can only be indicated or measured through a set of related observable variables (Rutledge et al., 2021; Lesnoff et al., 2021).

One of the primary uses of survival analysis in the banking industry is to assess credit risk (Dirick et al., 2017; Marletta & Nuovo, 2021). It is of utmost importance for banks to anticipate when credit payments might become overdue and to pinpoint customers with a higher risk of encountering credit payment delinquencies. Credit delinquency remains an ongoing concern in the banking sector, prompting banks to employ a variety of strategies and protocols to mitigate this risk. Survival analysis can be used to analyze the relationship between access to bank financing and start-up resilience (Castaldo et al., 2023). Studies in this domain may encompass latent variables, such as those associated with the psychological and behavioral aspects of customers. Different methods for establishing latent variables are available, including mean scores, total scores, and latent variable indicator models. The most notable distinction among these approaches resides in the assignment of weights to the indicators when forming latent variables. The selection of latent variable measurement methods can lead to varying outcomes in the resulting survival analysis models.

One of the commonly employed methods for assessing the credibility of prospective customers is through the evaluation of the 5C variables. The 5C approach is a widely utilized method for assessing the credibility of prospective customers in the banking sector. This approach involves assessing five variables, namely character, capacity, condition, collateral, and capital. The evaluation of the 5C variables is a procedure that banks can undertake to assess the credibility of prospective customers to ensure their accountability for credit repayment obligations.

Survival analysis as models capable of forecasting loan-related data. In line with this approach, parametric (Weibull) and semi-parametric models (Cox) have been employed in credit-related research (Jiang et al., 2019; Suárez et al., 2021). The focus of this survival analysis pertains to the credit mortality rate, representing the likelihood of a company facing bankruptcy and, consequently, the probability that the company will be unable to fulfill its loan obligations (Tan & Anchor,, 2016; Tan & Floros, 2018).

In this study, efforts were made to integrate the cox proportional hazard model with the latent variable indicator model to analyze the relationship between the 5C variable and credit risk. The variables studied tend to be reflective, so reflective indicator models are used. Previously, the background of this research has been explained from both statistical and non-statistical perspectives. The theory used has been outlined in the literature review. The data and research methods are elaborated in the research method section. The results and comprehensive discussion are detailed in the results and discussion section. The uniqueness (novelty) in this study lies in an integrative approach that tries to combine the Cox proportional hazard model with the latent variable indicator model to investigate the relationship between the 5C variable and credit risk. The emphasis on the use of reflective indicator models by the reflective properties of the variables studied adds a dimension of novelty to this study. Through modeling results that integrate survival analysis and latent variable indicator models, it is expected that a more significant contribution can be made to bank credit risk management, especially in the context of house ownership loans.

2 Literature Review

2.1 Survival Analysis

Survival analysis is a statistical method that focuses on studying the time until a specific event takes place. Kleinbaum & Klein (2005) defines it as a statistical approach designed to analyze data where the primary variable of interest is the duration until an event occurs. In this type of analysis, the central focus revolves around the concept of time. Survival time, which indicates how long an individual remains unaffected by a particular event, can be explained by various factors that influence it. When dealing with survival data, it is common to encounter incomplete subject information, primarily because some subjects have not experienced the event of interest (failure time) by the end of the study period. To address this issue, survival analysis incorporates specific considerations for handling incomplete data, a concept often referred to as data censoring (Katzman al., 2018). Data censoring arises from the inability to precisely determine the exact survival time of a subject, but available information can provide valuable insights into the subject's survival time (Kleinbaum & Klein, 2005). The survival function, denoted as $S(t)$, represents the probability that a subject will continue to survive (without experiencing the event) up to or beyond a specific time point, as defined by Lawless (2002). The survival function can be formulated as shown in equation (1)

$$S(t) = P(T \geq t) = \int_t^\infty f(x)dx \tag{1}$$

The hazard function, often referred to as the hazard rate, signifies the rate at which a subject faces the risk of death or an immediate event occurrence at time t, as described by Katzman al. (2018) and Kleinbaum & Klein (2005). It characterizes pessimism as it outlines the subject's likelihood of encountering an event at a specific moment in time.

