IN SEARCH OF A PARSIMONIOUS BANKRUPTCY MODEL FOR PRIVATE FIRMS

And the Cost to Lenders

JOHN-EDWARD OLINGSBERG† & OSCAR KÜNTZEL††

WINTER 2019

This paper comprehensively reviews in excess of 200 financial and non-financial covariates in search of a parsimonious bankruptcy prediction model for the private market. Financial and real-estate companies aside, the entire population of Swedish independent, limited liability companies are examined between 1998-2017 corresponding to 245,844 unique companies and 55,411 corporate default events. Predictors are atheoretically selected for one- to five-year forecasts and fitted as discrete-time duration-dependent hazard models with and without a frailty term. Each of these is cross-validated and tested against five competing covariate specifications: Altman's (1968) Z-score, Zmijewski (1984), Shumway (2001), Altman and Sabato (2007) and Dakovic et al. (2010). For the vast majority of prediction horizons, estimation windows and hold-out tests, we significantly outperform these contesting models along dimensions of model power, calibration and early-retrieval performance as measured by principally ROC curves, PR curves and their integrals. From our covariate vectors, we find evidence of intra-industry contagion at work for longer look-ahead predictions, suggest employees may be an underused denomination of company size and discover untaxed reserves as a highly promising indicator of bankruptcy in the context of Swedish Private Company Law. Finally, we gauge the practical merits of our empirically derived covariate specifications by backtesting all six one-year predictive models on simulated (equally-weighted) and actual private loans (contained in the data) during 2014-2016. The lender's perspective is adopted using the Basel III F-IRB approach. In both settings, the bank that assumes our bankruptcy hazard model makes considerably higher returns on risk-weighted assets than any of its competition.

Keywords: Private Default Prediction; Hazard Model; Basel III; Misclassification Costs; Parsimonious

JEL Classification Codes: G33; G34; G21; C55

---

† of the Stockholm School of Economics. Contact: 23323@student.hhs.se

†† of the Stockholm School of Economics. Contact: 23324@student.hhs.se

ACKNOWLEDGEMENTS

We extend our sincere gratitude to our supervisor Dr. Ramin Baghai for helpful feedback and input over the course of the paper. Further, we thank Dr. Kenth Skogsvik and Dr. Henrik Nilsson of the Stockholm School of Economics, Dr. Vineet Agarwal of the Cranfield School of Management, Dr. Jairaj Gupta of the University of Birmingham, Dr. Panagiotis Andrikopoulos of Coventry University, Dr. Thomas Oommen of Michigan Technological University, Miqaela Hill and Per Weidenman of Bisnode, and Olle Engdegård, Head of Predictive Modelling at UC.

Table of Contents

List of Tables and Figures

TABLES

FIGURES

List of Abbreviations

AFT - Accelerated Failure Time

AIC - Akaika Information Criterion

AME - Average Marginal Effect

AR - Accuracy Ratio

AUPRC - Area Under the Precision-Recall Curve

AUROC - Area Under the Receiver Operating Characteristics Curve

BIC - Bayesian Information Criterion

Bps – Basis Points

CAP - Cumulative Accuracy Profile

CCJ - County Court Judgement

CM_1 - CM_5 - The Composite Model for the One- to Five-Year Prediction Horizons

DTDDH - Discrete-Time Duration-Dependent Hazard

EAD - Exposure at Default (Notional Loan Value)

EBIT/DA - Earnings before Interest, Taxes / Depreciation and Amortisation

F-IRB - Foundation Internal Ratings-Based

FPR - False Positive Rate

GAM - Generalized Additive Model

LGD - Loss Given Default

MDA - Multiple Discriminant Analysis

NPV - Net Present Value

PD - Probability of Default

PH - Proportional Hazard

PPV - Positive Predictive Value (Precision)

PR - Precision-Recall

PR2 - Pseudo R-Squared

ROA - Return on Assets

ROC - Receiver Operating Characteristic

RORWA - Return on Risk-Weighted Assets

ROSE - Random Oversampling Examples

RWA - Risk-Weighted Assets

SEK - Swedish Crown (Krona)

SME - Small- and Medium-Sized Enterprises

SMOTE - Synthetic Minority Oversampling Technique

SNI - Swedish Standard Industrial Classification

TPR - True Positive Rate

UC - UC AB, A Private Business- and Credit-Information Company

1. Introduction

"Science has not yet mastered prophecy. We predict too much for the next year and yet far too little for the next 10". Although the late Neil Armstrong was addressing a joint congressional session in 1969 following mankind's first lunar exploits, his verbatim continues to ring true some odd 50 years later (National Academy of Engineering, 1993). Today, a five-day weather forecast, for instance, is accurate roughly 90% of the time. Extending the horizon to ten days or longer, however, sees accurate prognoses at a rate no different than predicting a (repeated) unbiased coin toss (NOAA, 2019). The same can be said of corporate *default predictions*. With its modern seminal inception a year prior to Apollo 11, Altman (1968), is able to identify over 90% of (sampled) defaulting public manufacturers between 1946-1965 one year in advance based on a parsimonious transformation of financial ratios into a '*Z-score*' index. Looking ahead three years or longer, however, his positive hit-rate using the same inputs falls well below the 50% mark[1].

