

Ledwon, Andreas V.; Jäger, Clemens C.

Article

Cox proportional hazards regression analysis to assess default risk of German-listed companies with industry grouping

ACRN Journal of Finance and Risk Perspectives (JOFRP)

Provided in Cooperation with:

ACRN Oxford Research Network, Oxford

Suggested Citation: Ledwon, Andreas V.; Jäger, Clemens C. (2020) : Cox proportional hazards regression analysis to assess default risk of German-listed companies with industry grouping, ACRN Journal of Finance and Risk Perspectives (JOFRP), ISSN 2305-7394, ACRN Oxford Research Network, Oxford, Vol. 9, Iss. 1, pp. 57-77,
<https://doi.org/10.35944/jofrp.2020.9.1.005>

This Version is available at:

<https://hdl.handle.net/10419/329591>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>



Contents lists available at SCOPUS

ACRN Journal of Finance and Risk Perspectives

journal homepage: <http://www.acrn-journals.eu/>

Cox Proportional Hazards Regression Analysis to assess Default Risk of German-listed Companies with Industry Grouping

Andreas V. Ledwon^{*,1}, Clemens C. Jäger²¹FOM University of Applied Sciences for Economics and Management, Germany & UCAM - Catholic University of Murcia, Spain²FOM University of Applied Sciences for Economics and Management, Germany

ARTICLE INFO

Article history:

Received 29 March 2020

Revised 22 April 2020

Accepted 4 May 2020

Published 27 May 2020

Keywords:

Insolvency

Default prediction

Cox proportional hazards regression

AG-CP

Recursive AUC

Walk-forward analysis

Industry grouping

ABSTRACT

This study evaluates three corporate failure prediction models using latest available data on corporate insolvencies for non-financial constituents represented in CDAX. We estimate semiparametric Cox proportional hazards models considering Andersen-Gill counting process (AG-CP) to explore the importance of accounting and financial ratios as well as industry effects that are useful in detecting potential insolvencies. The contribution of this paper is twofold. First, the literature on corporate default prediction is manifold and predominantly focused on U.S. data. Thus, academic contribution based on German-listed companies is limited. To our best knowledge, we are the first to conduct thorough comparative out-of-sample Cox regression models considering AG-CP based on a unique dataset for non-financial constituents subject to the German insolvency statute ("InsO"). Relying on a parsimonious accounting-based approach inspired by Altman (1968) and Ohlson (1980) is merely adequate. Shumway (2001) and Campbell et al. (2008) variable selection delivers the best discriminatory power and calibration results. In particular, a combination of pure accounting ratios augmented with market-driven information in Model (2) indicates superior accuracy rates in top deciles. However, in-sample empirical results underpin the importance towards market-based indicators, as all accounting ratios enter statistically insignificant. Secondly, we test to what extent industry variables improve discriminatory power and forecasting accuracy of fitted models. Contrary to the findings of Chava & Jarrow (2004), our research implies that industry grouping adds marginal predictive power and no overall improvement in accuracy rates when market variables are already included in the probability of default (PD) model.

Introduction

The prediction of corporate insolvency and the assessment of credit risk have been the subject of much academic and professional research over the last decades (Altman (1968), Ohlson (1980), Shumway (2001) and Campbell et al. (2008)). Although insolvency is a relatively rare event among firms publicly-traded in Germany, the need to understand its characteristics is essential not only for academics but also managers (Creditreform Wirtschaftsforschung, 2018). According to the German Stock Corporation Act (*Aktiengesetz*, hereinafter referred to as "AktG"), "the management board shall take suitable measures in particular surveillance measures to ensure that developments threatening the continuation of the company are detected in processes at an early stage" (*Aktiengesetz* (AktG), 2017, sec. 91 (2)). This paper investigates and tests three corporate failure prediction models using latest available data on corporate insolvencies for German-listed non-financial companies. We apply Cox proportional hazards regression models to explore the importance of accounting and financial ratios as well as industry effects as indicators of default for non-financial German-listed companies. All well-recognized models are predominantly

* Corresponding author.

E-Mail address: andreas.ledwon@fom-net.de

ORCID: 0000-0002-4406-8280

developed and tested on U.S. data using logit regression. Thus, we aim to fill this research gap and provide new evidence on insolvency prediction for German-listed companies by assessing out-of-sample adequacy. We account for the first unified insolvency statute in Germany (*Insolvenzordnung*, hereinafter referred to as “InsO”) and test to what extent industry variables improve discriminatory power of fitted models.

Our paper differs from previous research in terms of applied methodology and sample selection. First, Cox proportional hazards regression has been identified as promising methodology to analyze default prediction models for non-financial German-listed firms represented in CDAX. It offers an advanced approach for unbalanced panel data and provides $\exp(\text{coef})$, known as hazard ratios, which are straightforward and easy to interpret. As the traditional Cox proportional hazards model is limited in terms of a single survival time value for each firm with possible censoring, we follow Andersen & Gill (1982) by extending our model that the data for a firm is presented as multiple observations, each of which applies to an interval of observations, referred to as “start stop data” (Box-Steffensmeier & Jones, 2004, p. 98). Hence, the inclusion of the Andersen-Gill counting process (AG-CP) offers to incorporate time-dependent covariates which are indispensable in predicting insolvencies. With regard to sample selection, Mertens et al. (2018) test insolvencies in the period 1991-2015, whereas the earliest year for identifying insolvency in our paper is set to the year 2000 since InsO entered into force in 1999. Accounting for a unified code of insolvency is in line with sample selection methodology suggested by Hillegeist et al. (2004) which accounted for statutory changes in the U.S. Our sample data spans from 2000 to 2018. Thus, a long time horizon allows to analyze and validate low default rates (Sobehart et al., 2000). A rigorously collected sample enables statistically significant out-of-sample results following the state-of-the-art walk-forward procedure by Sobehart et al. (2000). In addition, we employ time-dependent AUC analysis suggested by Chambless & Diao (2006) to assess discriminatory power of presented models. To our best knowledge, we are the first to conduct a comprehensive comparative study of Cox proportional hazards models for non-financial German-listed companies in CDAX and performing time-dependent AUC analysis. Surprisingly, not much attention has been paid to industry effects in related academic literature so far. Following economic intuition, the inclusion of an industry grouping should improve discriminatory power and accuracy rates of fitted models. Divergent level of competition among industries as well as different accounting conventions and regulatory requirements should impact the likelihood of insolvency cases. It is of particular interest to analyze if and to what extent industry variables improve discriminatory power and forecast accuracy of fitted models of accounting-based and market-based indicators.

The remainder is structured as follows. Section 2 provides a literature overview focusing on the relevant streams of academic contributions. Section 3 introduces Cox proportional hazards regression for time-dependent covariates as part of semiparametric survival analysis and presents adequate model performance measures. Our sample is introduced in section 4. Empirical results and performance evaluation results based on time-dependent area under the ROC curve, predicted box plots, and out-of-sample out-of-time calibration are presented in section 5. Section 6 discusses the results of the study. Finally, section 7 presents conclusions and limitations along with recommendations for academics and practitioners alike.

Relation to Literature

A significant body of work has been dedicated to corporate distress and insolvency. Among these, Beaver (1966) lays empirical foundation in insolvency prediction with an univariate discriminant analysis. However, univariate analysis fails to provide information about which ratios are the most important in detecting bankruptcy potential, the weights that should be attached to the selected ratios, and how weights should be objectively established (Altman, 1968). Hence, Altman (1968) assessed the quality of the aforementioned ratio analysis. As a result, the well-known Z-Score was developed by performing Multiple Discriminant Analysis (MDA) composing working capital/total assets (*WCTA*), retained earnings/total assets (*RETA*), EBIT/total assets (*EBITTA*), market value of equity/book value of total liabilities (*METL*), and sales/total assets (*STA*). Multiple applications of MDA have been performed inter alia by Edmister (1972), Blum (1974), Taffler & Tisshaw (1977), Taffler (1983). In summary, the conducted MDA models focused mainly on non-financial firms until the 1980s. Moreover, dominant samples included U.S. corporates and only a minority of studies investigated UK data (Taffler & Tisshaw, 1977) or small and medium-sized enterprises (SMEs) defaults (Edmister, 1972). Relevant studies on the performance of probability of default (PD) models utilizing MDA approach for German firms have been initially estimated by Perlitz (1973) who performed one of the first MDA analysis for German-listed companies. Subsequently, Baetge et al. (1987) contributed MDA for German firms with a set of three identified variables capturing a firm’s capital structure, profitability and solvency.

As a reaction to the dominant usage of MDA in the 1970s, Ohlson (1980) and Zmijewski (1984) note several econometric issues associated with MDA. First, statistical requirements such as comparable variance-covariance matrices and normally distributed predictors for both groups, of failed and non-failed firms, are violated in MDA. Secondly, the output of an MDA model is a score with little intuitive interpretation. Lastly, the matching procedure typically used in MDA default prediction has been criticized to be arbitrary and hence falls within the topic of choice-based sample biases. Ohlson (1980) and Zmijewski (1984) laid foundation for an alternative, less restrictive model to assess the probability of bankruptcy. Utilizing maximum-likelihood optimization of a logit function, or a probit function in the case of Zmijewski (1984), the authors estimate the coefficients of predictor variables. Ohlson (1980) proposes the use of conditional logit analysis to avoid the discussed problems with respect to MDA. In context of proposed predictors, no new ratios have been developed. Nine independent variables have been chosen based on previous literature (Ohlson, 1980). Behr & Güttler (2007) construct a logit scoring model ranging between 1992 and 2002 for estimating PD of German SMEs using a unique data set on SME loans in Germany to foster knowledge about their default risk and apply adequate cost of debt.