The formulation of the hazard function can be expressed through equation (2) as stated by Lawless (2002).

$$h(t) = \lim_{\Delta t \to 0}[\frac{P(t \leq T < (t + \delta t)|T \geq t)}{\Delta t}] \tag{2}$$

2.2 Cox Proportional Hazard Model

ThecCox proportional hazard model is a type of semiparametric approach commonly used in survival analysis. This model enables us to assess the impact of a predictor variable on the duration until an event takes place (survival time). Nevertheless, a notable limitation of this model is its inability to precisely predict the functional form of the hazard rate based on the observed survival time, as noted by Kleinbaum & Klein (2005) and Rabe-Hesketh & Skrondal (2008). The mathematical expression for the Cox proportional hazard model can be found in equation (3).

$$h(t, x) = h_0(t)exp(\beta_1 x_1 + ... + \beta_p x_p) \tag{3}$$

The Hazard Ratio serves as a means to compare the risk of an event occurring in one individual with that of another. Stensrud & Hernán (2020) describe the Hazard Ratio as an indicator of the extent to which a predictor variable impacts the observed time. Utilizing the Hazard Ratio allows for a straightforward and comprehensible interpretation of the Cox regression coefficient. Equation (4) can be used to calculate the Hazard Ratio.

$$
\begin{aligned}
HR &= \frac{h(t,x^*)}{h(t,x)} \\
HR &= \frac{h_0(t)\exp(\beta_1 x_1^* + \ldots + \beta_p x_p^*)}{h_0(t)\exp(\beta_1 x_1 + \ldots + \beta_p x_p)} \\
HR &= \exp(\beta_1(x_1^* = x_1) + \ldots + \beta_p(x_p^* - x_p)) \\
HR &= \exp\left(\sum_{j=1}^{p} \beta_j (x_j^* - x_j)\right)
\end{aligned} \tag{4}
$$

2.3 Propotional Hazard Assumption

As mentioned by Collet (2003), the proportional hazard assumption can be assessed through various methods and tests, such as visual inspection and evaluating goodness of fit. In addition to the graphical approach for checking proportional hazard (PH) assumptions, there is also a numerical method available, which involves performing a global test, as explained by Kuitunen al. (2021). This global test involves examining the correlation between Schoenfeld residuals and the rank of survival time, calculated using the following correlation formula.

$$
r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{5}
$$

Where:
x : the rank of survival time
y : schoenfeld residuals of predictor variable
t : the i-th individual who experienced the incident

The hypothesis used in testing this correlation is as follows.
$H_0 : \rho = 0$ (Assuming PH is met)
$H_0 : \rho \neq 0$ (PH assumption is not met)
Reject H_0 when the p-value is less than the significance level α.

2.4 Selection the Best Model

In this phase, the objective is to choose and establish the superior model, either the AFT model or the PH models, based on the AIC values. A lower AIC value signifies a higher level of goodness-of-fit for the model (Pham, 2019).

2.5 Latent Variable Indicator Model

Measuring latent variables typically entails employing tools like questionnaires to gather data for each indicator or item present in the questionnaire. These indicators or items are commonly termed manifest variables or observable variables due to their capacity for direct observation or measurement, as explained by Solimun et al. (2017). Multiple techniques are available for computing latent variables in data analysis, encompassing mean scores, sum scores, scoring, reflective indicator models, and formative indicator models. When it comes to collecting data from latent or unobservable variables, two widely utilized techniques involve factor analysis and principal component analysis. Using this approach, the outcome provides factor scores and principal component scores, which serve as data representing latent variables. The constituents of these variables may differ; some of them reflect the latent variable, while others contribute to the construction of the variable, known as reflective and formative indicator models, respectively (Solimun et al., 2017). The decision between reflective and formative indicator models is of paramount importance as it has diverse implications for statistical analysis and the interpretation of research outcomes.

2.6 Reflectively Indicator Model

The reflective indicator model is a measurement approach applied in social and behavioral research to evaluate latent variables or specific constructs using indicators or items that are believed to represent these latent variables (Edwards, 2011). Within the reflective indicator model, the variability in latent variable values is presumed to be a function of the true score (indicator) plus an element of error (Solimun et al., 2017). In simpler terms, these indicators are seen as reflecting the fluctuations in the latent variable, rather than the opposite relationship. Data for latent variables using the reflective indicator model are frequently gathered through confirmatory factor analysis (Solimun et al., 2017; Hanafiah, 2020).