Admittedly an anecdotal piece of evidence, science has certainly not yet mastered prophecy. But that is not to say that science - both social and natural - has not markedly improved its predictive capacity in the five decades that have now elapsed since Apollo 11 and Altman's influential paper using *Multiple Discriminant Analysis*[2]. Indeed, a modern five-day weather forecast is as accurate as a one-day forecast was in 1980. 72-hour predictions of hurricane tracks are today more accurate than 24-hour forecasts were ca 40 years ago (Alley et al., 2019). Unfortunately, the degree of improvement in default discriminant classification since the late sixties has been somewhat less remarkable, conceivably in large part due to the high rate of change and innovation inherent to the social environment relative the climate and weather. This is perhaps best demonstrated by the continued use and reasonable success of Altman's MDA methodology and Z-score covariates to this day (Reisz and Perlich, 2007). Progress in early default-detection has been particularly gradual for forecasting-horizons exceeding one year. So much so, that finding respectable accuracy ratios for five-year or longer lead-times in a reputable scientific journal can be challenging[3]. The lion's share of the literature is predisposed to linger on one-year predictions yet. Even so, all prognoses specifications are not created equal. In the fall of 2012, the European Centre for Medium Range Weather Forecasting predicted Hurricane Sandy's devastation to the American East Coast two days prior to the U.S. National Weather Service (Miller, 2013). By way of modern modelling techniques and greater data accessibility, corporate default researchers are now regularly able to outperform Altman's Z-score[4] across different time-periods, geographies and industries for single-year look-ahead forecasts in the public company domain[5].

[1] These classification accuracy ratios pertain to the percentage of correctly identified defaults *within-sample*. Out-of-sample, the latter accuracies are - typically - lower
[2] Abbreviated MDA hereafter
[3] To the best of our knowledge, Duffie et al. (2007) are one of the select few that have documented listed-company out-of-sample accuracies in excess of 65% for longer (five-year) forecasts
[4] This not to say classification accuracies in excess of 90% are necessarily commonplace, however. Rather, comparative studies under the same pre-textual conditions show alternative model specifications are consistently able to dominate that of Altman (1968)
[5] See e.g. Shumway, 2001; Chava and Jarrow, 2004; Nam et al., 2008; Bauer and Agarwal, 2014

Setting aside prediction-length, what of the private sphere? If advances in public equity predictions have been piecewise, progress made in the private market have been lethargic. By almost any measure, the world is emphatically private[6], not public - and yet the mainstream corpus of default literature is unmistakably of the public, listed kind. The private default paucity of studies is presumably partly due to a simple lack of financial data that has become the point of departure for practically all financial distress predictions. Indeed, publicly available information on privates is often limited (Duan et al., 2018) and reporting requirements - if any - are typically more forgiving and lenient than those for listed companies. Although rather harmonised, much of the private U.S. market, for instance, is dictated by regional state laws, presenting some measure of hardship for the inter-state private bankruptcy researcher. Tax returns must be registered with the IRS federally, but these are under no legal obligation to be made public. Under the (amended) Securities Exchange Act of 1934, the SEC necessitates private company financial filings first if the private entity in question has assets in excess of $10 million and more than 500 unaccredited common-equivalent shareholders[7]; although this is not without exception (Securities Exchange Commission, 2018). By any stretch of the imagination, such a company would constitute a relatively large private, and only then is it obligated to catalogue and publicly disclose its yearly and quarterly financial performance.

Returning to the turf of accuracy ratios, private markets have generally seen a haircut to that of public predictions in the (rough) whereabouts of 15%[8] across forecast horizons. It is, simply put, more challenging to effectively excavate tell-tale signs of forthcoming distress in a setting characterized by infrequent, less stringent data: where listed companies provide (public) quarterly reports and often daily market information, private corporate default investigators must (in terms of financial ratios) make do with less comprehensive and periodic accounting data. And then, she must first have admission to said data. The virtues of public-company data become especially noticeable when considering the efforts made by academics to incorporate listed market information in foretelling private firm duress. Andrikopoulos and Khorasgani (2018), for instance, draw on listed SME-peers[9] in anticipating the defaults of non-listed SMEs. SME default research, moreover, is often operationalised by Basel- or EU-set criteria from which a considerable portion of the total private market dissipates[10].

In any case, though much fewer in number, there are studies conducted on private companies that achieve remarkable predictive accuracies rivalling that of public corporate defaults. Notably, Altman et al. (2016), are able to reach out-of-sample single-year (ten-year) accuracy ratios of ca 90% (46%) respectively by extrapolating a cross-sectional analysis of both financial and non-financial covariates from 2003 on Finnish non-listed firms between 2004-2013. These impressive results come with three caveats, however: (i) a cross-sectional (i.e. single-period) feed-in period for ten-year-ahead predictions is arbitrarily and econometrically problematic (ii) model covariates are un-parsimoniously large in number[11] (iii) model overfitting to some degree is a

[6] e.g. in OECD countries the percentage of SMEs vis-à-vis the total number of firms exceeds 97% (Altman and Sabato, 2007)
[7] Or more than 2,000 common-equivalent shareholders (accredited or not). These conditions are not exhaustive, nor are they without exemption (Securities Exchange Commission, 2018)
[8] See e.g. Altman and Sabato, 2007; Pederzoli and Torricelli, 2010; Duan et al., 2018
[9] Denotes “Small- and Medium-sized Enterprises”
[10] The European Commission, for instance, defines SMEs as businesses having less than 250 persons employed. Additionally, their annual turnover may not exceed EUR 50 million, unless their balance sheet falls below the EUR 43 million mark (Commission Recommendation 96/280/EC, updated in 2003/361/EC, enforced as of January 1, 2005)
[11] Specifically, all one-year to ten-year forecasts encompass a total of thirty-five (35) predictors

likely consequence of the latter[12]. Shelving concerns of the first and second kind, private models with a parsimonious count of predictors are curiously rare.

