Martina NOVOTNÁ, PhD
martina.novotna@vsb.cz
VSB – Technical University Ostrava, Czech Republic

Corporate Bankruptcy Risk across the Firm Life Cycle with Insights into Failure Dynamics

Abstract. *The prediction of corporate bankruptcy remains a critical topic in risk management, especially in the context of past economic crises and uneven economic development. Reliable forecasting requires the careful selection of modelling approaches that reflect their underlying assumptions and potential for practical application. This article applies and evaluates selected statistical methods to estimate bankruptcy models and identify key financial and structural factors affecting corporate survival. Using a longitudinal dataset of Czech firms across multiple industries, the study examines logistic regression models for repeated firm-year observations and survival analysis techniques, including Cox proportional hazards, Weibull, and flexible parametric models. The results show that profitability, leverage, liquidity, and firm age significantly shape bankruptcy risk. While logistic regression offers intuitive probability estimates, survival models provide richer insight into the timing and intensity of default, revealing a peak hazard between 11 and 14 years of firm age. The study contributes by integrating both modelling perspectives and offering practical implications for financial institutions, managers, and policymakers.*

Keywords: *bankruptcy prediction, hazard function, logistic regression, survival analysis, time-to-event modelling.*

JEL Classification: C50; G32; G33.

Received: 11 December 2025	Revised: 4 June 2026	Accepted: 11 June 2026

1. Introduction

The prediction of bankruptcy has been given considerable attention in the academic literature in recent years. Among the research, the study by Altman (1968) and the model known as Altman's or Z-score model are worth mentioning. More recently, Altman et al. (2018) provided a review of the model's effectiveness and importance globally, as well as its applications in finance and related areas. Their results suggest that while a general international model performs reasonably well, classification accuracy at the level of individual economies can be considerably improved with country-specific estimation.

For this reason, survival analysis can be considered a suitable alternative for examining corporate bankruptcy and estimating bankruptcy models. However, this area has not yet attracted adequate attention compared to the traditional methods mentioned above. In many cases, survival analysis was performed using the Cox proportional hazard model, but it is also common to use parametric models, such as

DOI: 10.24818/18423264/60.2.26.09

the Weibull model. Although some works use only one of these approaches and focus primarily on revealing significant factors in the probability of financial distress (e.g. default or bankruptcy), other studies use multiple approaches and verify the main results of the models by subsequent comparison.

The main objective of this study is to explore methods for estimating bankruptcy models and identifying the key factors that contribute to bankruptcy probability. In addition to financial variables, we assess the impact of industry, company size, and, most importantly, company age. By examining these factors, we aim to provide insights into the complex nature of bankruptcy prediction and to help understand the various factors that influence the likelihood of bankruptcy.

The contribution of this article lies in the following: First, the empirical evidence is based on the survival of Czech companies over a relatively long period, which is not sufficiently addressed in the literature. Although it can be assumed that the financial variables affecting the probability of financial distress are generally well established, verifying whether and to what extent they play a role in selected economies and sectors is desirable. Second, the study compares the results from logistic regression with those obtained from survival models. In addition, the application section employs various semi-parametric and parametric survival analysis models, including flexible models. Thus, the comprehensive approach provides a rigorous analysis of the factors affecting the survival of Czech companies.

The structure of this article is as follows. Firstly, we provide a literature review in Section 2 that summarises studies relevant to our research. In Section 3, we describe the methodology, including the data and the statistical methods. The results are presented in Section 4, followed by the discussion and conclusion.

2. Literature review

The analysis of company bankruptcy typically involves using credit score models, which are statistically derived to predict credit risk. One notable scoring model is Altman's (1968) Z-score model. LeClere (2000) provides an overview of early research on financial distress based on accounting indicators, highlighting the evolution of statistical models used for bankruptcy prediction. Over the years, further research in this field has been directed towards specific markets, exploring various types of predictors, and, above all, the application of new methods and their comparison (Muñoz-Izquierdo et al., 2019; Huang & Yen, 2019; Mai et al., 2019; Zoričák et al., 2020; Sun et al., 2021).

Earlier empirical work also documents a broad range of statistical, behavioural, and hybrid approaches to bankruptcy prediction, including studies on bank failure, credit scoring, and corporate liquidations (Lane, 1986; Grice & Dugan, 2001; Glennon & Nigro, 2005; Kelly et al., 2015; Louzada et al., 2014; Pereira, 2014; Feng et al., 2019).

Recent literature increasingly emphasises the limitations of static models and highlights the importance of time-dependent approaches. Survival analysis gained attention for its ability to incorporate censored data and model hazard rates over time

(Zivich et al., 2025). These models allow for a more nuanced understanding of bankruptcy risk, especially when firm age and time-varying covariates are considered. Comparative studies confirm that survival models often outperform logistic regression in capturing the dynamics of financial distress (e.g., Enserro & Miller, 2025). In accounting research, LeClere's (2000) review found a preference for qualitative response models such as probit or logistic regression, which estimate the probability of an event occurring. However, these models have a limitation in that they may only have data available immediately before the event. To overcome this limitation, the author suggests using survival analysis methods to examine events' occurrence and timing.

Laitinen and Kankaanpää (1999) discussed six of the most popular alternative methods for predicting financial failure, including survival analysis. Their research suggests that failure prediction accuracy varies with the technique used. However, they did not propose a single best method to predict financial failure. Agarwal and Audretsch (2001) focused on the effect of company size on survival and found that smaller companies are less likely to survive than larger companies. However, they also caution against making general pronouncements about the relationship between company size and survival, as this relationship can vary over the industry cycle and with the technological demands of the industry.