Factor analysis is a procedure that seeks to uncover associations among multiple variables that are mutually independent, thereby reducing the number of variables into one or more sets of variables that are fewer in quantity than the original variables. Factor analysis is a statistical technique extensively employed in various research domains, including the social sciences, economics, education, management, and psychology.

As an illustration, consider a normally distributed variable $X_1, X_2, ..., X_p$ characterized by a mean vector of μ and a variance-covariance matrix of Σ. This sets the stage for creating a model as outlined in equation (6).

$$\begin{aligned} X_1 &= c_(11)F_1 + c_(12)F2 + ... + c_(1p)Fp + \varepsilon_1 \\ X_1 &= c_(21)F_1 + c_(22)F2 + ... + c_(2p)Fp + \varepsilon_2 \\ &\vdots \\ X_1 &= c_(p1)F_1 + c_(p2)F2 + ... + c_(pp)Fp + \varepsilon_p \end{aligned} \tag{6}$$

When the input data matrix is labeled as S, the factor scores can be determined by applying the formula provided in equation (7).

$$\mathbf{S} - \mathbf{Fa} - \mathbf{c}'\mathbf{S^{-1}}(\mathbf{x_j} - \bar{\mathbf{x}}) \tag{7}$$

In the case where the input data matrix is denoted as R, the factor scores can be derived in accordance with the details presented in equation (8).

$$\mathbf{S} - \mathbf{Fa} - \mathbf{c}'\mathbf{R^{-1}}\mathbf{Z_j} \tag{8}$$

3 Research Method

The research data includes secondary data and simulation data. Secondary data were obtained through a research grant conducted by Fernandes et al. (2022) regarding customer assessments of the character, capacity, capital, collateral, condition, credit collectibility, and the Loan Payment Delay Time of each customer at Bank X. The research subjects consist of 1000 mortgage loan customers. The dependent variable in this study is the Loan Payment Delay Time for mortgage loans, measured in days. Indicator variables are measured using a Likert scale. The independent variable in this research is the assessment of the 5C variables. The variables and their indicators in this study can be detailed as shown in Table 1.

Table 1: Research Variable

Predictor Variable	**Indicator**	**Scale**
Censorship Status (Credit Collectibility Status) (d)	1 = Good Credit 0 = Bad Credit	Categorical
Credit Payment Delay Time (t)	Time	Continuous
Character (X_1)	Faith and responsibility ($X_{1.1}$) Nature or character/lifestyle ($X_{1.2}$) Payment Commitment ($X_{1.3}$)	Likert Likert Likert
Capacity (X_2)	Customer income ($X_{2.1}$) Ability to pay installment ($X_{2.2}$) Ability to complete credit on time ($X_{2.3}$)	Likert Likert Likert
Capital (X_3)	Permanent source of income ($X_{3.1}$) Have other business fields as a source of income ($X_{3.2}$) Have savings or deposits in the bank ($X_{3.3}$)	Likert Likert Likert
Collateral (X_4)	The selling value of the collateral that is pledged as collateral commensurate with/exceeding the credit ceiling ($X_{4.1}$) Collateral is physical or non-physical ($X_{4.2}$) Ownership of collateral and authenticity of documents ($X_{4.3}$)	Likert Likert Likert
Condition of Economy (X_5)	Business/enterprise/investment development ($X_{5.1}$) Economic fluctuations ($X_{5.2}$) Socio-economic conditions/family problems ($X_{5.3}$)	Likert Likert Likert

The model in this research can be seen in Figure 1.

Calculation of latent type predictor variables using Factor Analysis. The indicator model formed is reflective. The reflective indicator model means that the variable value is reflected by the indicator value. The results of factor score calculations are used in Cox proportional hazard modeling to determine the role of each predictor variable on credit repayment time. Data analysis in this research was carried out using R Studio software. The data analysis steps are described as follows.