Duan et al. (2018) and Dakovic et al. (2010) are two of the distinguished few that maintain a seemingly parsimonious tally of covariates - six (6)[13] and eight (8) in number respectively. The first, however, nests a further six predictors into one if its higher-level regressors - the *distance-to-default*[14] public firm equivalent - and achieves relatively modest albeit stable accuracy ratios over a three-year forecast horizon in the range of 53-56%. Moreover, the authors superimpose forward-looking (market) signals of duress from publics onto privates, assuming some degree of cross-sectional homogeneity between non-listed companies and their listed "peers". In a setting demarcated by high heterogeneity among firms, such a public-private read-across becomes particularly precarious. Post-GAM[15] and industry dummy extensions, Dakovic et al.'s (2010) research into Norwegian private limited companies also culminates in a less frugal thirty (30) combined predictors in their final model blueprint. Although the authors attain excellent accuracy ratios exceeding 78% over a four-year forecast horizon, their entire data survey also extends for a scant five years of time - a considerably shorter window than most. Consequently, as in Altman et al.'s (2016) research in Finland, for longer predictions ahead in time their feed-in sample (i.e. training data) becomes increasingly limited.

The paucity of parsimonious default models in the private market is so pronounced, we are unable to recount a single paper in the prevailing body of literature that with ten or less regressors has produced satisfactory accuracy ratios without appending its predictions with public market "equivalent" information[16]. We believe this is testimony to the hardship with which private company investigators have forecasted signals of distress and credit risk more generally. When competing on a comparable criterion (accuracy) in the shadow of the more data-privileged publics, it is no wonder the private strain of default research has, if anything, become less parsimonious and, indeed, less private with time. We attempt to address this gravitation towards longer model specifications and public market extrapolation by comprehensively reviewing in excess of 200 financial and non-financial covariates as a means for prudently predicting Swedish non-listed company defaults up to five years in advance between 1998-2017[17], limiting selected covariates to ten in number. In doing so, this paper - to the best of our knowledge - also constitutes (i) the most exhaustive investigation of the extant literature's catalogue of regressors[18] (ii) a large

[12] Model overfitting may be prevalent even in the event of excellent out-of-sample accuracy ratios. *A priori* unexpected increases in hit-rates with lead time, as Altman et al. (2016) detail in prediction years six-to-seven and nine-to-ten, may suggest a degree of noise-capturing and possible variable misspecification

[13] Not counting an additional two macro-level variables, however

[14] Briefly, first developed by Merton (1974), this credit risk metric gauges public companies' default risk by considering equity as a call option on the firm residual value

[15] Abbreviates "Generalized Additive Model": univariately investigates Kernel density functions to spline smooth regressors into distributionally-fitted predictors (for more on GAM, see Hastie and Tibshirani, 1990)

[16] With the exception of Pederzoli and Toricelli (2010) who achieve an accuracy ratio of ca 67% in sample using just four covariates. The authors, however, (i) perform no model validation tests outside of their sample (ii) consider solely a one-year forecast horizon (iii) suffer the same caveats as Altman et al. (2016) in using a single-period cross-sectional classification set-up

[17] Swedish national law unconditionally mandates all limited liability companies report financial statements for the fiscal year passed (Bolagsverket, 2018). Prior to 2014, the universe of Swedish private companies were also required to file audited financial accounts without exception (Bolagsverket, 2013)

[18] Commenting bankruptcy prediction studies from the infancy of the 1930's, Bellovary et al. (2007) identify 752 factors used across 165 studies over time. Of these, a staggering 674 are used in just two studies or less

private firm data-set in terms of firm-year observations, unique firms and number of defaults, second only to Altman et al.'s (2012) UK study on SMEs.

We compare our one-to-five-year models' predictive power and calibration with five (5) popular alternative specifications from the literature. Collectively, these come from different time-periods and geographies, are parsimonious and un-parsimonious and have been used in foretelling public-company, private-company and SME defaults. In operationalising the economic merit of our model, we both simulate and backtest all six (6) models' one-period predictions against Swedish private loan data contained in the sample itself between 2014-2016. The lender's perspective is assumed under the Basel III F-IRB approach.

2. Literature Review

2.1 Introduction

Although the extant literature on bankruptcy prediction now spans the better part of more than eight decades, we focus our efforts towards delineating in rough strokes the strains of default research that most closely neighbour this study. For a more thorough grounding and rendition of the previous bankruptcy literature most attuned to our enterprise, the reader is referred to Bellovary et al.'s (2007) literary review, as well as Gupta et al. (2017, pp. 437-447). In any case, it is beyond the scope of this paper to exhaustively detail the many parallel streams of default detection ranging across SMEs, private companies, listed corporations and financial instruments.

Thus far we have used the "default" terminology somewhat haphazardly. We use this taxonomy interchangeably with "bankruptcy" herein and, to be concrete, both these encapsulate non-voluntary default, bankruptcy, liquidation and organisational restructuring as delimited by Swedish Private Company Law. Much of modern research interprets legal bankruptcy in this broader sense, whether investigating public or private corporations (see e.g. Chava and Jarrow, 2004; Duffie et al., 2007). Some have even appended the legal definition of bankruptcy to include early evidence of financial distress by some *a priori* set cut-off rule[19], to good effect.