Although Altman & Hotchkiss (2006) conclude that from a statistical point of view logit regression seems to fit well with the characteristics of the default prediction problem where the dependent variable is binary, one has to distinguish between single-period classification models, also named static models, and survival analysis. A contemporaneous trend in the bankruptcy prediction literature is the utilization of hazard models, where, contrary to static models, the time-to-default of a firm is captured, and hence more firm-year observations are incorporated to explain bankruptcy (Zhang et al., 2018). In current literature, the term hazard model is used interchangeably with relating terms such as panel logit model, pooled logit model, or Cox proportional hazards regression analysis with time-varying covariates. All these aforementioned terms may be categorized within the field of survival analysis (Rodríguez, 2007a; Ruppert et al., 2003). Three issues related to static logistic regression models will be noted. Firstly, selection bias and the missing control for each firm's period at risk is criticized (Shumway, 2001). According to Shumway (2001), "static models ignore the fact that firms change through time, and hence produce bankruptcy probabilities that are biased and inconsistent estimates of the probabilities that they approximate" (Shumway, 2001, p. 101). In other words, it is crucial to account for the fact that some companies open insolvency proceedings after a few years of being in distress while other firms fail shortly after foundation. Hence, a static model approach does not consider duration dependence and firm age (Shumway, 2001). The third issue related to static models is the inefficiency of their out-of-sample forecasts. While hazard models include time series of available annual observations for each firm, static models incorporate only the last observation for each firm resulting in-sample selection bias, which comes from using only one, non-randomly selected observation per bankrupt firm (Hillegeist et al., 2004; Shumway, 2001).

Based on Kiefer (1988) and Lancaster (1990), Shumway (2001) defines a hazard model as "a binary logit model that includes all available information to determine each firm's bankruptcy risk at each firm year as a separate observation in time" (Shumway, 2001, p. 102). Each firm-year observation is included in the sample as long as a firm did not fail. Thus, bankrupt firms contribute only one failure observation to the logit model (Shumway, 2001). All firms that filed for any type of bankruptcy within 5 years of delisting are considered bankrupt in the respective analysis. Shumway's (2001) final sample contains 300 U.S. bankruptcies out of 3,182 firms between 1962 and 1992. Shumway (2001) estimates his models based on independent variables from prior studies, such as the forecasting models of Altman (1968) and Zmijewski (1984). However, compared to static models, computed hazard models produce divergent statistical inferences as half of well-established utilized accounting ratios to forecast bankruptcy in previous studies are not statistically related to failure. Consequently, Shumway (2001) introduces new market-driven independent variables represented by relative size (*RSIZE*), past excess returns (*EXRET*) and idiosyncratic standard deviation of each firm's stock returns (*SIGMA*). In result, the most accurate forecasts are generated with a hazard model that uses both market-driven and accounting variables to identify bankrupt firms resulting in 75% accuracy in the top decile of firms ranked annually by bankruptcy probability (Shumway, 2001).

Chava & Jarrow (2004) test the forecasting accuracy of bankruptcy hazard rate models for U.S. companies over the time period from 1962 to 1999 with yearly and monthly observation intervals and accounting for industry effects. Utilizing an expanded bankruptcy database including 1,461 bankruptcies, a superior forecasting performance of Shumway's (2001) model has been confirmed when compared to Altman (1968) and Zmijewski (1984) (Chava & Jarrow, 2004). Finally, Chava & Jarrow (2004) conclude that accounting variables add little predictive power when market variables are already included in the bankruptcy model. In addition, the authors argue that industry grouping significantly improve forecasting results as industry variables are statistically significant in-sample. However, the addition of industry effects increases the reported AUC ratio only marginally for non-financial public firms from 0.90 to 0.91 (Chava & Jarrow, 2004).

Beaver et al. (2005) apply hazard models following empirical suggestions by Shumway (2001) to forecast bankruptcy. A parsimonious model is applied capturing three key elements of the financial strength of a firm. Firstly, return on total assets measure takes profitability into account which has been proven to be an essential factor in previous empirical contributions. The second factor refers to Beaver's (1966) best single predictor cash-flow to total debt, which is in this analysis defined as net income to total assets (*NITA*). The third element is represented by the leverage ratio, total liabilities relative to total assets, and captured in this study by *TLTA* (Beaver et al., 2005). A key difference to Shumway's (2001) approach is that the identification and implementation of accounting-based variables have predictive power. Empirical results of Beaver et al. (2005) show that the models retain robustness of predictive power over time, while observing a slight decline in the predictive ability of accounting-based variables which is offset by improvement in the incremental predictive ability of market-related variables (Beaver et al., 2005).

Campbell et al. (2008) use hazard model methodology as developed by Shumway (2001) and applied by Chava & Jarrow (2004). The model obtaining the best fit includes both market-based and accounting data. First, new variables are introduced. The conventional way of measuring total assets is based on book value. However, Campbell et al. (2008) measure the equity component of total assets at market value by adding the book value of liabilities to the market value of equities referred to as *NIMTA*. First, the authors argue that *NIMTA* has stronger explanatory power, as market prices may include new information about the firm's prospects in more efficient and accurate manner (Campbell et al., 2008). Second, a measure of leverage is added by calculating total liabilities relative to total assets. Again, a market-valued version of this accounting series, defined as total liabilities divided by the sum of market equity and book liabilities, referred to as *TLMTA*, performs better than the traditional book-valued series (Campbell et al., 2008). Third, Campbell et al. (2008) add the ratio of cash and short-term investments to the market value of total assets, *CASHMTA*, in order to account for the liquidity position of a firm. Fourth, the market-to-book ratio, *MB*, takes the relative value placed on the firm's market equity to book value of equity into account. Fifth, the end of period log price per share of the firm, *PRICE* is included and truncated above USD 15 (Campbell et al., 2008). Finally, *EXRETAVG* is calculated as a stock excess return against the benchmark index while applying geometrically declining weights and $\phi = 2^{-\frac{1}{3}}$ in line with the literature (Campbell et al., 2008). In summary, Campbell et al. (2008) demonstrate a reduced-form econometric model to predict corporate bankruptcies and failures at short and long horizons. The applied model has greater explanatory power than the existing state-of-the-art models estimated by Shumway (2001) and Chava & Jarrow (2004) and includes the aforementioned additional variables with sensible economic motivation.

Mertens et al. (2018) test in a comparative study the structural Merton distance-to-default, Altman's (1968) Z-score as well as Campbell et al. (2008) hazard model. Various performance evaluation tools, including receiver-operating-characteristics analysis (ROC), calibration tests and a loan market simulation, are used, suggesting that the Campbell et al. (2008) model outperforms the other models fitting logit regressions. Although the performance evaluation metrics underpin that the failure score performs slightly worse when compared to U.S. data, the authors use it as a benchmark default risk model. Moreover, Mertens et al. (2018) do not recommend the application of Altman's (1968) Z-score and the distance-to-default approach as "the former has very weak discriminatory power and the latter is severely miscalibrated" (Mertens et al., 2018, p. 29).

Methodology

Cox Proportional Hazards Regression Model

A static logistic regression investigates responses by looking at one of only two possible values representing success and failure or, more generally speaking, by detecting the presence or absence of an attribute of interest (Rodríguez, 2007b). However, a static logistic regression does not provide information how the presence or absence of an attribute of interest is associated with intrinsic time effects of a well-defined event (Rodríguez, 2007a). In addition, censored observations are disregarded, meaning that for some units the event of interest has not occurred at the time the data is analyzed. Finally, the effect of explanatory variables in relation to its survival time is not taken into account and statistically assessed (Rodríguez, 2007a). Hence, static logistic regression is less effective in assessing the default risk predictors.

An alternative and promising approach, which has been primarily used in biostatistics and lately in financial credit risk literature, constitutes survival analysis. Survival analysis is used to analyze data in which the time until the event is of interest (Columbia University, 2004). The response is often referred to as a failure time, survival time, or event time (Columbia University, 2004). For simplification and standardization, the terminology "survival analysis" is

utilized, referring to the event of interest as “default” and to the waiting time as “survival time”. The Cox proportional hazards regression model comprises both a non-parametric aspect in the sense that it involves an unspecified function in the form of an arbitrary baseline hazard function, denoted as $h_0(t)$, and parametric model characteristics, as it allows modeling of the relationship between the failure rate and explanatory covariates (Cox, 1972). Hence, this type of methodology is often referred to as semiparametric model (Kalbfleisch & Prentice, 2002). The general form of the Cox proportional hazards regression model can be written as follows:

$$h(t, \mathbf{X}) = h_0(t) \exp \left[\sum_{i=1}^p \beta_i X_i \right], \quad (1)$$

where $h(t, \mathbf{X})$ is the expected hazard at time t for a firm with a given specification of a set of explanatory variables denoted by the bold $\mathbf{X} = (X_1, X_2, \dots, X_p)$. The baseline hazard function is called $h_0(t)$ and $\exp \left[\sum_{i=1}^p \beta_i X_i \right]$ represents the relative hazards, where the sum is over the p explanatory time-independent covariates \mathbf{X} .