1. Determine the input matrix for factor analysis
2. Obtain indicator loadings by carrying out factor analysis using the PCA method
3. Calculate factor scores for all indicators
4. Form an initial model and test the Cox proportional hazard assumption
5. If the assumptions have been met, then the coefficient significance test is carried out
6. Calculate and interpret the hazard ratio
7. Detailed and complete discussion of the modeling results

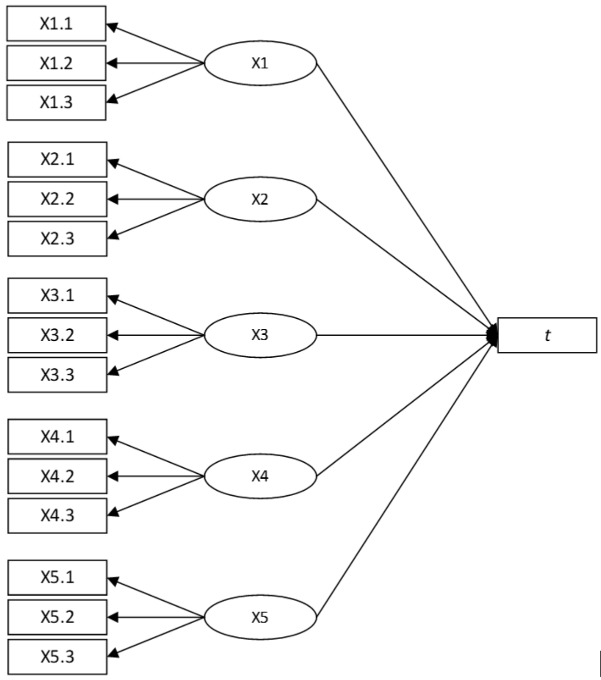

Figure 1: Research Model

4 Result and Discussion

4.1 Reflectively Indicator Model

The latent variables involved are reflective in nature, meaning they reflect or represent concepts that cannot be directly measured. Therefore, factor analysis is used to create an indicator model that can depict these latent variables.

Table 2: Factor Analysis Result

Variable	Estimated Factor Loading		Rotated Estimated Factor Loading		Communalities
	F_1	F_2	F_1	F_2	
Character (X_1)	0.78	-0.29	0.81	0.21	0.89
	0.76	0,65	0.25	0.97	1.00
	0.78	-0,35	0.84	0.16	0.73
Capacity (X_2)	0.76	0.64	0.81	0.35	0.89
	0.78	-0,33	0.81	0.21	1.00
	0.78	-0,30	0.88	0.16	0.79
Capital (X_3)	0.76	-0.54	0.93	0.12	0.88
	0.78	-0,03	0.60	0.51	0.62
	0.76	0.58	0.88	0.94	0.92
Collateral (X_4)	0.75	0.64	0.22	0.96	0.98
	0.77	-0,43	0.88	0.11	0.78
	0.79	-0.19	0.75	0.31	0.66
Condition (X_5)	0.76	-0.57	0.94	0.13	0.89
	0.76	-0.57	0.14	0.94	0.90
	0.79	0.00	0.56	0.55	0.62

The rotational $F1$ loading factor is an indicator model in this study. Based on the table, we can see that the character variable, the Intention and responsibility indicator ($X_1.1$) has an outer loading of 0.840. This shows that indicators $X_{1.1}$ play the most role in reflecting character variables. In the capacity variable, the indicator of the customer's ability to pay on time $X_2.3$ has an outer loading of 0.840. This shows that indicators $X_{2.3}$ play the most role in reflecting capacity variables. In the capital variable, the indicator of a fixed source of income $X_{3.1}$ which has an outer loading of 0.930. This shows that indicators $X_{3.1}$ play the most role in reflecting capital variables. In collateral variables, collateral indicators are physical or non-physical $X_{4.2}$ which have an outer loading of 0.880. This shows that indicators $X_{4.2}$ play the most role in reflecting collateral variables. In the variable condition of economy, the development indicator $X_{5.1}$ that has an outer loading of 0.940. This shows that indicators $X_{5.1}$ play the greatest role in reflecting the condition of economy variables.

4.2 Proportional Hazard Assumption Test

Cox proportional hazard modeling is built upon the underlying assumption of proportional hazards, which indicates that the ratio of individuals in various groups remains constant over time and remains unaffected by the passage of time. The results of the assessment of the Cox proportional hazard assumption using the Global Test are presented in Table 3. Table 3. provides a detailed account of these outcomes.

Table 3: Proportional Hazard Assumption Test Result

Variable	Test Statistics	p-value
Character (X_1)	0.23	0.63
Capacity (X_2)	1.36	0.24
Capital (X_3)	1.41	0.23
Collateral (X_4)	1.75	0.36
Condition of Economy (X_5)	1.73	0.19

Testing the Cox proportional hazard assumption uses the following hypothesis test.
$H_0 : \rho = 0$ (Assuming PH is met)
$H_1 : \rho \neq 0$ (PH assumption not met)
According to the information presented in Table 3, it is observed that all predictor variables exhibit a p-value greater than 0.05. Consequently, the decision made is to retain the null hypothesis (H_0). This leads to the conclusion that the hazard ratios for traits such as character, capacity, capital, and collateral are not influenced by the passage of time, or, in other words, the proportional hazard assumption holds true.