2.1.1 A Brief Historical Recount of the Evolution of Default Detection

Bankruptcy prediction's most simple design likely takes the form of univariate factor analysis. Most distress research employed this now archaic mode of forecasting up until the 1960s. Beaver, (1966) is likely the most iconic univariate study of this era. With his simple approach, he detects 92% of bankrupt and non-bankrupt firms one year in advance using only a single predictor - net income to total debt. His legwork lays the foundation for what only a couple years later becomes Altman's (1968) seminal Z-score. The latter makes use of multiple predictors as opposed to a single one and is the cornerstone in the famously cited *Multiple Discriminant Analysis* (MDA) prediction method. Albeit principally constructed with the manufacturing industry in mind (understandably, as of the time) Altman's five Z-score variables continue to have widespread use across industries and geographies in both its original MDA format and, more so, in *non-linear regression* (e.g. Duan et al., 2018; Reisz and Perlich, 2007; Agarwal and Taffler, 2008a). The latter methodology saw its inception through Ohlson (1980), where the multivariate conditional

[19] Gupta et al. (2017, p.438) and Keasey et al. (2014) classify a company as under financial duress if "it reports earnings of less than its financial expenses for two consecutive years, has net worth/total debt less than one, and experiences negative growth in net worth for the same two consecutive time periods"

probability model (or simply, the *logit* binary response model) made its entrance into the default prediction space. This modelling technique came as a direct critique to MDA, its many idiosyncrasies and predictive problems.

MDA, Ohlson (1980) showed, relies on equal covariance and normality. As a consequence of the former, MDA allows exclusively for matched-pair research where the count of defaults and non-defaults must equate. In the context of rare-event studies, such as defaults, this is by today's standards an exceptionally limiting data constraint. MDA's linear parameterization and its latter assumption of normality may also produce probabilities of default both less than 0 and above 1. Ohlson was the first to pioneer a solution to both of these limitations, why conditional probability models of different sorts have become the hallmark for the lion's share of modern default studies (Bellovary et al., 2007). Some odd years later, Zmijewski (1984), introduced his take on the multivariate conditional probability model in bankruptcy forecasting - the *probit* - this time reverting to the use of the standard normal distribution as opposed to the logistic link function. Nevertheless, up to this point the reader may be surprised to learn that all modes of bankruptcy prediction have focused entirely on financial data in the period immediately preceding the time of default. In other words, the literature has employed a single-period classification model. Shumway (2001) calls this breed of predictive functions *static* insofar as they do not incorporate time-varying covariates in foretelling bankruptcy. He suggests making use of multiple-period financial data instead, through what he describes as a duration model with time-varying regressors, or simply the *discrete duration hazard model*[20]. Single-period classification is both biased and inconsistent, he argues, while anecdotally demonstrating that in a ten-year data-sample the hazard model incorporates approximately ten times the data as that of the former. Forecasts, resultantly, are both more precise and yield superior predictions while unbiased (*ibid.*). Survival analysis, therefore, ought to be the researcher's preferred *modus operandi* on both theoretical and empirical grounds.

Thereafter, several studies have supplemented Shumway's discrete duration hazard model with some economically and statistically meaningful additions, largely borrowed from the domain of medicine from which survival analysis originates. Nam et al. (2008), for instance, extended Shumway's (2001) discrete duration-*independent* hazard model to that of one with duration-dependence where the baseline hazard rate - a theoretical must in the discrete-time space - is temporally dependent, i.e. varies with the passage of time. Gupta et al. (2017) and Dakovic et al. (2010) include random effects on the observational unit level to account for heterogeneity among firms - the equivalent to what is often called *shared frailty* in the jurisdiction of survival analysis.

More recently yet, heuristic algorithms including artificial neural networks, random forests and support-vector machine have seen a surge in use (see e.g. Ciampi and Gordini, 2013; Ribeiro et al. 2012). So far, however, there is little to suggest that any machine learning construct is consistently able to outperform hazard models and non-linear regression more generally speaking (Bellovary et al., 2007). Ciampi and Gordini (2013), for instance, detect just 1.8% more defaults on the aggregate level relative logit regression in their holdout sample (i.e. control group) of Italian SMEs.

[20]Often referred to under a wider umbrella of nomenclature, including but not limited to *survival analysis*, *event history analysis* and *hazard models*

2.2 Selected Competing Default Models

2.2.1 Altman's MDA Z-score - 1968

In response to Beaver's (1966) univariate prediction regime - and call to future multivariate works - Altman (1968) developed a multivariate discriminant analysis comprised principally of univariate ratio analysis about input centroids (group means). His review of financial variables for the American manufacturing industry produces a five-factor multivariate linear parametrization of default probabilities in a matched-pair setting of thirty-three defaults and non-defaults. Combined, these covariates deliver an arbitrary Z-score which can then be measured against a set cut-off threshold for classifying defaulters and non-defaulters.

Altman's Z-score MDA predictors are: (i) working capital to total assets (WC/TA), (ii) retained earnings to total assets (RE/TA), (iii) EBIT[21] to total assets (EBIT/TA), (iv) market value of equity to book value of total liabilities (MEQ/TL) and (v) sales to total assets (S/TA). According to Altman (1968, pp.594-596) the factors capture (i) liquidity and size, (ii) cumulative profitability over time, (iii) firm productivity, (iv) leverage and (v) capital turnover as management's ability to meet competition, respectively. Albeit a classification model originally devised for public companies, the literature is ripe with examples where the Z-score predictors are benchmarked against private companies (see e.g. Chava and Jarrow, 2004; Dakovic et al., 2010) [22] - often through the now commonly used logistic regression.

2.2.2 Zmijewski's Conditional Probability Binary Response Model - 1984

Following the well-versed footsteps of Ohlson (1980), Zmijewski (1984) highlights the inherent oversampling bias of the matched-pair study vis-à-vis the rarity of the bankruptcy event in the real world. Samples are selected in a non-random fashion, and defaults overrepresented relative the population rate and census proportion.

Zmijewski's (1984) proposed regressors under the probit non-linear regression are: (i) net income to total assets (NI/TA), (ii) total liabilities to total assets (TL/TA) and (iii) current assets to current liabilities (CA/CL). He contends his predictors embody (i) profitability, (ii) leverage and (iii) liquidity, respectively. Akin to Altman (1968), Zmijewski (1984) developed his model with the public market in mind. Although his covariate selection is used less regularly in the realm of private companies, it enjoys considerable attention in the public domain (e.g. Shumway, 2001; Chava and Jarrow, 2004).