As time-independent variables remain constant over time for a given firm, our model setup considers both time-independent and time-dependent variables whose value differ over time. Hence, an extended version of the Cox model considering AG-CP is required. In our study, modeling survival analysis presupposes to set a discrete time scale to mirror the occurrence of default defined as intervals (Hosmer et al., 2011). Although insolvency dates have been tracked as exact dates, annual data intervals have been selected due to the following reasons. Discrete time of intervals of one year ensure comparability of accounting-, market-based- and macroeconomic indicators. Moreover, to enable comparability with the studies conducted, yearly intervals have been chosen to predict corporate default. Hence, the estimation of the survival function is based on 19 consecutive yearly intervals denoted as $[t_0; t_1]; [t_1; t_2]; \dots; [t_{18}; t_{19}]$. Every firm i joining the study at observation time t_0 is categorized as active. Andersen-Gill counting process (AG-CP) is applied as time marches onward. Thus, start/ stop intervals for each firm year observation are taken into account. As firms have multiple observations in our data setup, we account for correlation within each firm and utilize a cluster variance represented by the argument *cluster(IDENT)*.

The event variable default y_{it} is binary for firm i at time t assuming only two values coded as one or zero. If insolvency according to InsO sec. 17-19 is documented, a change of state occurs, i.e. $y_{it} = 1$ and the firm disappears from the sample in the year following the event. In addition, firms are removed from the study without filing for insolvency inter alia due to M&A activity, spin-off or squeeze-out. Likewise, a firm that survives to the last period t_{19} cannot have failed in previous periods and thus does not change its state from zero to one. This universal characteristic of survival data is known as right-censoring. Utilized accounting data are lagged as calendar years have been selected. In some cases, observable accounting data in the year prior to bankruptcy are not available and consequently substituted with the previous year's accounting data as suggested by Chava & Jarrow (2004). This measure attempts to illustrate that accounting-based information is available to the market at the time of estimation (Chava & Jarrow, 2004). Extending the Cox proportional hazards regression model to fit both time-independent and time-dependent explanatory covariates, we can write:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp \left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j(t) \right], \quad (2)$$

where $h(t, \mathbf{X}(t))$ is the expected hazard at time t for a firm with a given specification of a set of time-independent explanatory variables $\mathbf{X}(t) = (X_1, X_2, \dots, X_{p_1})$, denoted by X_i , and time-dependent explanatory variables $\mathbf{X}(t) = (X_1(t), X_2(t), \dots, X_{p_2}(t))$ denoted by $X_j(t)$. The baseline hazard function is called $h_0(t)$ multiplied by the exponential function $\exp \left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j(t) \right]$.

In addition, this paper applies Breslow's method to estimate the cumulative baseline hazard rate. In order to estimate a Cox proportional hazards regression model, an indicator of financial distress and a relevant set of

explanatory variables is required which are presented in section 4. Next, model performance measures will be presented.

Assessment of Model Discrimination, Calibration and Diagnostics

The receiver operating characteristic (ROC) curve is a statistical measure ensuring performance validation of a binary classifier while providing an accuracy rate (Krzanowski & Hand, 2009). When focusing on the research question in this paper, the binary classifier insolvency is affected by various independent variables which influence the validity of the model. Per definition, the ROC curve may be described as a plot of true positive rate (TPR) on the y-axis versus false positive rate (FPR) on the x-axis in relation to a cut-off threshold C (Krzanowski & Hand, 2009). The decision results in consideration of the cut-off value C affect a PD's model performance and its ROC curve shape. One may conclude that the larger the area under the ROC curve (AUC), the higher the discriminatory power of PD (Engelmann et al., 2003). Hence, AUC is a common method in the computation of validity criteria. The area AUC is 0.5 for a random model without discriminative power and is 1.0 for a perfect model (Engelmann et al., 2003). Consequently, a suitable AUC ratio is between 0.5 and 1.0 for any reasonable PD model. However, the traditional approach of AUC curve analysis considers the event (insolvency) status and marker value for a firm as fixed over time. Thus, the essential factor default is not considered in terms of discriminatory power. In consequence, companies that are financially stable in an early stage of the study may file for insolvency at a later stage due to longer study follow-up. Thus, an AUC ratio as a function of time is more appropriate. That said, this paper applies dynamic AUC for right-censored time-to-event data suggested by Chambless & Diao (2006). For this out-of-sample discrimination the data set is divided into a training and testing sample with 70% to 30% data partition. The *iauc* summary measure is given by the integral of AUC on $[1,19]$ weighted by the estimated probability density of the time-to-event outcome. As the primary aim of this paper is to assess the discriminatory power of tested PD models, box plots are utilized to visualize the sensitivity of predictions. As the survival probability for a subject is equal to $\exp(-\text{expected value})$, predicted PD is equal to $1 - \exp(-\text{expected value})$.

In light of the fact that insolvency is a relatively rare event among publicly traded German firms, a rigorous out-of-sample out-of-time calibration procedure, also known as walk-forward testing, is performed. The state-of-the-art walk-forward procedure according to Sobehart et al. (2000) provides realistic, reasonable and rigorous out-of-sample test of the forecasting ability of different prediction models. In a real-world application of a default forecasting model, "walk-forward testing provides a framework for generating statistics that allow researchers to test the predictive power of a model on data not used to fit it" (Stein, 2007, p. 94). The procedure will be descriptively introduced. Firstly, we select one year for simplification, for example 2002. We fit a Cox regression model using all the data available in or before the selected year. Once the model's form and parameters are established for the selected time period, predictions containing all firm year observations available during the following year 2003 are generated. The predictions of PDs per firm year observation are part of the result set. Next, we move the estimation window one year further, i.e. 2003. Hence, all data available in 2000 to 2003 is used to fit models. These fitted models are now used to predict estimated values for the year 2004, respectively. This process is repeated, adding new predictions to the result set each year. In a nutshell, the walk-forward testing is comprised of an expanding estimation window. Figure 1 demonstrates graphically the aforementioned procedure. Dark circles represent in-sample data and white circles represent out-of-sample out-of-time testing data. The testing results for each year of prediction are stored to create a result set which is used to analyze the performance of the model in more detail. According to Stein (2007), "this approach simulates, as closely as possible given the limitations of the data, the process by which the model will actually be used. Each year, the model is refit and used to predict the credit quality of firms one year hence" (Stein, 2007, p. 94). Besides a rigorous and realistic data calibration, practitioners benefit from applying walk-forward testing as reparametrized models on a periodical basis provide information on economic changes (Stein, 2007).



Figure 1. Schematic of the walk-forward testing approach

Source: representation based on (Sobehart et al., 2000, p. 9)

Finally, model diagnostics help to identify whether fitted Cox regression models in this paper adequately describe the illustrated results (Fox & Weisberg, 2018). Hence, three types of diagnostics are applied:

- (1) testing the proportional hazards assumption
- (2) detecting influential observations or outliers
- (3) examining nonlinearity

To assess various aspects of model diagnostics residuals are examined. A residual is calculated for each observation such as for each firm providing a measure of the difference between the actual value and predicted values (Fox & Weisberg, 2018). In particular, there are three main types of residuals, to check the aforementioned model assumptions. Firstly, scaled Schoenfeld residuals are performed and plotted to check the proportional hazards assumption, followed by deviance residuals to examine influential observations and outliers and concluded with Martingale-residuals to assess nonlinearity.

Sample

Earliest year for identifying insolvency has been set to 2000 since InsO entered into force in 1999. This approach is in line with sample selection methodology suggested by Hillegeist et al. (2004) which accounted for statutory changes in the U.S. The raw sample comprises active and dead firms of the German equity market from 2000 to 2018. However, extensive data cleansing had to be conducted in order to increase data efficiency. Suggestions provided by Ince & Porter (2006) and Brückner (2013) have been followed and expanded to the specific needs of conducted research. All German non-financial firms with equity listings at Deutsche Börse in Frankfurt and represented in CDAX in the time-period 2000 to 2018 have been investigated. Following Fama & French (1992), our sample

selection excludes financial constitutes from the sample universe because the high leverage which is normal for banks and the financial service sector dilutes estimation results. Yearly frequency has been chosen to allow for data comparison to previous studies. The cleansed sample consists of 488 firms through 19 discrete yearly time intervals leading to 6,622 firm-year observations. Through careful web searches the delisting causes of 184 inactive firms have been identified. 97 insolvency proceedings according to InsO sec. 17-19 have been reconstructed as depicted in table 1. Reasons for leaving the stock market are manifold. The remaining 87 inactive firms have left the stock market *inter alia* due to M&A activity, spin-off or squeeze-out. All exits have been tracked with the exact date and respective source. In case of insolvencies, the date of opening based on ad-hoc announcements had to be manually retrieved. In order to foster transparency, all main court decisions are published and can be researched on the internet on the online record of the Ministry of Justice (*Insolvenzordnung (InsO)*, 2018, sec. 9). In practice, detailed information such as insolvency court in charge as well as detailed contact details are necessary in order to retrieve respective data. Consequently, data could not be gathered in all cases from the website: <http://www.insolvenzbekanntmachungen.de>. In result, the website: <http://www.dgap.de> provides alternative input. In few cases, press releases have been utilized to determine exits. Table 1 provides descriptive statistics summarizing the properties of the introduced insolvency indicator.