4.3 Cox Proportional Hazards Regression integrated with the reflective indicator model

Cox regression analysis is performed on the factor scores derived from factor analysis as a crucial step in the effort to comprehend how the factors identified in the factor analysis contribute to changes in the timing of credit payment delays.

Table 4: Results of Cox Proportional Hazards Regression Modeling

Variable	B	$exp(B)$	$se(B)$
Character (X_1)	-0.437	0.645	0.0353
Capacity (X_2)	-0.456	0.633	0.0362
Capital (X_3)	-0.484	0.616	0.0337
Collateral (X_4)	0.099	1.104	0.0367
Condition of Economy (X_5)	-0.125	0.882	0.0357

The results listed in Table 4 can be written into a mathematical model as follows.

$$h_i(t) = h_0(t)exp(-0.437X_1 - 0.456X_2 - 0.484X_3 + 0.099X_4 - 0.125X_5)$$

4.4 Hazard Ratio

The hazard ratio of the Cox proportional hazard model is described as follows.

4.4.1 Character (X_1)

$$HR = \frac{h_A}{h_B} = exp(\beta_1) = exp(-0.437) = 0.645$$

The hazard ratio for the character variable shows a value of 0.645, meaning that Iicreasing the customer's character value will have the potential to speed up the customer's credit repayment time 0.645 times faster.

4.4.2 Capacity (X_2)

$$HR = \frac{h_A}{h_B} = exp(\beta_2) = exp(-0.456) = 0.634$$

The hazard ratio for the capacity variable shows a value of 0.634, meaning that increasing the customer capacity value will have the potential to speed up customer credit payment times by 0.634 times faster.

4.4.3 Capital (X_3)

$$HR = \frac{h_A}{h_B} = exp(\beta_3) = exp(-0.484) = 0.616$$

The hazard ratio for the capital variable shows a value of 0.616, meaning that increasing the customer capital value will have the potential to speed up customer credit payment times by 0.616 times faster.

4.4.4 Collateral (X_4)

$$HR = \frac{h_A}{h_B} = exp(\beta_4) = exp(0.099) = 1.104$$

The hazard ratio in the collateral variable shows a value of 1.104, meaning that an increase in the value of customer collateral will slow down the credit payment time to 1.104 times longer.

4.4.5 Condition of Economy (X_5)

$$HR = \frac{h_A}{h_B} = exp(\beta_5) = exp(-0.125) = 0.882$$

The hazard ratio for the condition of economy variable shows a value of 0.882, meaning that increasing the customer condition of economy value will have the potential to speed up customer credit payment times by 0.882 times faster.

4.5 Discussion

Based on an analysis using survival on latent variable data, it can be concluded that the assessment variable 5C can affect the customer credit payment delay time at Bank X. The following discusses the influence of each 5C variable on the customer credit payment time.

Based on calculations, it is obtained that, an increase in the value of 1 unit for character variables can reduce a person's risk 0.645 times to be late, in other words not to be late. Based on the calculation of the indicator model, character variables are reflected most strongly by the item $X_1.1$, namely the will and responsibility of a person. Thus, to lower one's risk of being late can focus on assessing one's intentions and responsibilities.

Customer character refers to the behavior and track record of an individual or company regarding financial obligations, including their debt repayment history and overall financial responsibility. Customers with good character tend to be more compliant with credit agreements and agreed-upon terms. They are more likely to respect payment deadlines and promised payment amounts. Customer character also reflects an individual's ability to manage their finances. People who are diligent in managing their finances and have stable income are more likely to avoid payment delays.

Based on calculations, it is obtained that, an increase in the value of 1 unit for the capacity variable can reduce a person's risk 0.634 times to be late, in other words not to be late. Based on the calculation of the indicator model, the capacity variable is reflected most strongly by the item $X_2.3$, namely a person's ability to complete credit on time. Thus, to lower one's risk of being late can focus on assessing one's ability to pay credit on time.