2.2.3 Shumway's Discrete Duration Hazard Model - 2001

Recall Shumway's (2001) critique; prior bankruptcy prediction methods have ultimately only come in one shape or form - as single-period classification models. These "static" models consider but a handful of the multiple-period data available to the distress researcher. The latter, he maintains, delivers consistent estimates when used in its hazard model format, explicitly accounting for time. Econometrically speaking, survival analysis is sensitive to the period at risk and allows for time-varying covariates as well as temporal duration dependencies, while

[21] Denotes "Earnings Before Interest and Taxes"
[22] In doing so, the market value of equity is typically substituted for the book value of equity

generating more efficient out-of-sample predictions by harvesting multiple as opposed to a single data collection period.

Shumway (2001) finds a combination of accounting and market variables as the most promising predictive information set. Baseline hazard rate aside, his five predictors are: (i) net income to total assets (NI/TA), (ii) total liabilities to total assets (TL/TA), (iii) the logarithm of the size of the firm's market value of equity relative its stock exchange's total market capitalization (LNRELSIZE), (iv) the idiosyncratic standard deviation of the firm's stock returns (SIGMA) and (v) the firm's leading period's annualized cumulative monthly return less the stock market index return on which it trades (ABSRETURN). Like his peers Altman (1968) and Zmijewski (1984), Shumway (2001) advanced his predictive efforts on the premise of public market information and listed companies. He asserts his elected covariates represent (i) profitability, (ii) leverage, (iii) forward-looking prospects and firm size, (iv) cash flow variability and/or operating leverage and (v) performance relative to peers and the general marketplace. As a rule of thumb, Shumway's (2001) covariate specification is used more regularly than Zmijewski's (1984) yet less so than Altman's (1968) Z-score regressors in the private default literature. Chava and Jarrow (2004), for instance, draw on Shumway's chosen predictors excluding market variables (iii)-(v)[23] in an attempt to investigate different models' forecasting ability using monthly bankruptcy data and extending model predictions to include the financial sector.

2.2.4 Altman and Sabato's U.S. SME Default Model - 2007

Noticing an apparent lack of U.S. SME default prediction dating back to Edmister (1972), Altman and Sabato (2007) investigate the capacity of generic corporate bankruptcy models in foretelling bankruptcies of these smaller organisational entities. In line with previous literature's findings, they discover that a SME-tailored prediction model outperforms a large corporate corresponding model to the tune of almost 30% in accuracy. Using a simple market simulation, the lender armed with the SME-adapted prediction model reaps an approximate 50 basis point capital alleviation in Basel II A-IRB enforced capital requirements (*ibid.*).

The authors detect five pertinent SME default predictors: (i) short term debt to equity book value (STD/EQ), (ii) cash to total assets (CASH/TA), (iii) EBITDA[24] to total assets (EBITDA/TA), (iv) retained earnings to total assets (RE/TA) and (v) EBITDA to interest expenses (EBITDA/IE). These, the researchers purport, express the firm's financial profile in terms of (i) leverage, (ii) liquidity, (iii) profitability, (iv) coverage and (v) activity, respectively.

2.2.5 Dakovic, Czado and Berg's non-linear predictors with random effects - 2010

In examining the functional relationships between explanatory variables and the probability of default, Dakovic et al. (2010) are able to augment their forecasting ability. The authors accomplish this through examining the Kernel density estimators of their covariates under study through the GAM framework, highlighting important non-linear relationships across different parts of the covariate-probability-of-bankruptcy distribution. An industry-specific random intercept is included as a means of differentiating between industry heterogeneity and, implicitly, a shared industry frailty term (*ibid.*).

[23] The result of which - NI/TA and TL/TA - they term their *private firm model*, relying only on accounting information
[24] Denotes "Earnings Before Interest, Taxes, Depreciation and Amortization"

In the Norwegian private market[25] between 1996-2000, the researchers identify eight explanatory covariates: (i) the number of auditor remarks (REVANM), (ii) firm age measured in years (AGE), (iii) a dividends paid dummy indicator (DIV), (iv) book value of equity to total assets (BEQ/TA), (v) logarithmised total assets (LNSIZE), (vi) cash and marketable securities to current liabilities (CASH/CL), (vii) return on assets to total assets (EBIE/TA) and (viii) current liabilities to total assets (CL/TA). Briefly, these are postulated to capture (i) accounting quality, (ii) past experience, (iii) management activities, (iv) solidity, (v) size, (vi) liquidity, (vii) profitability and (viii) leverage, respectively (*ibid.*).

Including industry dummy variables as well as the spleen-smoothed expansion of the functional relationships observed through GAM produces a total of thirty (30) regressors. More, Bellovary et al. (2007, p.12) remind us, may not equate to better predictions: Beaver (1966) enjoys a 92% accuracy ratio using one variable, Jo et al. (1997) settle for 86% using 57 predictors.

2.3 Non-financial predictors - Accounting Quality, Corporate Governance, Innovation and Frailty

As Bellovary et al. (2007) indicate, the archive of covariates studied under default prediction is expansive - 752 unique indicators across some 165 studies between, roughly, 1930-2007. Once dominated by the presence of exclusively financial predictors, the universe of default prediction has come to value qualitative information as indispensable to bankruptcy prediction the more readily available it has become. This is particularly true for the less transparent private market (e.g. Altman et al., 2012; Lehmann, 2003). Although the extent of variation across non-financial measures is extensive, popular categories of the qualitative sort have mirrored facets of accounting quality and corporate governance. These include, among others, auditor remarks, changes of auditors (Altman et al., 2016; Senteney et al., 2006; Dakovic et al., 2010), reporting compliance and on-time filings (Altman et al., 2012; Altman et al., 2016), board size / composition, owner concentration, multiple directorships, affiliated directors and payment slips (Daily and Dalton, 1994, Ciampi, 2015; Altman et al., 2016), credit analyst judgment and management quality (Lehmann, 2003; Edmister, 1972).