Table 1. Sample statistics: total number of active firms and insolvency rate

Year	Active Firms	Insolvency Proceedings	Insolvency Rate (%)
2000	346	0	0.00
2001	356	6	1.69
2002	341	13	3.81
2003	339	1	0.29
2004	340	6	1.76
2005	347	3	0.86
2006	372	2	0.54
2007	381	3	0.79
2008	377	6	1.59
2009	360	11	3.06
2010	350	10	2.86
2011	347	3	0.86
2012	343	3	0.87
2013	328	10	3.05
2014	306	3	0.98
2015	303	5	1.65
2016	298	5	1.68
2017	300	3	1.00
2018	304	4	1.32
Full Sample	488	97	19.88

The total insolvency rate during the selected sample period was 19.88%. However, focusing on average annual insolvencies the rate fluctuates between 0.00% and 3.81% highlighting that insolvency is a relatively rare event among publicly traded German firms. The annual insolvency rates show fluctuation and correlation to the overall economy, with peaks during the abandoned segment “Neuer Markt” in 2002 (3.81%), where many of included firms changed voluntarily to the regulated market before filing for insolvency proceedings (Burghof & Hunger, 2004) and post-crisis effects in relation to the global financial crisis affecting insolvencies in 2009 (3.06%) and Eurozone crisis 2013 (3.05%) respectively.

We considered the well-known four-digit Standard Industrial Classification (SIC) as a uniform classification system retrieved from Thomson Reuters Datastream (TDS). The industries in the sample are grouped into 8 major divisions based on the SIC. A detailed overview is displayed in table 2. The SIC codes from 1800-1999 are not used and as outlined before SIC codes 6000-6799 (Finance, Insurance and Real Estate) have been excluded in our study.

The Manufacturing sector accounts for more than half of the insolvencies in our sample (50.52%), followed by the service sector (31.96%).

Table 2. Insolvencies sorted by SIC Code Division

SIC Codes	SIC-Division	Insolvency Proceedings	Insolvency Rate (%)
0100-0999	Agriculture, Forestry and Fishing	1	1.03
1000-1499	Mining	2	2.06
1500-1799	Construction	4	4.12
2000-3999	Manufacturing	49	50.52
4000-4999	Transportation, Communications, Electric, Gas and Sanitary Service	3	3.09
5000-5199	Wholesale Trade	3	3.09
5200-5999	Retail Trade	4	4.12
7000-8999	Services	31	31.96

The four-digit SIC code allows to cascade testing on a finer level as depicted in table 3. Hence, a more detailed view of dominant insolvencies on SIC-Industry level is presented. Insolvencies in the Manufacturing division are dominated by Industrial Machinery & Equipment (10 filings) and Electronic & Other Electric Equipment (13 filings) accounting for 10.31% and 13.40% respectively. The Business Service industry (21 filings) predominantly accounts for the insolvency proceedings with 21.65% within the SIC-Division Services.

Table 3. Insolvencies sorted by SIC Code Division and Industry

SIC Codes	SIC-Division/ SIC-Industry	Insolvency Proceedings	Insolvency Rate (%)
0100-0999	Agriculture, Forestry and Fishing	1	1.03
08	Forestry	1	1.03
1000-1499	Mining	2	2.06
10	Metal, Mining	1	1.03
12	Coal Mining	1	1.03
14	Nonmetallic Minerals, Except Fuels	0	0.00
1500-1799	Construction	4	4.12
15	General Building Contractors	3	3.09
16	Heavy Construction, Except Building	1	1.03
2000-3999	Manufacturing	49	50.52
22	Textile Mill Products	1	1.03
23	Apparel & Other Textile Products	4	4.12
24	Lumber & Wood Products	1	1.03
26	Paper & Allied Products	4	4.12
27	Printing & Publishing	3	3.09
28	Chemical & Allied Products	4	4.12
30	Rubber & Miscellaneous Plastics Products	2	2.06
32	Stone, Clay, & Glass Products	1	1.03
33	Primary Metal Industries	3	3.09
34	Fabricated Metal Products	2	2.06
35	Industrial Machinery & Equipment	10	10.31
36	Electronic & Other Electric Equipment	13	13.40
37	Transportation Equipment	1	1.03
4000-4999	Transportation, Communications, Electric, Gas and Sanitary Services	3	3.09
45	Transportation by Air	2	2.06
48	Communications	1	1.03

SIC Codes	SIC-Division/ SIC-Industry	Insolvency Proceedings	Insolvency Rate (%)
5000-5199	Wholesale Trade	3	3.09
50	Wholesale Trade – Durable Goods	3	3.09
5200-5999	Retail Trade	4	4.12
52	Building Materials & Gardening Supplies	1	1.03
56	Apparel & Accessory Stores	1	1.03
57	Furniture & Homefurnishing Stores	1	1.03
59	Miscellaneous Retail	1	1.03
7000-8999	Services	31	31.96
73	Business Services	21	21.65
78	Motion Pictures	5	5.15
79	Amusement & Recreation Services	1	1.03
87	Engineering & Management Services	4	4.12

Despite the fact that various academics studied corporate defaults (Altman (1968), Ohlson (1980), Shumway (2001) and Campbell et al. (2008)), to the best of our knowledge, insolvencies for non-financial German-listed constituents represented in CDAX have never been analyzed and evaluated comprehensively based on advanced Cox proportional hazards model and time-dependent AUC analysis and industry effects. Hence, we aim to fill this research gap and provide new evidence with a comparative study of well-established Cox regressions starting from accounting variables inspired by Altman (1968) and Ohlson (1980), augmented with market-based variables proposed by Shumway (2001) and Campbell et al. (2008) based on a unique up-to-date database considering the inception of InsO. The exogenous variables used in this study were significant predictors of corporate default in previous empirical research. Utilized variables are categorized into liquidity, solvency, profitability as well as other indicators. In addition, variables are grouped in reference to its data origin, namely accounting-based indicators, market-based ratios and market-based ratios augmented by macroeconomic indicators in order to test for accuracy rates. In total, 14 ratios have been constructed using yearly data from TDS.

Liquidity ratios, represented by *CASHTA* and *CASHMTA*, provide information on a company's ability to fulfill its due obligations paying attention on a short-term focus. Short-term liquidity is measured as cash holdings divided by a company's total assets. If a firm is not having a sufficient positive cash balance and is unable to refinance, it is an indicator for prospective default. The conventional way of measuring total assets is based on book value. However, Campbell et al. (2008) measure the equity component of total assets at market value by adding the book value of liabilities to the market value of equities referred to as *CASHMTA*. A high *CASHMTA* ratio indicates high liquidity and hence a lower probability of bankruptcy. Campbell et al. (2008) argue that augmented market-driven ratios have stronger explanatory power, as market prices may include new information about the firm's prospects in a more efficient and accurate way (Campbell et al., 2008). In a nutshell, a more accurate valuation of a company's total assets may be performed as market equity capitalization is available in real time and reflects recent news to the firm. In addition, it allows to reflect financing capacity either through equity issuance as well as the ability to secure short-term financing (Campbell et al., 2008).

Measures of solvency contain three computed ratios. A measure of leverage is added by calculating total liabilities relative to total assets, named as *TLTA*. Again, a market-driven version of this accounting series, defined as total liabilities divided by the sum of market equity and book liabilities, referred to as *TLMTA*, is presented. The third introduced market-driven variable is the idiosyncratic standard deviation of each firm's stock returns, denoted as *SIGMA*. Shumway (2001) argues that *SIGMA* is strongly related to bankruptcy, both statistically and logically, as firms with a high volatility of cash flows are more likely to be affected by the event of bankruptcy. In other words, *SIGMA* relates to operating leverage (Shumway, 2001). *SIGMA* is calculated as the annualized standard deviation of the residual of a daily regression to its benchmark index CDAX.

Net income or historical losses divided by a company's total assets, denoted as *NITA*, represents one of the profitability ratios used in the empirical analysis. This accounting ratio has been enriched by its market-driven indicator *NIMTA*. Furthermore, past excess return, *EXRET*, is calculated as a firm's past excess return in year $t - 1$ minus the value-weighted CDAX benchmark index return in year $t - 1$. Each firm's annual returns are calculated by cumulating monthly returns. When some of a firm's monthly returns are missing, the value-weighted benchmark index return is substituted for the missing returns (Shumway, 2001).

$$EXRET_{i,t} = \log(1 + R_{i,t}) - \log(1 + R_{CDAX,t}) \quad (3)$$

$EXRETAVG$ is calculated as a stock excess return against the CDAX while applying geometrically declining weights and $\phi = 2^{-\frac{1}{3}}$ in line with the literature (Campbell et al., 2008).

$$EXRETAVG_{i,t} = 1 \frac{1-\phi}{1-\phi^{12}} (EXRET_{t-1} + \dots + \phi^{11} \times EXRETAVG_{t-12}) \quad (4)$$

Finally, relative size, denoted as $RSIZE$, is calculated as the natural logarithm of each firm's market capitalization divided by the total size of CDAX.