Machine learning (ML) and explainable AI (XAI) have reshaped bankruptcy prediction. Ensemble models (XGBoost, LightGBM, NGBoost) deliver top predictive performance (Nguyen et al., 2023), but suffer from limited interpretability. SHAP (Shapley Additive Explanations) has become essential for explaining these models. For example, Nguyen et al. (2023) use it to clarify the role of financial ratios, while Fasano et al. (2025) combine deep recurrent networks with XAI to achieve high accuracy and actionable insights for SMEs. These advances build on earlier machine-learning research applied to financial distress prediction (Mai et al., 2019).

Further research applies or compares survival analysis methods for predicting financial distress. Studies range from Cox models and CART trees (Gepp & Kumar, 2015) to analyses of firm-specific uncertainty (Byrne et al., 2016), technology-based firms' resources (Löfsten, 2016), and new ventures in crises (Simón-Moya et al., 2016). Other work includes dynamic-logit survival models for distress (Christopoulos et al., 2018), website-based survival signals (Blazquez et al., 2018), hotel survival determinants (Gemar et al. 2019), and crisis-driven failure risks moderated by R&D human capital (Martinez et al. 2019).

In summary, recent research highlights the value of integrating traditional statistical methods with modern machine learning and time-to-event models. Survival analysis is increasingly viewed as a powerful tool for capturing the dynamic nature of bankruptcy risk, particularly when combined with time-varying covariates and explainable AI. This article contributes by applying logistic regression and survival models to a unique longitudinal sample of Czech firms, revealing the temporal dynamics of corporate failure. Survival modelling has also been applied to financial distress in other contexts, showing that time-varying indicators can meaningfully improve predictive accuracy (Muñoz-Izquierdo et al., 2019).

The article is structured as follows: Section 2 reviews the relevant bankruptcy prediction literature; Section 3 outlines the dataset and methods; Section 4 presents the logistic and survival model results; Section 5 discusses the findings in relation to prior research; and Section 6 concludes with implications for practitioners and policy makers.

3. Research methods

Based on the research objectives and the literature review, the following hypotheses are formulated:
H1: Bankruptcy risk varies significantly by industry and firm size, with small firms in cyclical sectors being most vulnerable.
H2: Key financial indicators from the accounting statements significantly predict bankruptcy.
H3: Firm age is negatively associated with bankruptcy risk, with younger firms more prone to failure.

Research sample

The empirical analysis is based on a longitudinal database of Czech firms from 1988–2015 obtained from Magnusweb. The data have the structure of an unbalanced firm-year panel, as individual firms may contribute repeated annual observations over time. The logistic models are estimated on 64,810 firm-year observations after complete-case selection. Within this sample, 708 firm-year observations correspond to bankruptcy events. For the survival analysis, the data are reorganised for time-to-event modelling and, therefore, differ from the firm-year structure used in the logistic specification. The survival sample contains 10,093 firms, 62,107 survival records, and 708 observed failures. Time to failure averages 8,216 days, 95% CI [8,182.7, 8,249.6]. The lower number of records in the survival setting reflects the transformation of the original firm-year data into a time-to-event structure and the corresponding requirements of survival modelling. Firms that did not go bankrupt during the observation window are treated as right-censored observations.

The 22 original financial indicators were checked for outliers and winsorised (5th/95th percentiles), (Ludwig-Mayerhofer, 2020). After removing insignificant or highly correlated variables, key predictors were selected (Table 1). Two binary categorical variables capture industry (agriculture/utility vs. industrials/services) and size (micro vs. small/medium/large), following OECD thresholds (<10 employees for micro). These classifications support testing whether industry and size are related to bankruptcy risk.

Table 1. Description of variables

Variable	Description	Mean	Median
roa	Return on assets	0.0345	0.0288
roc	Returns on costs	0.1088	0.0500
cla	Coverage of long assets (equity+long term liab)/fixed assets	2.4538	0.9920
er	Equity ratio (equity/total assets)	0.4484	0.4608
liab_t	Liability turnover (days)	196.3546	85.2195
td	Debt ratio (debt/equity)	0.5320	0.5141
cr	Current ratio	3.5654	1.9546
lnta	Ln (total assets)	9.5717	9.569
rec_to_ca	Receivables to current assets	0.5011	0.4934
t_life	Age of the company	4598.008	4644

Source: Authors' processing.

Logistic regression analysis

Logistic regression is used to estimate the probability of bankruptcy in the firm-year data. Since the dependent variable is binary, the model expresses the relationship between bankruptcy occurrence and the selected explanatory variables through the log-odds function. The estimated coefficients are interpreted using odds ratios, where $OR = \exp(\beta)$. An odds ratio greater than one indicates that a higher value of the explanatory variable is associated with higher odds of bankruptcy, while an odds ratio below one indicates lower odds of bankruptcy, ceteris paribus (Hosmer et al., 2013).

In finance, logistic regression is widely used for binary default prediction (e.g., Menard, 2010; Hosmer et al., 2013; Harrell, 2015). Logistic models classify firms into two outcome categories using one or more predictors, assume a binomial distribution, and estimate parameters by maximising the likelihood of the observed data. Covariates in our study are selected following Hosmer et al.'s (2013) purposeful selection approach, beginning with univariable tests and retaining variables with $p < 0.25$ for multivariable modelling.

We assume a collection of p independent variables denoted by the vector $\mathbf{x}' = (x_1, x_2, \cdots, x_p)$, where each of these variables is at least interval scaled. If we denote the conditional probability that the outcome is present by the expression $\Pr(Y = 1|\mathbf{x}) = \pi(x)$, then the logit of the multiple logistic regression model has the following form:

$$g(x) = ln\left[\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_{p,} \quad (1)$$

and the multiple logistic model is given by:

$$\pi(x) = \frac{e^{g(\mathbf{x})}}{1+e^{g(\mathbf{x})}}. \quad (2)$$

Because the dataset contains repeated observations for the same firms, the standard logistic model is used only as a benchmark. To account for within-firm dependence, two correlated binary-response specifications are estimated: a cluster-specific random-effects model and a population-averaged model. The random-effects model contains 10,196 firm-level groups, with one to nine observations per firm and an average of 6.4 observations. The estimated intra-firm correlation is practically zero (rho ≈ 4.49e-06), and the LR test of rho = 0 is not statistically significant ($p = 0.490$). This helps explain why the standard, random-effects, and population-averaged estimates are very similar.