Customer capacity affects credit payment delay time because customers' financial capabilities and financial management skills directly impact their ability to meet credit payment obligations on time. The customer's income level is a crucial factor affecting their ability to pay credit on time. The higher the income, the greater the likelihood that they have sufficient financial resources to pay their credit installments on time.

Based on calculations, it is obtained that, an increase in the value of 1 unit for the capital variable can reduce a person's risk 0.616 times to be late, in other words not too late. Based on the calculation of the indicator model, the capital variable is reflected most strongly by the item $X_3.1$, namely a fixed source of income. Thus, to reduce the risk of someone being late, they can focus on assessing the customer's source of income.

Customers with sufficient capital tend to have more financial resources available to meet their financial obligations. With enough capital, they are more likely to pay credit installments on time. They may have a financial reserve or a better financial strategy to deal with unexpected financial emergencies or difficulties. In this case, they are more likely to maintain on-time credit payments. On the other hand, customers with limited capital may face financial difficulties, which can lead to credit payment delays.

Based on calculations, it is obtained that, an increase in the value of 1 unit for collateral variables can increase a person's risk 1,104 times to be late. Based on the calculation of the indicator model, the collateral variable is reflected most strongly by the item $X_4.2$, namely the type of collateral is physical or non-physical. Thus, to lower the risk of someone being late can focus on assessing the type of guarantee.

Collateral or security can provide security to lenders. With collateral or security, lenders have assets that can be used as security if borrowers fail to repay their credit. This makes lenders more comfortable in providing credit and may be more inclined to offer lower interest rates. However, it's important to remember that even with collateral or security, payment delays can still occur. Other factors, such as changes in the borrower's financial situation, unforeseen financial issues, or changes in market conditions, can also affect the timing of credit payment delays.

Based on calculations, it is obtained that, an increase in the value of 1 unit for the variable condition of economy can increase a person's risk 1,104 times to be late. Based on the calculation of the indicator model, the condition of economy variable is reflected most strongly by the item $X_5.1$, namely business development / investment. Thus, to reduce the risk of someone being late can focus on assessing the customer's business development and investment.

A poor economic condition can result in a decrease in household income. When income decreases, customers may face difficulties in meeting their credit payments, especially if their credit installments are high. If the prices of goods and services increase (inflation), customers may experience additional pressure on their budgets. Price increases can raise the cost of living, reducing the remaining money available for credit payments.

5 Conclusion

Based on the results of the analysis, it was obtained that the model that can present the relationship between the 5C assessment variable and the time of delay in paying credit on home ownership loans at Bank X is as follows.

$$h_i(t) = h_0(t)\exp\Big(-0.437 \cdot \text{Character} - 0.456 \cdot \text{Capacity} \\ -0.484 \cdot \text{Capital} + 0.099 \cdot \text{Collateral} - 0.125 \cdot \text{Condition}\Big)$$

All 5C variables can be considered to reduce the risk of late payment of customer credit Home ownership loans at Bank X. The customer's character is reflected most strongly by the intention and responsibility of a customer. Customer capasity is reflected most strongly by the customer's ability to complete credit on time. Customer capital is reflected most strongly by the customer's steady income source. Customer collateral is reflected most strongly by the type of guarantee is physical or non-physical. The condition of economy is reflected most strongly by business development and customer investment. These items can be the main focus for banks to assess customers in order to minimize the risk of bad loans.

Acknowledgement
I would like to thank all parties who have helped in preparing this journal. I would like to thank Mr. Adji Achmad Rinaldo Fernandes for permission to use research data.

References

Austin, P.C., Latouche, A., & Fine, J.P. (2020). A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model. *Statistics in Medicine*, *39*(2), 103–113.

Campigotto, F., Weller, E. (2014). Impact of informative censoring on the Kaplan-Meier estimate of progression-free survival in phase II clinical trials. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Biology*, *32*(27), 3068–3074.

Castaldo, A., Pittiglio, R., Reganati, F., & Sarno, D. (2023). Access to bank financing and start-up resilience: A survival analysis across business sectors in a time of crisis. *The Manchester School*, *91*(7), 141–170.

Collet, D. (2003). *Modelling Data in Medical Reseaarch, second edition (2nd ed.)*. Chapman and Hall.

Dirick, L., Claeskens, G., & Baesens, B. (2017). ATime to default in credit scoring using survival analysis: A benchmark study. *Journal of the Operational Research Society*, *68*(6), 652–665.