Other variables considered for inclusion in our models are $PRICE$, accounting for tendency of distress companies trading at low share prices, the market-to-book ratio, MB , as correction factor for utilized market-driven variables in Model (3) and $YEAR$ as continuous variable controlling for varying insolvency rates in survival study time. $PRICE$ is computed as end of period log price per share of the firm, while including and truncating above EUR 15. Following economic intuition, distressed companies tend to have low share prices and hence declining value of equity, reflecting a corporate crisis. Previous research argued that variations above USD 15 do not seem to affect failure probability and thus the ratio is capped above EUR 15 (Campbell et al., 2008). Next, MB takes into account the relative value placed on the firm's market equity to the adjusted book value of equity ($BE_{adjusted}$). Following Campbell et al. (2008), book value of equity is adjusted by the difference between market capitalization and book value of equity to adjust mismeasured and too large values of MB as indicated in equation 5.

$$BE_{adjusted,i,t} = BE_{i,t} + 0.1(ME_{i,t} - BE_{i,t}) \quad (5)$$

Hence, MB acts as a correction factor, as the aforementioned variables are all measured based on market value (Campbell et al., 2008). If the book value is relevant, probability of insolvency increases with MB (Campbell et al., 2008). $YEAR$, accounts for varying insolvency rates for the given time span from 2000 to 2018 as continuous variable. Finally, industry grouping (IND) according to the four-digit SIC code is used to analyze industry effects. Table 4 summarizes computed exogenous variables in this paper while providing a short description and expected coefficient regression signs derived from the literature review.

Table 4. Definition of exogenous variables and expected signs

Variables	Category	Description	Expected sign of coefficient
Solvency			
$SIGMA$	Market	$SIGMA$ is computed as the annualized standard deviation of the residual of a daily regression against the benchmark index CDAX.	+
$TLMTA$	Market	$\frac{total\ liabilities_{i,t}}{market\ capitalization_{i,t} + total\ liabilities_{i,t}}$	–
$TLTA$	Accounting	$\frac{total\ liabilities_{i,t}}{total\ assets_{i,t}}$	–
Liquidity			
$CASHMTA$	Market	$\frac{cash\ \&\ short\ term\ investments_{i,t}}{market\ capitalization_{i,t} + total\ liabilities_{i,t}}$	–
$CASHTA$	Accounting	$\frac{cash\ \&\ short\ term\ investments_{i,t}}{total\ assets_{i,t}}$	–
Profitability			
$EXRET$	Market	$\log(1 + R_{i,t}) - \log(1 + R_{CDAX,t})$	–
$EXRETAVG$	Market	$\frac{1-\phi}{1-\phi^{12}} (EXRET_{t-1} + \dots + \phi^{11} EXRETAVG_{t-12})$	–

Variables	Category	Description	Expected sign of coefficient
<i>NIMTA</i>	<i>Market</i>	$\frac{EBITDA_{i,t}}{\text{market capitalization}_{i,t} + \text{total liabilities}_{i,t}}$	–
<i>NITA</i>	<i>Accounting</i>	$\frac{EBITDA_{i,t}}{\text{total assets}_{i,t}}$	–
<i>RSIZE</i>	<i>Market</i>	$\log\left(\frac{\text{market capitalization}_{i,t}}{\text{market capitalization}_{CDAX,t}}\right)$	–
Other variables			
<i>PRICE</i>	<i>Market</i>	$\min[\log(PRICE_{i,t}); \log(15)]$	–
<i>MB</i>	<i>Market</i>	$\frac{\text{market capitalization}_{i,t}}{\text{book value of equity}_{adjusted,i,t}}$	+
<i>YEAR</i>	<i>Macro</i>	<i>YEAR</i> , accounts for varying insolvency rates for the given time span from 2000 to 2018 as continuous variable.	–
<i>IND</i>	<i>Macro</i>	Categorical industry grouping according to the four-digit SIC code is used to analyze industry effects.	+ / –

Empirical Results and Evaluation

Before presenting empirical results, we discuss the selection and economic intuition of the exogenous covariates within the employed Cox regression models. In the first column, we construct and estimate a reduced-form Model (1) inspired by Altman (1968) and Ohlson (1980). In the second column, we follow Shumway (2001) and estimate a model comprising six variables: *NITA*, *TLTA*, *EXRET*, *SIGMA*, *RSIZE* and *YEAR*. Model (2) takes assets in the traditional way into account, using book values. In column three, we apply Campbell et al. (2008) approach in Model (3), while adjusting book values to market-driven variables as already explained in section 2 and adding *CASHMTA*, *EXRETAVG*, *PRICE* and *MB*, accordingly.

Model (1) confirms the economic intuition and negative expected coefficient sign for *TLTA* and *CASHMTA*, which enter statistically significant at the level of 0.01. Holding the other covariates constant, one unit increase in *TLTA* increases the hazard by a factor of 1.03, or 3%. In contrast to this minor effect, *CASHMTA* concludes a lower risk of insolvency with a provided hazard ratio of 0.08. Finally, *YEAR* concludes an increasing insolvency risk of 6% per annum for the given survival study time at the statistical level of 0.05.

All variables for Model (2) enter with expected signs, however in comparison to Model (1) the importance of market data is distinctly emphasized as all accounting variables enter insignificantly. *EXRET* provides evidence that a firm's past excess return is a strong insolvency predictor. *SIGMA* is strongly related to bankruptcy, both statistically and logically, as firms with a high volatility of returns are more likely to be affected by the event of filing for insolvency. *RSIZE* proves that increasing company size relative to the benchmark index reduces the risk of filing for insolvency. Model (2) underpins that with an increase in *YEAR*, constituents are less likely to be affected by insolvencies.

Following Campbell et al. (2008), Model (3) reconfirms the findings of Model (2) and underpins the statistical benefit of substituting accounting-based ratios by adjusted market-driven ratios. In addition, Model (2) and (3) show unambiguously the importance of past excess returns (*EXRET* and *EXRETAVG*) and volatility (*SIGMA*). Furthermore, the price (*PRICE*) as well as the relation between book value of equity and market capitalization (*MB*) play an essential role in default prediction as all aforementioned variables enter highly statistically significant. Results are presented in table 5. The applied Cox proportional-hazards regression models are fitted using the statistical programming language R considering AG-CP as well as a cluster variance represented by the argument *cluster(IDENT)* which accounts for correlation within each firm. In addition, this paper applies Breslow's method to estimate the cumulative baseline hazard rate. The presented standard errors are robust.

Table 5. In-sample results on entire sample without industry grouping

Exogenous variables	Model (1)	Model (2)	Model (3)
<i>NITA</i>	0.98 (0.94-1.02)	0.98 (0.93-1.03)	
<i>NIMTA</i>			0.87*** (0.82-0.92)
<i>TLTA</i>	1.03*** (1.01-1.05)	1.00 (0.97-1.02)	
<i>TLMTA</i>			1.51*** (1.15-1.98)
<i>CASHTA</i>	0.08*** (0.01-0.44)		
<i>CASHMTA</i>			0.80 (0.46-1.37)
<i>EXRET</i>		0.23*** (0.18-0.30)	
<i>EXRETAVG</i>			0.00*** (0.00-0.00)
<i>SIGMA</i>		1.93*** (1.50-2.49)	1.72*** (1.29-2.30)
<i>RSIZE</i>		0.52*** (0.39-0.70)	0.87 (0.61-1.24)
<i>PRICE</i>			0.44*** (0.25-0.77)
<i>MB</i>			1.00*** (1.00-1.00)
<i>YEAR</i>	1.06** (1.01-1.13)	0.85*** (0.76-0.95)	0.96 (0.89-1.04)
<i>n</i>	6,622	6,622	6,622
<i>Number of events</i>	97	97	97
<i>Concordance</i>	0.64 se = 0.03	0.97 se = 0.01	0.97 se = 0.01
<i>Likelihood ratio test</i>	20.35***	434.00***	387.90***
<i>Wald test</i>	53.75***	554.30***	509.60***
<i>Score (logrank) test</i>	24.26***	1,643.00***	1,339.00***
<i>Schoenfeld global test</i>	0.22	0.30	0.44

Note:

exp(coef) are presented for each variable.

*** p<0.01, ** p<0.05, * p<0.1

All results are based on the entire sample.

Lower .95 and upper .95 confidence intervals are provided in parentheses.

Next, we provide empirical results with categorial industry grouping according to SIC Division. First, no changes in coefficients of preceding variables are observed allowing to test industry effects. Contrary to Chava & Jarrow (2004), empirical in-sample results indicate that German insolvency prediction is merely influenced by the industry to which the company belongs, as the *IND* variable has proven to be significant in only in Model (2). In particular *IND* (2000 – 3999), *IND* (4000 – 4999) and *IND* (5000 – 5199) enter statistically significant at level of 0.05. The Service sector represented by *IND* (7000 – 8999) is statistically significant at the level of 0.01. However, focusing on reported Concordance, only minor improvements can be reported for Model (1) and Model (2) whereas Model (3) shows no change when industry grouping is added.