A fixed-effects logit specification was considered as an alternative. However, this approach uses only firms with within-firm variation in bankruptcy outcomes, thereby substantially reducing the usable sample. Moreover, fixed effects absorb time-invariant structural characteristics, such as industry and size, which are of direct interest in this study. For these reasons, the fixed-effects logit is not adopted as the main specification.

The distinction between the binary-response and survival parts is as follows: the logistic models estimate bankruptcy probability in firm-year observations, whereas survival models use the failure-time structure to estimate when bankruptcy occurs and how the bankruptcy hazard evolves over firm age. This distinction is important because logistic regression can identify whether firm age is associated with bankruptcy probability, but survival analysis is needed to describe the timing and intensity of bankruptcy risk.

Survival analysis

Survival analysis models the time until bankruptcy while properly handling right-censored observations (Hosmer et al., 2008; Harrell, 2015). In this part of the analysis, firms are followed from entry until bankruptcy or censoring. The effects of explanatory variables are interpreted through hazard ratios, where $HR = \exp(\beta)$. A hazard ratio greater than one indicates a higher bankruptcy hazard, meaning a higher instantaneous likelihood of failure at time t, conditional on survival up to that time. Conversely, a hazard ratio below one indicates a lower bankruptcy hazard. Unlike odds ratios in logistic regression, hazard ratios reflect the relative change in the hazard of bankruptcy over time rather than the odds of bankruptcy occurrence.

The Cox proportional hazards model is used as the semi-parametric benchmark (Hosmer et al., 2008; Cleves et al., 2010; Harrell, 2015). The proportional hazards assumption is assessed before retaining the final Cox specification; variables with evidence of violation are not retained in the final Cox model. The Weibull model is then estimated as a parsimonious parametric benchmark (Hosmer et al., 2008; Cleves et al., 2010). In contrast to the Cox model, the Weibull model specifies the baseline hazard parametrically and assumes a monotonic hazard pattern over time. To examine whether the baseline hazard follows a non-monotonic pattern over the firm life cycle, flexible parametric Royston–Parmar models are estimated following Royston and Lambert (2011). In this framework, the model with one degree of freedom corresponds to the Weibull specification, while higher degrees of freedom allow a more flexible hazard shape.

The Cox proportional hazards model does not specify the baseline hazard parametrically, but models the effect of covariates multiplicatively on the baseline hazard. The hazard function can be written as follows:

$$h(t, x, \beta) = h_0(t) r(x, \beta). \tag{3}$$

The hazard function in (3) consists of two parts: the baseline hazard $h_0(t)$, which describes how the hazard changes over survival time, and the relative risk component $r(x, \beta)$, which captures the effect of explanatory variables. No parametric form is imposed on $h_0(t)$, although this flexibility may come at the cost of lower efficiency compared with fully parametric models.

Flexible parametric models are estimated on the hazard scale using one to four degrees of freedom for the baseline spline. The preferred flexible specification corresponds to four degrees of freedom for the baseline restricted cubic spline, implying three interior knots. Since the knots were not set manually, the default `stpm2` knot placement was used in Stata. The interior knots were located at 1,571, 2,921 and 4,485 days of firm age, with boundary knots at 45 and 6,330 days. This specification is retained because it captures the non-monotonic shape of bankruptcy hazard over firm age while preserving interpretability (Royston and Lambert, 2011).

The general flexible model can be described as follows:

$$\ln H(t) = f(t; \gamma), \tag{4}$$

where $f(t; \gamma)$ represents a nonlinear spline-based function of time t, having a parameter vector γ.

Model performance is compared primarily using log-likelihood, AIC, and BIC. The measures of explained variation are reported only as supplementary indicators and are not interpreted as direct equivalents of the adjusted R^2 used in linear regression. According to the information criteria, the more parsimonious specifications perform well, while the flexible model with three interior knots is retained for interpreting the non-linear shape of bankruptcy hazard over firm age.

4. Results

Logistic models

The logistic regression uses panel data, treating each firm as a cluster of correlated yearly observations. Financial variables are time-varying, while industry and size remain constant; firm age serves as the time variable. We estimate a standard logistic model alongside cluster-specific and population-averaged versions. The dataset includes 10,196 firms and 64,810 observations over the observed period.

Variable selection follows purposeful selection based on univariable screening and multivariable diagnostics. The final logistic specifications include financial ratios, firm age, and structural characteristics. The standard model is reported only as a benchmark; the cluster-specific and population-averaged models are used to demonstrate robustness to within-firm dependence. The estimates are very similar across the three specifications because the estimated intra-firm correlation in the random-effects logit model is practically zero.

Table 2. Estimated coefficients of logistic models (t-continuous)

	Standard model	Cluster-specific model			Population average model		
Variable	**Coef.**	**Coef.**	**Change (%)**	**SE**	**Coef.**	**Change (%)**	**SE**
roa	-0.9150*	-0.9150**	-2.9240		-0.9341**	-2.8121	
roc	-0.1751*	-0.1751*	14.2171		-0.1760*	14.3525	
cla	0.0280**	0.0280**	0.1371		0.0281**	0.0991	

	Standard model	Cluster-specific model		Population average model	
er	-3.2108**	-3.2108**	-6.7588	-3.2231**	-6.8145
cr	-0.0818**	-0.0818**	-16.6729	-0.0825**	-16.4603
lnta	0.1323**	0.1323**	-3.5370	0.1337**	-3.7682
liab_t	-0.0003**	-0.0003**	-2.4247	-0.0003**	-2.5716
rec_to_ca	0.9431**	0.9431**	-0.9009	0.9467**	-0.9656
t_life	-0.0004**	-0.0004**	-4.0816	-0.0004**	-4.0816
2.industry_b	-1.0078**	-1.0078**	1.2639	-1.0089**	0.9677
2.size_b	0.4831**	0.4831**	-3.9203	0.4807**	-4.1965
_cons	-3.8037**	-3.8037**	-0.0762	-3.8249**	-0.2635

* p<0.10, ** p<0.05

Source: Authors' processing.