Notably, the literature on innovation and survivability is more sparsely populated. Buddelmeyer et al. (2008) persuasively suggest that innovation is an inherently risky endeavour operating closer to the tails of the survival distribution, at least relative most other traditional predictors. Innovative processes may milden the *ex ante* likelihood of default when executed well, but can equally amplify the firm's distress because of its ingrained riskiness. While Cefis and Marsili (2005) retrieve a positive correlation between survivability and successful innovation, Buddelmeyer et al. (2009) nuance this discovery by distinguishing between successful and unsuccessful innovative practices novel to the firm and novel to the economy. Audretsch (1995), on the other hand, convincingly highlights that new entrants to innovative industries experience a barrier to survival as a consequence of said industry's innovative nature. Concurrently, however, this higher hazard rate diminishes as the post-entry time (i.e. period at risk) elongates and the entrant gains experiences (Audretsch, 1995, p.454). In analytically quantifying innovative efforts, Zoltan and Audretsch (1989) empirically show that the registered count of patents do well in capturing innovative activity, co-varying with company skilled labour and R&D spent

[25] Admittedly, the authors point out that "most" of the 98,421 firms under scrutiny are not registered at any exchange (Dakovic et al., p.1739)

resources. As Buddelmeyer et al. (2009, p.262) point out, however, patents - particularly when recently submitted or approved - contain both legal and market uncertainty, and the count of patents is insensitive to the current stage of the patent lifecycle. Granted patents, nonetheless, ought to be considerably less risky than those awaiting approval; although what separates a "good" from a "bad" patent continues to be elusive and hard to decipher externally from outside the company.

Frailty[26] - i.e. correlated defaults - has received more recent attention in light of today's higher incidence of systemic risks and liquidity-commonality characteristic of recessions and flash crashes. Lang (1992) proclaims that bankruptcy announcements exert a simultaneous competitive and contagion effect on industry peers. The first encapsulates a positive gain of wealth from a changed intra-industry competitive hierarchy while the latter constitutes a deadweight loss to the survivor from shared cash flow characteristics with the defaulter. In industries demarcated by low levels of leverage and competition, the competitive effect dominates the contagion effect (Lang, 1992, pp.59-60). Controlling for firm-observable factors, Duffie et al. (2009) demonstrate that modern conventional loss estimates systematically underestimate the default risk innate to collateralized loan portfolios. Some unobservable *latent factors* (shared distress exposure) is at work among U.S. corporations, they argue, why failing to allow for heterogeneity among firms downward biases credit risk measures including the popular value-at-risk metric (*ibid.*). Das et al. (2007) echo this contention, presenting evidence of default clustering among U.S. corporates between 1979-2004. In response to the latter study, however, Lando and Nielson (2010) show that the paper's underlying methodology is unable to account for contagion-events present in the explanatory variables that themselves deterministically measure firm default - i.e. contagion in the covariates goes unnoticed. Adjusting for this *ex ante*, the authors find no evidence of *latent frailty* operating outside of the regressors themselves.

On the aggregate, the literature on innovation's and frailty's relationship with survivability remains somewhat fleeting and contended.

2.4 Survival Analysis and Discrete-Time Hazard Models

For our purposes, three commonly encountered phenomena in survival analysis are especially important to flesh out. These are left-truncation, censoring and delayed entry. *Left-truncation* is the denomination for when a unit enters the data mid-study, i.e. not at the start of the observational period. In the domain of bankruptcy research, such is rather commonplace – e.g. dormant firms become active, or new companies are formed. *Censoring* is the consequence of having access only to a partial window of the lifetime of the subject; observations prior to and after the start and end of the observational period's time-frame are, ergo, *censored*. *Right-censoring* is particularly prevalent in the context of defaults, where the time-to-event (default) is unknown as it has not occurred within the time-interval limiting the length of the study. *Delayed entry* manifests when, as often is the case, the earliest available information in the dataset on a company does not coincide with its inception and incorporation. The company, to use survival analysis jargon, has been *at risk* for periods preceding the start of the data available. Although the incidence of these lies outside the researcher's control and is to such an extent an unavoidable

[26] Often synonymously called *contagion*. Often econometrically captured through random effects

nuisance[27], recognizing these data characteristics while ameliorating their impact to results and inference thereof is of growing importance (Allison, 2010).

In its most general form, the survival function, denoted $S(t)$ below, embodies the probability a firm will survive beyond a time t. The hazard rate, $h(t)$, a direct derivation of the survival rate, measures the instantaneous default probability at t (see Equations (1) and (2) below).

$$S(t) = \Pr(T \geq t) = 1 - F(t) = \int_t^{\infty} f(u)du \tag{1}$$

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{2}$$

T is the survival time, a continuous random variable following the cumulative (probability) density function $F(t)$ and $f(t)$ respectively.