Table 6. In-sample results on entire sample with industry grouping

Exogenous variables	Model (1) with <i>IND</i>	Model (2) with <i>IND</i>	Model (3) with <i>IND</i>
<i>NITA</i>	0.99 (0.96-1.02)	0.98 (0.94-1.02)	
<i>NIMTA</i>			0.90** (0.81-0.99)
<i>TLTA</i>	1.04*** (1.02-1.05)	0.99 (0.97-1.02)	
<i>TLMTA</i>			1.56*** (1.20-2.04)
<i>CASHTA</i>	0.06*** (0.01-0.37)		
<i>CASHMTA</i>			0.94 (0.57-1.55)
<i>EXRET</i>		0.24*** (0.19-0.31)	
<i>EXRETAVG</i>			0.00*** (0.00-0.00)
<i>SIGMA</i>		1.98*** (1.54-2.56)	1.69*** (1.22-2.33)
<i>RSIZE</i>		0.47*** (0.34-0.65)	0.82 (0.57-1.17)
<i>PRICE</i>			0.41*** (0.24-0.68)
<i>MB</i>			1.00*** (1.00-1.00)
<i>YEAR</i>	1.07** (1.01-1.13)	0.82*** (0.74-0.92)	0.92* (0.84-1.01)
<i>IND (1000-1499)</i>	1.93 (0.31-11.95)	1.80 (0.40-7.99)	2.80 (0.47-16.70)
<i>IND (1500-1799)</i>	1.21 (0.20-7.54)	0.60 (0.17-2.12)	2.98* (0.87-10.19)
<i>IND (2000-3999)</i>	0.62 (1.01-1.13)	0.25** (0.08-0.81)	1.17 (0.40-3.44)
<i>IND (4000-4999)</i>	0.24 (0.04-1.63)	0.13** (0.02-0.89)	0.36 (0.06-2.22)
<i>IND (5000-5199)</i>	0.62 (0.09-4.49)	0.21** (0.05-0.88)	0.69 (0.19-2.53)
<i>IND (5200-5999)</i>	0.76 (0.12-4.68)	0.69 (0.19-2.46)	2.21 (0.62-7.79)
<i>IND (7000-8999)</i>	0.83 (0.17-4.07)	0.17*** (0.05-0.59)	0.58 (0.19-1.76)
<i>n</i>	6,622	6,622	6,622
<i>Number of events</i>	97	97	97
<i>Concordance</i>	se = 0.03	se = 0.01	se = 0.01
<i>Likelihood ratio test</i>	29.40***	448.00***	404.90***
<i>Wald test</i>	72.90***	579.70***	575.60***
<i>Score (logrank) test</i>	34.64***	1,651.00***	1,361.00***
<i>Schoenfeld global test</i>	0.36	0.81	0.34

Note:

exp(coef) are presented for each variable.

*** p<0.01, ** p<0.05, * p<0.1

All results are based on the entire sample.

Lower .95 and upper .95 confidence intervals are provided in parentheses.

The proportional hazards assumption is checked by assessing scaled Schoenfeld residuals global test and graphical diagnostics based on the identical methodology (Zhang et al., 2018). Results of Schoenfeld's global test conclude no violation of the proportional hazards assumption for fitted models with respective p-values of 0.22 and 0.36 for Model (1), 0.30 and 0.81 for Model (2), and 0.44 and 0.34 for Model (3). Graphical diagnostics of dfbeta, dfbetas, deviance

and Martingale residuals do not show severe influential observations or outliers as well as nonlinearity for fitted Model (2) and Model (3). Once more, weak discriminatory power of Model (1) is further confirmed through violation of influential observations or outliers as well as indicating nonlinearity. Finally, Goodness of fit of applied models is assessed by the likelihood ratio test, Wald test, and score (logrank) test. These three statistics are asymptotically equivalent. Statistical significance in Model (2) and Model (3) show significant improvement relative to the null. Overall, respective results present robust empirical evidence in explaining corporate insolvencies for German non-financial firms.

Area under the Receiver Operating Characteristic (ROC) curve, known as AUC, is used to measure the ability of survival models to predict future risk. In a nutshell, the AUC ratio ranges between 0 and 1. The baseline of 0.5 describes a total random model. It is crucial to mention that the estimated AUC ratio is adjusted for time-dependency and censoring as both aspects affect estimation results of survival curves and coefficients of survival regression analysis (Chambless & Diao, 2006). In order to evaluate the forecasting accuracy of employed Cox regression models, a recursive computation of dynamic AUC for right-censored time-to-event data suggested by Chambless & Diao (2006) is performed. For out-of-sample model discrimination, the data set is divided into a training and testing sample with 70% to 30% data partition, remembering company identification. The *iauc* summary measure is given by the integral of AUC on [1,19] weighted by the estimated probability density of the time-to-event outcome presented as illustrated below.

Table 7. AUC for right-censored time-to-event data

Time Point	AUC Model (1)	AUC Model (1) with <i>IND</i>	AUC Model (2)	AUC Model (2) with <i>IND</i>	AUC Model (3)	AUC Model (3) with <i>IND</i>
<i>iauc</i>	0.64	0.65	0.87	0.88	0.89	0.89

In figure 2 black lines represent the AUC for models without industry grouping, whereas red lines represent the AUC for models with industry grouping. Model (1) illustrates a weak discriminatory power of *iauc* of 0.64 without industry grouping. The inclusion of industry grouping marginally improves discriminatory power to 0.65. Model (2) confirms with an *iauc* of 0.87 high discriminatory power slightly increases when adding industry grouping (0.88). Model (3) shows superior results with superior *iauc* of 0.89. Summarizing, marginal improvement of industry grouping can be concluded when dealing with accounting variables and the mixed approach of accounting and market-based indicators suggested by Shumway (2001). Increasing discriminatory power of each model in each time interval can be observed. That said, the discriminatory power rises with growing time intervals. During the time of the abandoned segment “Neuer Markt” in 2002, all three models reveal shortcomings with forecasting accuracy of survival and insolvencies as many included firms changed voluntarily to the regulated market before filing for insolvency proceedings. Comparing our out-of sample results to Chava & Jarrow (2004), a marginal improvement of discriminatory power can be validated for non-financial German firms represented in CDAX.

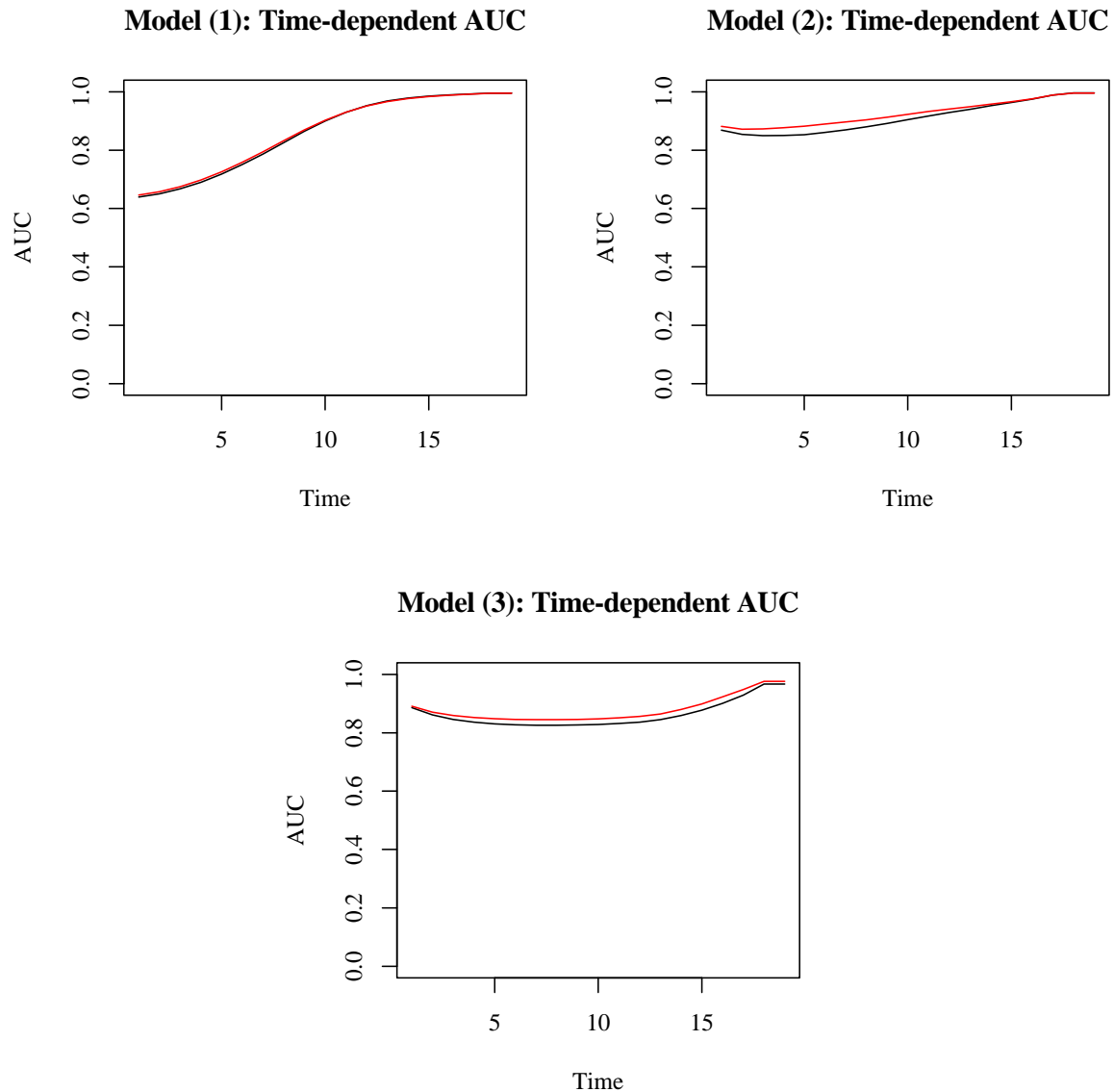


Figure 2. Time-dependent AUC Model (1) – (3)

As the purpose of this paper is the assessment of insolvency for non-financial German firms listed in CDAX and testing for industry effects, box plot analysis is applied to provide a better understanding of discriminatory power of predicted insolvency. The black bars represent the median, whereas the edges of each box represent the 25th and the 75th percentiles. In a nutshell, a well-performing model should show no overlapping of predictions of the binary coded event variable insolvency. That said, the weak discriminatory power of Model (1) is reconfirmed when compared to Model (2) and (3). Taking industry grouping into account does not show extensive shifts of boxplots and predictions.