To better understand the relationship between age and bankruptcy, instead of the continuous variable *t_life*, we use the categorical variable *t_med*, which divides company age into four groups based on quartiles of the median. The estimated coefficients of models with the time categorical variable are summarised in Table 3.

Table 3. Estimated coefficients of logistic models (t-categorical)

	Standard model	Cluster-specific model		Population average model	
Variable	**Coef.**	**Coef.**	**Change SE (%)**	**Coef.**	**Change SE (%)**
roa	-1.0263**	-1.0263**	-2.4840	-1.0465**	-2.3716
roc	-0.1858**	-0.1858*	13.0614	-0.1863*	13.1800
cla	0.0282**	0.0282**	0.4950	0.0283**	0.4717
er	-3.2046**	-3.2046**	-5.0611	-3.2142**	-5.1089
cr	-0.0844**	-0.0844 **	-2.9483	-0.0850 **	-3.1537
lnta	0.1115**	0.1115**	-17.4629	0.1131**	-17.2716
liab_t	-0.0003 **	-0.0003 **	-2.0710	-0.0003**	-2.2189
rec_to_ca	0.9312**	0.9312**	-1.1474	0.9343**	-1.2035
2.t_med	-0.6343**	-0.6343**	0.7156	-0.6344**	0.8841
3.t_med	-0.5634**	-0.5634**	-0.3667	-0.5586**	-0.4650
4.t_med	-3.6654**	-3.6654**	0.0203	-3.6491**	-0.8526
2.industry_b	-1.0017**	-1.0017**	1.5750	-1.0024**	1.2888
2.size_b	0.5024**	0.5024**	-3.9063	0.4996**	-4.1738
_cons	-4.3297**	-4.3297**	-1.2701	-4.3483**	-1.4470

* p<0.10, ** p<0.05

Source: Authors' processing.

Estimates are highly consistent across standard, cluster-specific, and population-averaged models, reflecting the near-zero intra-firm correlation in the random-effects specification. Our results confirm that bankruptcy probability is significantly driven by financial condition, age, and structural factors. Specifically, firms in industrials and services face substantially lower bankruptcy odds than those in agriculture and utilities (OR = 0.37). Interestingly, non-micro firms show a higher risk profile than micro firms (OR = 1.65), ceteris paribus.

The categorical age specification confirms that older firms have lower bankruptcy odds than the youngest group. The negative coefficients for the higher age categories support the hypothesis that firm age has a protective effect. Nevertheless, logistic regression cannot capture the timing and shape of bankruptcy risk over the life cycle, which motivates the following survival analysis.

Overall, logistic regression highlights the roles of financial condition, firm age, and structural characteristics, but cannot account for the timing of risk escalation – addressed later through survival analysis.

Survival models

The survival analysis starts with the Cox proportional hazards model. After purposeful covariate selection and testing of the proportional hazards assumption, variables violating the assumption are excluded from the final Cox specification. This step provides a semi-parametric benchmark for the subsequent Weibull and flexible parametric survival models. The final Cox model (see Table 5) is then used to define the main set of predictors for the parametric survival models. The survival and cumulative hazard curves are interpreted for combinations of industry and size at mean financial values.

Figure 1 displays the survival and cumulative hazard functions, which help illustrate the relationship between a company's age and the probability of a bankruptcy event. Curves S1-S4 are plotted for the mean values of five financial variables in the model and four variants according to the sector (i) and company size (s) as follows: V1 (i=1, s=1), V2 (i=1, s=2), V3 (i=2, s=1), V4 (i=2, s=2). Thus, for example, the S3 curve is the survival function based on mean values of financial variables and micro companies from services and industrials. The baseline survival function S0 is plotted for zero values of the variables. The corresponding cumulative hazard curves are then displayed on the right side and labelled H0-H4.

The survival curves show that the riskiest group is V2 (small firms in agriculture/utilities), while the most resilient group is V3 (micro firms in industrials/services). These patterns are consistent with the logistic regression results.

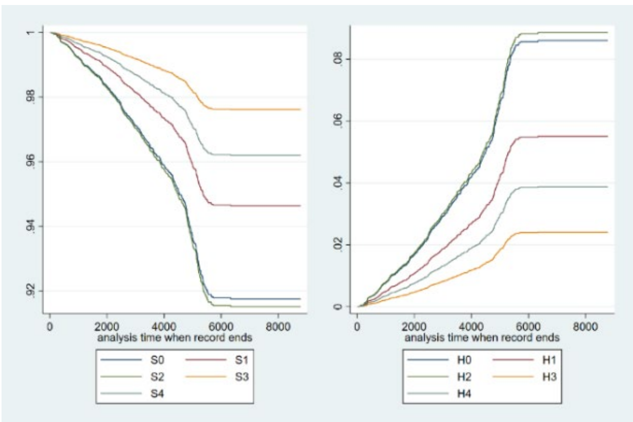

Figure 1. Survival and cumulative hazard functions

Source: Authors' own creation.