Survival time, the time-to-event, is almost exclusively retrieved at discrete points in time in the default literature – typically in annual or monthly intervals. When the occurrence of an event is known precisely, Gupta et al. (2017, p.439) maintain, a *continuous-time* hazard model may be reconcilable with the choice of survival model. Precluded on these grounds alone in the vast majority of cases, most practitioners and researchers opt for the *discrete-time* hazard model instead. Doing so, however, theoretically mandates the inclusion of a *baseline hazard rate* temporally chronicling the hazard function's development over time. Therein lies a long backlog of different baseline specifications, ranging from time dummies (Beck et al., 1998), foreign exchange rate volatility (Nam et al., 2008) to the natural logarithm of age (Shumway, 2001) and the preceding period's realized rate of default (Hillegeist et al., 2004). These, the reader should notice, can be duration *dependent* (e.g. the realized rate of default) or duration *independent* (e.g. company age). Finally, in the discrete (continuous) time setting using the logit or probit link functions, the binary dependent - failure and non-failure - is equivalent to the *odds ratio* (*hazard rate*).

Different survival analysis models also rely on different assumptions regarding the relationship between the hazard function and the vector of covariates / predictors. Two prevailing bodies of these are accelerated failure time (AFT) models and proportional hazards (PH), both of which come in semi-parametric and fully-parametric formats as well as continuous and discrete time (Nam et al., 2008; Gepp and Kumar, 2008, pp.15-17). The (Cox) PH adaptation being eminently more regularly used, AFT models are best left as an unexplained adjacent to PH models for brevity.

[27] Without the loss of, at times, significant portions of data

As such, Cox's extended form[28] semi-parametric PH model in continuous time simplifies from Equation 2 to:

$$h(t|x_{i,t}) = \alpha_0(t) \cdot e^{x_{i,t}' \cdot \beta + c} \tag{3}$$

Where $\alpha_0(t)$ is the baseline hazard rate and $e^{x_{i,t}' \beta + c}$ parametrically captures some *a priori* specification of predictors x_i for firm $i = 1, 2, 3, \ldots, N$ at time $t = 1, 2, 3, \ldots, t$ with constant c. β coefficient(s) are estimated analogously to the maximum likelihood procedure in the asymptote following Shumway's (2001) derivations and rationale.

We streamline the remainder of the discussion further to the *discrete time* application of equation 3, given its wildly more prominent use over Cox *continuous time* models in the literature (Gupta et al., 2017; Gepp and Kumar, 2008).

Although the discrete analogue of Equation 3 is strictly speaking comparable to the binary response model using the complementary log-log link function (*cloglog*), Beck et al. (1998) and Shumway (2001) show "that the cloglog link and the more familiar logit link are almost identical when probabilities of failure are sufficiently small" (Nam et al., 2008, p.496). Using the logit function, Equation 3 condenses to:

$$\Pr(y_{i,t} = 1) = \frac{1}{1 + e^{-(\alpha_t + x_{i,t} \cdot \beta + c)}} \tag{4}$$

Where $y_{i,t} = 1$ is the instantaneous probability of firm i defaulting at time t. When $\alpha_0(t)$ is duration dependent, this Cox concoction can be neatly summed up as a *discrete-time duration-dependent hazard model* (*DTDDH Model*) running under *the Cox semi-parametric proportional hazards* form[29].

As Gupta et al. (2017, p.438) articulate, the lion's share of existing bankruptcy studies using survival analysis are overly lax in (i) explicating how delayed entry is handled (ii) omitting or failing to specify the underlying baseline hazard rate (iii) investigating and incorporating contagion / frailty as well as recurrent events (i.e. multiple defaults) (iv) testing the proportional hazard assumption in the parametric and semi-parametric Cox model variations (v) defending the choice between discrete-time or continuous-time hazard models. *Latent* frailty, contagion and recurrent events can be relatively easily integrated into the logit regression through random effects (i.e. random intercepts specified on an observational unit sort), and issues pertaining to delayed entry can be circumvented by simply discarding time periods' likelihood contribution prior to the onset of the study (Gupta et al., 2017, p.441). Meanwhile, proportional hazard may be tested for parametrically or non-parametrically. Non-parametric tests of equality between survival functions (i.e. non-parametric PH tests) typically assume the Wilcoxon-Breslow-Gehan test and log-rank test. When predictors are continuous and time-varying, however, it is both inappropriate not to mention incredibly taxing on time to compare survival functions for every assumed value of the predictors. Nam et al. (2008), nevertheless, do just this when univariately

28 The extended form allows for time-varying covariates, while the reduced form does not

29 Henceforth, Cox semi-parametric proportional hazards DTDDH is used interchangeably shortened to *DTDDH* as well as *logit model, binary response model* and *hazard model*

studying possible model covariates. Instead, as Gupta et al. (2017, p.446) report, (scaled) *Schoenfeld residuals* from the survival analysis model fitting may be used in a chi-squared test to detect violations of the proportional hazards assumption. Scaled Schoenfeld residuals are centred at $\hat{\beta}$ for each regressor such that the residuals' slope, when plotted against time[30], may not be significantly different from zero in satisfying the null hypothesis that PH is adhered to. This slope test is analogous to verifying that the log hazard-ratio function is constant over time. It should be reiterated, however, that the extant literature it littered with examples where the proportional hazards assumption is both unchecked and violated. As Allison (2010) counsels, one must be concerned with both the inclusion of economically meaningful covariates and infractions of the PH assumption.