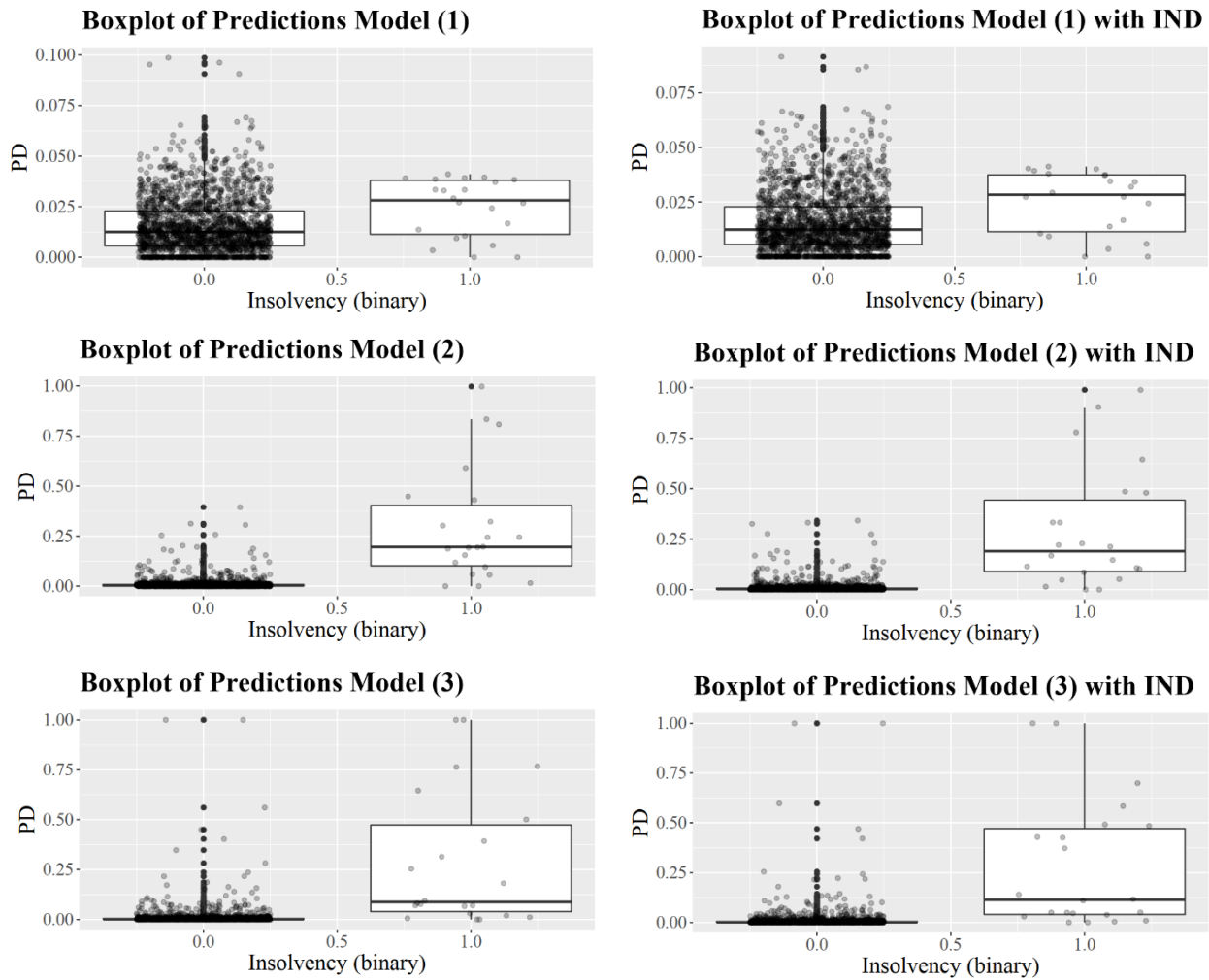


Figure 3. Boxplots of Predictions for Model (1) – (3)

After presenting out-of-sample tests to assess discriminatory power, state-of-the-art walk-forward analysis suggested by Sobehart et al. (2000) for out-of-sample out-of-time prediction is applied to perform rigorous model calibration. The result set of the expanding estimations is utilized to perform model calibration. A well calibrated PD model should provide high accuracy rates (AR) in top deciles. Following Shumway (2001) and Chava & Jarrow (2004), the firms with the highest probability of default in each year are placed into the first decile in descending order. Secondly, year by year, beginning from 2002, we count the number of firms in each decile that historically filed for insolvency. The number of firms in each decile that filed for insolvency across the study from 2002 until 2018 are aggregated in the result set and for each decile the percentage of bankrupt companies occurring in that decile are reported in table 8.

Table 8. Model calibration and decile ranking

Decile	AR Model (1) (%)	AR Model (1) with <i>IND</i> (%)	AR Model (2) (%)	AR Model (2) with <i>IND</i> (%)	AR Model (3) (%)	AR Model (3) with <i>IND</i> (%)
1	25.27	25.27	97.80	97.80	93.41	90.11
2	18.68	17.58	1.10	1.10	3.30	6.59
3	18.68	12.09	0.00	0.00	1.10	2.20
4	8.79	19.78	0.00	0.00	1.10	0.00
5	7.69	7.69	0.00	0.00	0.00	0.00
6	4.40	3.30	0.00	0.00	0.00	0.00
7	5.49	4.40	0.00	0.00	0.00	0.00
8	2.20	3.30	1.10	1.10	0.00	0.00
9	6.59	4.40	0.00	0.00	1.10	0.00
10	2.20	2.20	0.00	0.00	0.00	1.10

Model (1) shows severe miscalibration. In the first decile only 25.27% of insolvencies are correctly identified. When focusing on insolvent companies above the probability median, 79.12% of insolvencies can be estimated. Including industry grouping, we observe an increase in accuracy rates of 3.30%, achieving 82.42% in the top 5 deciles. Nonetheless, Model (1) ranks 20.88% and 17.58% of estimated insolvencies within deciles 6-10. Hence, out-of-sample calibration reassures the weak discriminatory power. Following Shumway (2001) variable selection of accounting-based and market-driven covariates in Model (2) leads to 97.80% correct predictions in the first decile with and without industry grouping and 98.90% in top deciles. According to the market-driven approach by Campbell et al. (2008), Model (3) confirms high accuracy rates in the top decile of 93.41% and 90.11%. Within top 5 deciles, Model (3) confirms with and without industry grouping accuracy rates of 98.90%, too. However, no improvement of calibration results is present when including industry grouping. Comparing our results to U.S. studies from Shumway (2001), Chava & Jarrow (2004) and Campbell et al. (2008) as well as to German findings by Mertens et al. (2018), satisfactory calibration results in top deciles can be concluded for Model (2) and Model (3).

Discussion

In this study we employ and test several Cox proportional hazards models for German-listed companies with a newly compiled sample of active and insolvent non-financial entities considering the inception of the current insolvency statute. Based on the previous academic contribution, we focused on applying a new methodology, namely Cox proportional hazards regression models considering the Andersen-Gill counting process (AG-CP), to account for right-censoring and time-dependent covariates. Model (2) and (3) with market-based variables proposed by Shumway (2001) and Campbell et al. (2008) proved to provide most accurate results in terms of discriminatory power and model calibration.

First, relying on a parsimonious accounting-based approach is merely adequate. Exogenous variables presented, inter alia by Altman (1968) and Ohlson (1980) show the worst discriminatory power in our analysis represented in Model (1). The performed walk-forward testing reconfirms a severe miscalibration. Moreover, nonlinearity has been detected as part of model diagnostics.

Second, Shumway (2001) and Campbell et al. (2008) variable selection deliver the best results for German-listed companies with high *iauc* ratios. Whereas Mertens et al. (2018) propose to use Campbell et al. (2008) variable selection as benchmark model for Germany, our research is the first to test Shumway (2001) variable selection on the German market in Model (2). With such, we incorporate an approach taking into account traditional accounting-based figures measured at book value augmented with market-driven indicators. This allows us to draw more detailed conclusions about the importance of hybrid models. That said, a mixture of pure accounting ratios with market-driven information in Model (2) indicates superior performance in top deciles in out-of-sample calibration. However, in-sample empirical results underpin the importance towards market-based indicators, as all accounting ratios enter statistically insignificant.

Performing Cox proportional hazards regression shows an advanced approach to handle unbalanced panel data and offers straightforward and easy interpretation of hazard ratios. Hence, our analysis supports the recommendation of using Cox proportional hazards regression analysis considering AG-CP as a benchmark model for German-listed companies. In particular, we highlight the importance of time-dependent evaluation measures as our study captures a firm's time-to-default. Thus, time-dependent AUC analysis suggested by Chambless & Diao (2006) is used to ensure comparability and accounts for right-censoring.