Figure 2 shows kernel-smoothed hazard functions, revealing that the hazard rate is not constant: it rises gradually, peaks sharply around 4,000–5,000 days (≈11–14 years), and then declines. This confirms that bankruptcy risk is highest in mid-life stages of firms. Consistent with earlier results, V3 exhibits the lowest peak hazard, whereas V2 remains the riskiest. These temporal dynamics are not identifiable through logistic regression, underscoring the added value of survival analysis.

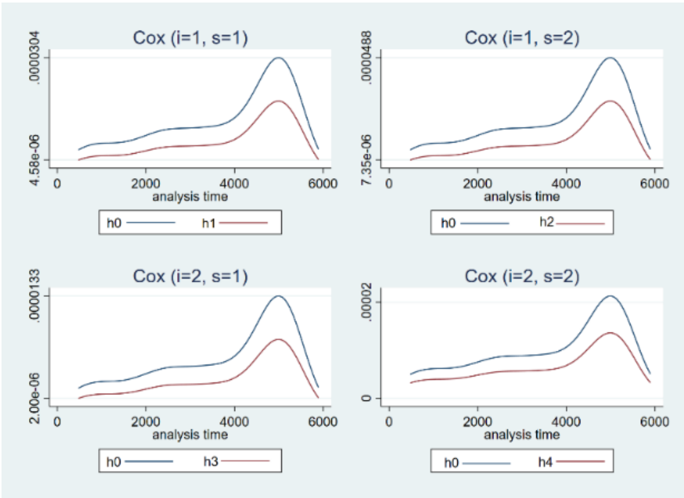

Figure 2. Smoothed hazard functions
Source: Authors' own creation.

After identifying the most influential variables using the Cox model, we develop additional survival models, such as the Weibull model and flexible models with interior knots (1–3 knots). The estimated models are comparable; however, some spline coefficients are not statistically significant in models with one and two interior knots. The Weibull/df1 model is the most parsimonious according to AIC and BIC, whereas PH(4) is retained as the flexible specification for interpreting the non-monotonic shape of the baseline hazard (see Table 4). The use of both models is therefore intentional: the simpler Weibull/df1 specification provides a parsimonious benchmark, while PH(4) provides a richer description of hazard dynamics over firm age. This directly addresses model-selection concerns without claiming that one single model is universally superior.

Table 4. Comparison of Weibull and flexible model

Model	**d.f.**	***p***	**AIC**	**BIC**	**Explained variation**
Weibull	1	1.30 (0.044)	5249.653	5330.982	0.402 (0.018)
PH(4)	4	1.720 (0.242)	5340.890	5449.534	0.400 (0.018)

Source: Authors' processing.

Figure 3 presents the baseline cumulative and hazard functions (with covariates set to zero). Both models indicate that bankruptcy intensity varies with firm age. The Weibull model shows a monotonically increasing hazard (p = 1.3), rising steeply during the first ~2,000 days (≈5.5 years). The spline model (hdf4) identifies the highest risk between roughly 7.7 and 14.2 years (2,800–5,200 days), after which the hazard declines and stabilizes around 16.5 years.

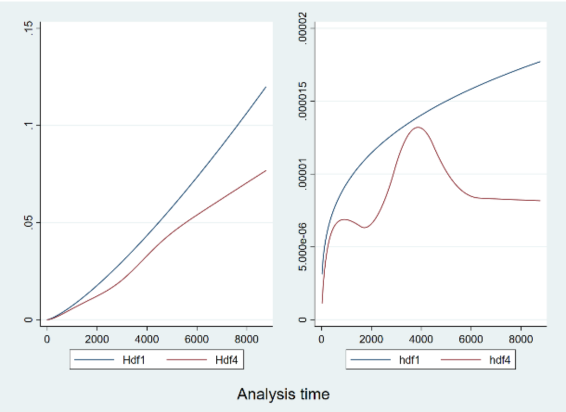

Figure 3. Cumulative and baseline functions

Source: Authors' own creation.

Table 5 shows three survival models' coefficient estimates and corresponding hazard ratios (all coefficients are statistically significant at p<0.05). The coefficients and hazard ratios of the parametric models are similar and support the main findings of the Cox model, suggesting only slight differences. The bankruptcy hazard is increased by *cla* and *td*, and negatively affected by *roa*, *er*, and *liab_t*. For example, a 1-unit increase in *td* increases the bankruptcy hazard by approximately 270%, while a 1-unit increase in *roa* decreases the hazard by 70% in the Weibull model.

Table 5. Estimated coefficients and hazard ratios

	Cox		Weibull		PH4	
Variable	**Coef.**	**HR.**	**Coef.**	**HR.**	**Coef.**	**HR.**
roa	-1.4490	0.2348	-1.2286	0.2927	-1.3415	0.2615
cla	0.0274	1.0278	0.0280	1.0284	0.0253	1.0256
er	-2.3494	0.0954	-2.4469	0.0866	-1.9986	0.1355
liab_t	-0.0004	0.9996	-0.0004	0.9996	-.0004	0.9996
td	1.2650	3.5432	1.2999	3.6689	1.6175	5.0403
2.industry_b	-0.8289	0.4365	-0.8251	0.4382	-0.8097	0.4450
2.size_b	0.4732	1.6051	0.4651	1.5922	0.4588	1.5822
cons			-13.885	0.0000	-16.547	
rcs1 (p)			1.2958	1.2958	1.72	
rcs2					0.61	
rcs3					-2.31	
rcs4					2.95	

Source: Authors' processing.

These effects are consistent with economic theory and prior research. We compare the fit of the Cox and the Weibull model based on the Cox-Snell residual plots in Figure 4. The overall fit of both models is comparable.