Edwards, J.R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, *14*(2), 370-388.

Fernandes, A.A.R., Hutahayan, B., Kartikasari, D.P. (2022). *Evaluasi dan Pengembangan Pengukuran Risiko (Risk Metric) dan Model Risiko (Risk Model) Berbasis Pemodelan Statistika dengan Fleksibilitas Tinggi.* (Penelitian Dasar Kompetitif Nasional, Brawijaya University, 2023).

Ferreira, J.C., & Patino, C.M. (2016). What is survival analysis, and when should I use it?. *Jornal Brasileiro de Pneumologia*, *42*(1), 77.

Hanafiah, M.H. (2020). Formative vs. reflective measurement model: Guidelines for structural equation modeling research. *International Journal of Analysis and Applications*, *18*(5), 876-889.

Hougaard, P. (1995). Frailty models for survival data. *Lifetime Data Analysis*, *1*(3), 255–273.

Jiang, C., Wang, Z., & Zhao, H. (2019). A prediction-driven mixture cure model and its application in credit scoring. *European Journal of Operational Research, Elsevier*, *277*(1), 20-31.

Katzman, J.L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Medical Research Methodology*, *18*(1), 1–12.

Kleinbaum, D.G., Klein, M. (2005). *Survival analysis: a self-learning text.* New York: Springer.

Katzman, J.L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). Testing the proportional hazards assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: methodological perspectives and review. *BMC Musculoskeletal Disorders*, *22*(1), 1–7.

Lawless, J.F. (2002). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.

Lesnoff, M., Roger, J.M., & Rutledge, D.N. (2021). Monte Carlo methods for estimating Mallows's Cp and AIC criteria for PLSR models. Illustration on agronomic spectroscopic NIR data. *Journal of Chemometrics*, *35*(10), 1–21.

Lin, H.M., Liu, S. T.H., Levin, M.A., Williamson, J., Bouvier, N.M., Aberg, J.A., Reich, D., & Egorova, N. (2023). Informative Censoring—A Cause of Bias in Estimating COVID-19 Mortality Using Hospital Data. *Life*, *13*(1), 210.

Liu, Y. (2017). Sensitivity analyses for informative censoring in survival data: A trial example. *Journal of Biopharmaceutical Statistics*, *27*(4), 595–610.

Marletta, A., Nuovo, A. (2021). The Evaluation of Credit Risk Using Survival Models: An Application on Italian SMES. *Statistica & Applicazioni*, *XIX*(2), 161–178.

Perron, B.E., & Gillespie, D.F. (2015). Latent Variables. *Key Concepts in Measurement*, 93–118.

Pham, H. (2021). A New Criterion for Model Selection. *Mathematics 2019*, *7*(12), 1215.

Suárez, R.P., Abad, R.C., & Fernández, J.M.V. (2021). Probability of default estimation in credit risk using a nonparametric approach. *Test*, *30*(2), 383–405.

Rabe-Hesketh, S., Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. Texas: Stata Press.

Rutledge, D.N., Roger, J.-M., & Lesnoff, M. (2021). Different Methods for Determining the Dimensionality of Multivariate Models. *Frontiers in Analytical Science*, *1*, 754447.

Šiaulys, J., Puišys, R. (2022). Survival with Random Effect. *Mathematics 2022*, *10*(7), 1097.

Lawless, J.F. (2002). *Metode statistika multivariat pemodelan persamaan struktural (SEM) pendekatan WarpPLS.*. Malang: UB Press.

Stensrud, M.J., Hernán, M.A. (2020). Why test for proportional hazards?. *JAMA Guide to Statistics and Methods*, *323*(14), 1401–1402.

Tan, A., Anchor, J. (2016). Stability and profitability in the Chinese banking industry: evidence from an auto-regressive-distributed linear specification. *Investment Management and Financial Innovations*, *13*(4), 120–128

Tan, Y., Floros, C. (2018). Risk, competition and efficiency in banking: Evidence from China. *Global Finance Journal*, *35*, 223-236.

Zhang, Z., Reinikainen, J., Adeleke, K.A., Pieterse, M.E., & Groothuis-Oudshoorn, C. G. M. (2018). Time-varying covariates and coefficients in Cox regression models. *Annals of Translational Medicine*, *6*(7), 121–121.