In contrast to this research, Mertens et al. (2018) tested insolvencies in the period 1991-2015. Earliest year for identifying insolvency has been set to the year 2000 in our paper since InsO entered into force in 1999. This approach is in line with sample selection methodology suggested by Hillegeist et al. (2004) which accounted for statutory changes in the U.S. Our sample data spans from 2000 to 2018. This long time horizon allows to analyze and validate low default rates (Sobehart et al., 2000). Thus, a comparably smaller number of observations does not downgrade our empirical results as data cleansing measures aimed to proxy non-financial constitutes represented in CDAX and hence increased targeted accuracy which is highlighted in the out-of-sample *iauc* results.

Finally, our research indicates that German insolvency prediction is gradually influenced by the industry to which the company belongs when focusing on discriminatory power. Out of-sample tests do not show an overall improvement in model calibration. When including industry grouping, we observe an increase in accuracy rates in the top 5 deciles only for our accounting-based Model (1). The presented decile ranking indicates even decreasing calibration results in the top decile for Model (3). Contrary to the findings of Chava & Jarrow (2004), our research implies that industry grouping adds marginal predictive power and no overall improvement in accuracy rates when market variables are already included in the PD model.

Conclusion

Our study contributes to the growing body of literature by developing and applying Cox proportional hazards regression for time-dependent covariates to assess corporate insolvency for non-financial constitutes represented in CDAX. Our research concludes the following theoretical and managerial implications:

Accounting-based indicators and market-based ratios as well as industry effects are analyzed with regards to their usefulness in detecting potential insolvencies. In addition to the theory, we conclude that the hybrid approach of Shumway (2001) variable selection of pure accounting ratios and market-driven indicators provides most accurate model calibration. In terms of discriminatory power, we observe a marginal increase if book value of total assets is substituted with its market-valued version. Moreover, substantial industry effects are observed in Model (1) focusing on predictability which relies on accounting ratios. Hence, one may assume that most predictive power to forecast insolvencies lies within market-driven indicators. In summary, Model (2) and (3) proposed by Shumway (2001) and Campbell et al. (2008) proved to provide most accurate results in terms of discriminatory power and model calibration. Next, our newly compiled sample from the year 2000 to 2018 of active and insolvent non-financial entities considers the inception of the first unified insolvency statute in Germany. This sample selection period allows to report results which are not biased by this major regulatory change in 1999.

This research provides an opportunity for practitioners to explore the application of the extended Cox proportional hazards regression with AG-CP to predict corporate insolvency. In particular, our model setup may be expanded to recurring events related to insolvency proceedings and other corporate events. This study can be utilized by practitioners to compare expected rates of insolvencies to a set of peers as we account for industry grouping. Hence, one may validate its performance not only based on industry peers but also against external ratings. Finally, our approach offers a framework for generating statistics that allow practitioners to test the predictive power of a model on data not used to fit it. Practitioners benefit from the state-of-the-art walk-forward procedure according to Sobehart et al. (2000) as reparametrized models on a periodical basis provide information on economic changes in a realistic and reasonable way.

Although our findings are robust, further examination on German PD models may provide additional insights into German insolvency characteristics and predictability. Future research should investigate a tailored variable selection for the German market as our approach is predominantly based on U.S. literature. Finally, effects of the 2012 implemented act for further facilitation of the restructuring of companies (*Gesetz zur weiteren Erleichterung der Sanierung von Unternehmen*, known as "ESUG") should be investigated. According to Moraht & Lütcke (2012), the key aim of the legislative change lays in strengthening creditors' rights through earlier involvement and greater influence in the selection of the insolvency administrator. From a debtors' perspective the ESUG creates incentives to apply for the opening of insolvency proceedings at an early stage in order to enhance the chances of successfully

restructuring the company. Hence, self-administration has been strengthened, protective shield proceedings introduced, and the insolvency plan procedure streamlined.

References

- Aktiengesetz (AktG)*. (2017).
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609.
- Altman, E. I., & Hotchkiss, E. (2006). *Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt. Foundations and Trends® in Finance* (3rd ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Andersen, P. K., & Gill, R. D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics*, 10(4), 1100–1120.
- Baetge, J., Huß, M., & Niehaus, H.-J. (1987). Die Beurteilung der wirtschaftlichen Lage eines Unternehmens mit Hilfe der statistischen Jahresabschlußanalyse. *Betriebswirtschaftliche Steuerungs- und Kontrollprobleme* (pp. 19–31). Wiesbaden: Gabler Verlag.
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4, 71–111. [Accounting Research Center, Booth School of Business, University of Chicago, Wiley].
- Beaver, W., McNichols, M., & Rhie, J. W. (2005). Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies*, 10(1), 93–122.
- Behr, P., & Güttler, A. (2007). Credit Risk Assessment and Relationship Lending: An Empirical Analysis of German Small and Medium-Sized Enterprises. *Journal of Small Business Management*, 45(2), 194–213.
- Blum, M. (1974). Failing Company Discriminant Analysis. *Journal of Accounting Research*, 12(1), 1–25. [Accounting Research Center, Booth School of Business, University of Chicago, Wiley].
- Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event History Modeling: A Guide for Social Scientists (Analytical Methods for Social Research)* (1st ed.). Cambridge: Cambridge University Press.
- Brückner, R. (2013). *Three Essays on the German Capital Market*. Humboldt-Universität zu Berlin.
- Burghof, H.-P., & Hunger, A. (2004). *Access to Stock Markets for Small and Medium-Sized Growth Firms: The Temporary Success and Ultimate Failure of Germany's Neuer Markt*. SSRN.
- Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *Journal of Finance*, 63(6), 2899–2939.
- Chambless, L. E., & Diao, G. (2006). Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Statistics in Medicine*, 25, 3474–3486.
- Chava, S., & Jarrow, R. A. (2004). Bankruptcy Prediction With Industry Effects. *Review of Finance*, 8(4), 537–569.
- Columbia University. (2004). Slides: Lecture 15 Introduction to Survival Analysis. *BIOS 515 - Analysis*. Retrieved March 3, 2020, from <http://www.stat.columbia.edu/~madigan/W2025/notes/survival.pdf>
- Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187–220.
- Creditreform Wirtschaftsforschung. (2018). Insolvenzen in Deutschland, Jahr 2018.
- Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *Journal of Financial and Quantitative Analysis*, 7(02), 1477–1493.
- Engelmann, B., Hayden, E., & Tasche, D. (2003). Measuring the discriminative power of rating systems. *Banking and Financial Supervision*, 2(01/2003), 1–24.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427–465.
- Fox, J., & Weisberg, S. (2018). *An R Companion to Applied Regression* (3rd ed.). SAGE Publications, Inc.
- Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies*, 9, 5–34.
- Hosmer, D. W., Lemeshow, S., & May, S. (2011). *Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition*. *Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition*.
- Ince, O. S., & Porter, R. B. (2006). Individual equity return data from Thomson datastream: Handle with care! *Journal of Financial Research*.
- Insolvenzordnung (InsO)*. (2018).
- Kalbfleisch, J. D., & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Statistics. Wiley.
- Kiefer, N. (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, 26, 646–679.
- Krzanowski, W. J., & Hand, D. J. (2009). *ROC Curves for Continuous Data*. *ROC Curves for Continuous Data* (1st ed.). Chapman and Hall/CRC.
- Lancaster, T. (1990). *The Econometric Analysis of Transition Data* (1st ed.). Cambridge: Cambridge University Press.
- Mertens, R., Poddig, T., & Fieberg, C. (2018). Forecasting Corporate Defaults in the German Stock Market. *Journal of Risk*, 20(6), 29–54.
- Moraht, J., & Lütcke, N. (2012). The new ESUG Law for Further Facilitation of the Restructuring of Businesses. *CMS Update Banking & Finance*, (May 2012).

- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(Spring), 109–131.
- Perlitz, M. (1973). Die Prognosefähigkeit von Kennzahlen. *Die Prognose des Unternehmenswachstums aus Jahresabschlüssen deutscher Aktiengesellschaften* (pp. 65–165). Wiesbaden: Gabler Verlag.
- Rodríguez, G. (2007a). Lecture Notes for Survival Models. *Lecture Notes on Generalized Linear Models*. Retrieved April 29, 2019, from <https://data.princeton.edu/wws509/notes/c7.pdf>
- Rodríguez, G. (2007b). Lecture Notes for Logit Model for Binary Data. *Lecture Notes on Generalized Linear Models*. Retrieved April 29, 2019, from <https://data.princeton.edu/wws509/notes/c3.pdf>
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric Regression* (1st ed.). Cambridge: Cambridge University Press.
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74(1), 101–124.
- Sobehart, J., Keenan, S., & Stein, R. (2000). Benchmarking Quantitative Default Risk Models: A Validation Methodology. *Technical Report, Moody's Investor Service, Inc.*, (March 2000), 1–20.
- Stein, R. (2007). Benchmarking default prediction models: pitfalls and remedies in model validation. *The Journal of Risk Model Validation*, 1(1), 77–113.
- Taffler, R. J. (1983). The Assessment of Company Solvency and Performance Using a Statistical Model. *Accounting and Business Research*, 13(52), 295–308. Routledge.
- Taffler, R. J., & Tisshaw, H. J. (1977). Going Going gone four factors which predict. *Accountacy*, 88(1003), 50–54.
- 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. AME Publishing Company.
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research (Wiley-Blackwell)*, 22, 59–82.



© 2020 by the authors. Licensee ACRN Publishing, Austria, Editor in Chief Prof. Dr. Othmar M. Lehner. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>)