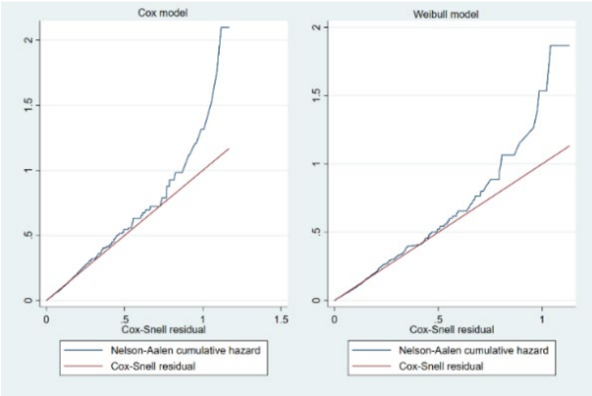

Figure 4. Cox Snell residual plots
Source: Authors' own creation.

5. Discussion

This study offers a comprehensive assessment of corporate bankruptcy risk by combining logistic regression with survival analysis. Drawing on a large panel dataset of Czech firms, we identify key financial and structural determinants of failure and show that time-to-event models provide a more detailed understanding of how risk evolves over the life cycle of a firm.

Our results are broadly consistent with previous studies that emphasise the predictive importance of profitability, liquidity, leverage, and activity ratios (Laitinen, 2005; Agarwal & Taffler, 2008; Gepp & Kumar, 2015; Byrne et al., 2016; Altman et al., 2018; Blazquez et al., 2018; Christopoulos et al., 2018; Gemar et al., 2019). In line with Laitinen (2005), Byrne et al. (2016), and Pierri & Caroni (2017), we confirm that higher profitability and a stronger equity position lower the probability of bankruptcy across all estimated models. We also identify several additional financial indicators, such as return on costs, current ratio, receivables to current assets, and ln (total assets), that significantly contribute to bankruptcy prediction.

Some variables display model-specific effects. While liabilities turnover would theoretically be expected to increase financial stress, our logistic model indicates a positive effect, whereas the Cox and Weibull models show a small negative association. This mixed pattern is also noted in earlier empirical work, where turnover variables often depend on the modelling framework and sector composition (e.g., Agarwal & Taffler, 2008).

The positive effect of receivables on current assets supports the interpretation of Blazquez et al. (2018), suggesting liquidity pressure or inefficient working-capital management, especially among smaller firms. Firm size also appears to matter, although its effect depends on the way it is measured. In the logistic models, the continuous size proxy ln(total assets) is positively associated with bankruptcy odds, which differs from studies suggesting that larger firms tend to face lower bankruptcy risk (Altman et al., 2018). The categorical size indicator reveals that non-micro firms face higher bankruptcy risk than micro-firms, a finding confirmed across both logistic and survival models. While this contrasts with studies by Simón-Moya et al. (2016) or Martinez et al. (2019), it aligns with Laitinen (2005) and Blazquez et al. (2018), who also found lower risk profiles among the smallest entities. This disparity may be explained by organisational scale. As noted by Becker-Blease et al. (2010), employee-based metrics often capture organisational complexity. In this context, larger firms may face greater structural rigidities and higher fixed costs, whereas micro-firms benefit from the entrepreneurial dynamics and flexibility essential for survival in new ventures (Agarwal & Audretsch, 2001; Allen & Rose, 2006).

Industry effects remain important. Although Altman et al. (2018) identify heightened risk in construction and manufacturing, our broader industry categories do not allow replication. Instead, we find that industrial and service firms have lower bankruptcy probabilities than agriculture and utilities. Simón-Moya et al. (2016), on the other hand, report a higher failure risk for new service ventures, indicating that sector effects are sensitive to firm age and sample composition.

The most notable contribution of this study concerns the role of firm age. While logistic regression confirms that younger firms are more likely to fail, supporting previous findings (Laitinen, 2005; Byrne et al., 2016; Pierri & Caroni, 2017; Altman et al., 2018), the survival models reveal a more detailed, non-linear pattern. Bankruptcy risk peaks between approximately 11 and 14 years and then declines. Similar evidence on the importance of early-stage financial characteristics for firm survival is reported by Fuertes-Callén et al. (2022), who show that survival models can capture distinct risk patterns over the firm life cycle. Such non-linear age effects can only be adequately modelled using flexible survival specifications, such as the Royston–Parmar model. Although early Z-score research implicitly reflected age effects through measures such as retained earnings relative to total assets, later work showed that explicitly incorporating firm age improves interpretation (Altman, 1968; Altman et al., 2018).

From a methodological perspective, the combination of logistic and survival approaches yields a richer understanding of bankruptcy risk. Logistic models estimate the probability of failure in repeated firm-year observations, while survival models uncover the timing and intensity of risk. The standard logistic model is retained as a benchmark, but the random-effects and population-averaged estimates confirm that the main findings are robust to the repeated-observation structure. The fixed-effects logit was considered but not retained as the main model because it substantially restricts the usable sample and absorbs structural variables such as industry and size. Although some variables reflect broader structural characteristics, such as industry, the present study retains a unified regression and survival framework to ensure comparability across specifications. A multilevel extension may represent a useful avenue for future research.

The empirical design is not an aggregate time-series, IV, or GMM model. Therefore, classical unit-root/non-stationarity diagnostics and overidentification tests are not directly applicable to the main research design. The relevant concerns are instead repeated observations, overspecification, and overfitting. These are addressed through correlated binary-response models, survival-time modelling with censoring, purposeful variable selection, information criteria, and diagnostic checks.

6. Conclusions

This article provides a comprehensive analysis of corporate bankruptcy prediction using both logistic regression and survival analysis. Based on a large longitudinal dataset of Czech firms, the study confirms that financial indicators, particularly profitability, leverage, liquidity, and asset structure, play a crucial role in determining bankruptcy risk. In addition, firm-specific characteristics such as age, industry, and size significantly influence the probability of failure.

While logistic regression models offer interpretable estimates of bankruptcy probability, survival analysis provides deeper insights into the timing and intensity of risk. The results show that younger firms are more vulnerable, but the highest hazard of bankruptcy occurs between 11 and 14 years of firm age. This non-linear pattern, captured by flexible parametric survival models, highlights the importance of time-to-event modelling in financial risk assessment.

The combination of logistic regression and survival analysis is methodologically valuable, especially when working with longitudinal data and time-varying covariates. Logistic models estimate the likelihood of bankruptcy, whereas survival models allow for the modelling of hazard rates over time, offering a more dynamic perspective on corporate failure.

The findings have several practical implications. For financial institutions, survival models can enhance credit risk assessment by identifying critical periods in a firm's lifecycle. For corporate managers, the results suggest that mid-life firms may require targeted financial strategies to mitigate risk, particularly in terms of liquidity and receivables management. For policymakers, the study provides evidence that firm age, size, and sector are relevant predictors of bankruptcy, which may inform the design of support programs and regulatory interventions. Finally, for researchers, the article demonstrates the benefits of integrating traditional statistical methods with time-to-event analysis and opens avenues for future work combining survival models with machine learning and explainable AI techniques.

In conclusion, the study contributes to the literature by offering a robust empirical framework for bankruptcy prediction, grounded in both statistical rigour and practical relevance. The results underscore the importance of modelling not only the probability of failure, but also its timing, which is essential for effective financial decision-making and policy design.

Acknowledgements: *This research was funded by the VSB – Technical University of Ostrava under the Research Project No. SP2026/003 and by the Czech Science Foundation (GAČR) under the Research Project No. GA23-07128S.*

References

[1] Agarwal, R., Audretsch, D.B. (2001), *Does entry size matter? The impact of the life cycle and technology on firm survival. The Journal of Industrial Economics*, 49(1), 21-43, https://doi.org/10.1111/1467-6451.00136.

[2] Agarwal, R., Taffler, R. (2008), *Comparing the performance of market-based and accounting-based bankruptcy prediction models. Journal of Banking & Finance,* 32(8), 1541-1551, https://doi.org/10.1016/j.jbankfin.2007.07.014.

[3] Allen, L.N., Rose, L.C. (2006), *Financial survival analysis of defaulted debtors. Journal of the Operational Research Society,* 57, 630-636, https://doi: 10.1057/palgrave.jors.2602038.

[4] Altman, E.I. (1968), *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance,* 4(23), 589-609, https://doi: 10.1111/j.1540-6261.1968.tb00843.x.

[5] Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K., Suvas, A. (2018), *Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-Score model*. Retrieved from SSRN: https://ssrn.com/abstract=2536340 (6. 1. 2023), https://doi: 10.2139/ssrn.2536340.

[6] Becker-Blease, J.R., Kaen, F.R., Etebari, A., Baumann, H. (2010), *Employees, firm size and profitability in US manufacturing industries. Investment Management and Financial Innovations*, 7(2), 7-23.

[7] Blazquez, D., Domenech, J., Debón, A. (2018), *Do corporate websites' changes reflect firms' survival? Online Information Review*, DOI:10.1108/OIR-11-2016-0321.

[8] Byrne, J.P., Spaliara, M.-E., Tsoukas, S. (2016), *Firm survival, uncertainty, and financial frictions: Is there a financial uncertain-ty accelerator? Economic Inquiry*, 1(54), 375-390, https://doi: 10.1111/ecin.12240.

[9] Christopoulos, A.G., Dokas, I.G., Kalantonis, P., Koukkou, T. (2018), *Investigation of financial distress with a dynamic logit based on the linkage between liquidity and profitability status of listed firms. Journal of the Operational Research Society*, 1476-9360, https://doi: 10.1080/01605682.2018.1460017.

[10] Cleves, M., Gould, W.W., Marchenko, Y.V. (2010), *An Introduction to Survival Analysis Using Stata*. College Station, TX, USA: Stata Press.

[11] De Leonardis, D., Rocci, R. (2008), *Assessing the default risk by means of a discrete-time survival ana-lysis approach. Applied Stochastic Models in Business and Industry*, 4(24), 291-306, https://doi: 10.1002/asmb.705.

[12] Enserro, D.M., Miller, A. (2025), *Improving the Estimation of Prediction Increment Measures in Logistic and Survival Analysis. Cancers,* 17(8), 1259, https://doi.org/10.3390/cancers17081259.

[13] Fasano, F., Adornetto, C., Zahid, I., La Rocca, M., Montaleone, L., Greco, G., Cariola, A. (2025), *The Dilemma of Accuracy in Bankruptcy Prediction: A New Approach Using Explainable AI Techniques. European Journal of Innovation Management*, 28(11), 1-22, https://doi.org/10.1108/EJIM-06-2024-0633.

[14] Feng, M., Shaonan, T., Chihoon, L., Ling, M. (2019), *Deep learning models for bankruptcy prediction using textual disclosures. European Journal of Operational Research*, 274, 743-758, https://doi: 10.1016/j.ejor.2018.10.024.

[15] Fuertes-Callén, Y., Cuellar-Fernández, B., Serrano-Cinca, C. (2022), *Predicting startup survival using first years financial statements. Journal of Small Business Management,* 60(6), 1314–1350*,* https://doi.org/10.1080/00472778.2020.1750302.

[16] Gemar, G., Soler, I.P., Guzman-Parra, V.F. (2019), *Predicting bankruptcy in resort hotels: a survival analysis. International Journal of Contemporary Hospitality Management*, 31(4), 1546-1566, https://doi: 10.1108/IJCHM-10-2017-0640.

[17] Gepp, A., Kumar, K. (2015), *Predicting financial distress: A comparison of survival analysis and decision tree techniques. Procedia Computer Science*, 54, 396-404, https://doi: 10.1016/j.procs.2015.06.046.

[18] Glennon, D., Nigro, P. (2005), *Measuring the default risk of small business loans: A survival analysis approach. Journal of Money, Credit and Banking*, 37(5), 923-947.

[19] Grice, J.S., Dugan, M.T. (2001), *The limitations of bankruptcy prediction models: Some cautions for the researcher. Review of Quantitative Finance and Accounting*, 17, 151-166.

[20] Harrell, F.E. (2015), *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer, New York, USA.

[21] Hosmer, D.W., Lemeshow, S., May, S. (2008), *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. Hoboken: John Wiley & Sons, Hoboken, USA.

[22] Hosmer, D.W., Lemeshow, S., Sturdivant, R.X. (2013), *Applied Logistic Regression*. Hoboken: Wiley, Hoboken, USA.

[23] Huang, Y.-P., Yen, M.-F. (2019), *A new perspective of performance comparison among machine learning algorithms for financial distress prediction. Applied Soft Computing Journal*, 83, 105663, https://doi: 0.1016/j.asoc.2019.105663.

[24] Kelly, R., O. Brien, E., Stuart, R. (2015), *A long-run survival analysis of corporate liquidations in Ireland. Small Business Economics*, 44(3), 671-683, https://doi: 10.1007/s11187-014-9605-1.

[25] Laitinen, E.K. (2005), *Survival analysis and financial distress prediction: Finnish evidence. Review of Accounting and Finance*, 4(4), 76-90, https://doi: 10.1108/eb043438.

[26] Laitinen, T., Kankaanpää, M. (1999), *Comparative analysis of failure prediction methods: the Finnish case. The European Accounting Review*, 8(1), 67-92, https://doi: 10.1080/096381899336159.

[27] Lane, W.R. (1986), *An application of the Cox proportional hazards model to bank failure. Journal of Banking & Finance*, 4(10), 511-531, https://doi: 10.1016/S0378-4266(86)80003-6.

[28] LeClere, M.J. (2000), *The occurrence and timing of events: Survival analysis applied to the study of financial distress. Journal of Accounting Literature*, 19, 158-189.

[29] Löfsten, H. (2016), *Business and innovation resources. Determinants for the survival of new technology-based firms. Management Decision*, 1(54), 88-106, https://doi: 10.1108/MD-04-2015-0139.

[30] Louzada, F., Cancho, V., de Oliveira Jr., Bao, Y. (2014), *Modeling time to default on a personal loan portfolio in presence of disproportionate hazard rates. Journal of Statistics Applications & Probability*, 3(3), 1-11, https://doi: 10.2139/ssrn.2416547.

[31] Ludwig-Mayerhofer, W. (2020), *Winsorizing and Trimming. Internet Guide to Stata.* Retrieved from: https://wlm.userweb.mwn.de/Stata/wstatwin.htm (18. 12. 2022).

[32] Mai, F., Tian, S., Lee, C., Ma, L. (2019), *Deep learning models for bankruptcy prediction using textual disclosures*. *European Journal of Operational Research*, 274(2), 743-758, https://doi.org/10.1016/j.ejor.2018.10.024.

[33] Martinez, M.G., Zouaghi, F., Marco, T.G., Robinson, C. (2019), *What drives business failure? Exploring the role of internal and external knowledge capabilities during the global financial crisis. Journal of Business Research*, 98, 441-449, https://doi: 10.1016/j.jbusres.2018.07.032.

[34] Menard, S.W. (2010), *Logistic Regression: From Introductory to Advanced Concepts and Applications*. Los Angeles: Sage Publications, Los Angeles, USA.

[35] Muñoz-Izquierdo, N., Laitinen, E.K., Camacho-Miñano, M-M., Pascual-Ezama, D. (2019), *Does audit report information improve financial distress prediction over Altman's traditional Z-Score model? Journal of International Financial Management & Accounting*, 31(1), 65-97, https://doi: 10.1111/jifm.12110.

[36] Nguyen, H.H., Viviani, J.-L., Jabeur, S.B. (2023), *Bankruptcy Prediction Using Machine Learning and Shapley Additive Explanations. Review of Quantitative Finance and Accounting*, 65, 107-148, https://doi.org/10.1007/s11156-023-01192-x.

[37] Pereira, J. (2014), *Survival analysis employed in predicting corporate failure: A forecasting model proposal. International Business Research*, 5(7), 9-20, https://doi: 10.5539/ibr.v7n5p9.

[38] Pierri, F., Caroni, C. (2017), *Bankruptcy prediction by survival models based on current and lagged values of time-varying financial data. Communications in Statistics: Case Studies, Data Analysis and Applications*, 3(3-4), 62-70, https://doi: 10.1080/23737484.2018.1431816.

[39] Royston, P., Lambert, P.C. (2011), *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*. College Station, TX, USA: Stata Press.

[40] Simón-Moya, V., Revuelto-Taboada, L., Ribeiro-Soriano, D. (2016), *Influence of economic crisis on new SME survival: reality or fiction? Entrepreneurship & Regional Development*, 28, 157-176, https://doi: 10.1080/08985626.2015.1118560.

[41] Sun, J., Fujita, H., Zheng, Y., Ai, W. (2021), *Multi-class financial distress prediction based on support vector machines integrated with the decomposition and fusion methods. Information Sciences*, 559, 153-170, https://doi: 10.1016/j.ins.2021.01.059.

[42] Zivich, P.N., Cole, S.R., Shook-Sa, B.E., De-Monte, J.B., Edwards, J.K. (2025), *Estimating Equations for Survival Analysis with Pooled Logistic Regression*. arXiv preprint, https://doi.org/10.48550/arXiv.2504.13291.

[43] Zoričák, M., Gnip, P., Drotár, P., Gazda, V. (2020), *Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. Economic Modelling*, 84, 165-176, https://doi: 10.1016/j.econmod.2019.04.003.