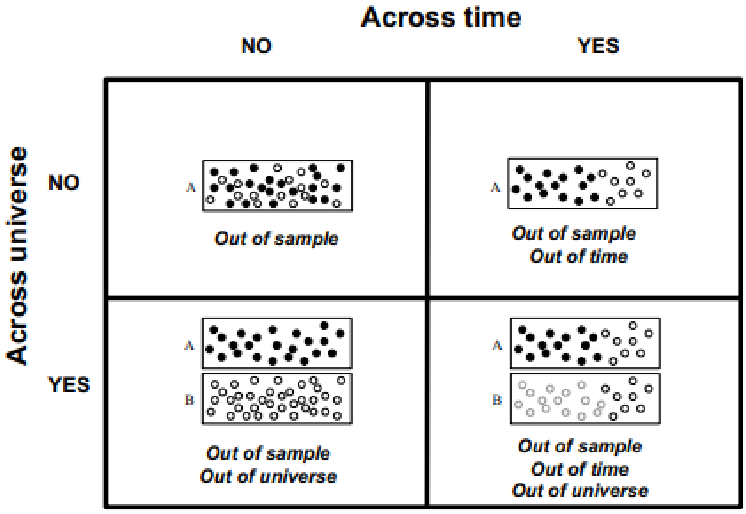
Since Shumway (2001), corporate bankruptcy research has welcomed the field of medicine's survival analysis technique for predicting corporate non-survivors. Shumway (2001) derived the eloquent result that hazard rate estimation – in the discrete time setting –converges to the simple binary response, conditional probability logit model with multiple periods. The discrete hazard rate, he explains, bears the same asymptotic variance-covariance matrix and likelihood function (Shumway, 2001, pp.109-111) as the multi-period conditional probability logit. Recall that the primary advantage of survival analysis relative traditional logit analysis techniques employed by Ohlson (1980), Zmijewski (1984) and others is that it allows the researcher to capture data's time-series features through time-varying covariates. The former constructs, in other words, are unable to recall the information set's multiple periods, why they *ex post* have come to be known as single-period (or cross-sectional) classification schemes. This is one of several points of contention Shumway (2001) would caution in light of Altman et al.'s (2016) and Pederzoli and Toricelli's (2010) prediction studies in Finland and Italy respectively.

2.5 Model Validation - Receiver Operating Characteristics and Precision Recall Curves

The Lachenbruch (i.e. *jackknife*) method of model validation repeatedly retains an observation from the estimation sample to be classified post-estimation by the discriminant model. Albeit a satisfactory method when burdened with a small sample size, a superior measure of external validity makes use of a *hold-out* sample containing a set of observations not yet seen or estimated across some dimension. Model validation through these efforts are a means of examining the model specification's external validity while concurrently probing for model overfitting in the estimation window. These can be further categorised into hold-out samples within and outside of the sample's time-period. For convenience, Table 1 illustrates a schematic of these variants visually.

[30] Or a functional form thereof, e.g. the natural logarithm of time

Table 1: Schematic View of Hold-Out Sample Types



Stein (2007, p.92)

Hold-out testing may, as can be deduced from the above, account for external variance across the observational unit and across time. In the table above, dark circles indicate training data (i.e. the estimation sample), white circles are testing data and grey circles highlight (optional) data points that can be used for additional testing. For robustness, the simplest of hold-out cross-validation checks (upper-left quadrant) is typically insufficient in capturing, at the very least, the temporal variance of default predictions. Especially so in a setting characterised by bankruptcy clustering, cyclicality, recurrent events and / or frailty – i.e. under any circumstance where defaults cannot be assumed to hold the properties of a stationary process (R. Stein, 2007). Arguably the most common validation technique of the hold-out kind is what Stein (2007) labels the *out-of-sample out-of-time* test, where data is tested after some fixed point in time using some sub-sample (typically the same observational units as those in the estimation window, in addition to a smaller set of previously unobserved units). The most demanding test (*B* in the lower-right quadrant) considers exclusively new observations across both time and unit, i.e. unobserved companies across an unobserved interval of time. This version, unsurprisingly, ought to be the researcher's preferred choice of cross-validation technique using hold-out samples. As Stein (2007) reminds us of, however, this most stringent test of external validity is only feasible in large datasets given the rarity of the default event – else, too many defaults may be left outside of the estimation window. In any case, although Stein (2007) terms this most rigorous test format as *out of universe out of time*, we prefer the taxonomy of *hold-out out-of-time*. Herein, therefore, *hold-out* always refers to an observational unit (company) never before seen – analogous to *out of universe*.

In spite of the many merits of cross-validating models, the reader may be surprised to learn that, as per Bellovary et al. (2007), less than 50% of surveyed bankruptcy research up to 2007 made use of hold-out testing for validating model performance outside of the sample used to fit said model[31] (i.e. external validity). Hold-out validation tests have taken two main routes in the bankruptcy literature: that of Cumulative Accuracy Profiles (*CAP*) and that of Receiver Operating

[31] Specifically, 77 of 165 studies made use of some means of hold-out testing; 5 of 11 during the 2000's (Bellovary et al., 2007, p.8)

Characteristic (*ROC*) curves. The latter camp condenses *contingency tables'*[32] *True Positives* (*TPs*, i.e. correctly identified defaults), *True Negatives* (*TNs*, i.e. correctly identified non-defaults), *False Positives* (*FPs*, i.e. misclassified defaults) and False Negatives (*FNs*, i.e. misclassified non-defaults) for all possible *cut-offs*[33]. In a contingency table (Table 2 below), a predetermined cut-off point (e.g. $\Pr(y_{i,t} = 1) = 0.5$) determines the number of TPs, TNs, FPs and FNs.

Table 2: Hypothetical Contingency Table for a Cut-Off Threshold

		Actual		
		Default	**Non-default**	Total
Predicted	**Default**	*TPs*	*FPs*	Σ = TPs + FPs
	Non-default	*FNs*	*TNs*	Σ = TNs + FNs
	Total	Σ = Actual Positives	Σ = Actual Negatives	Σ = N

In order to create a common yardstick from which to compare model power and accuracy irrespective of the selected cut-off and tacit costs of misclassification, the ROC conveniently plots the *False Positive Rate* (*FPR*) against the *True Positive Rate* (*TPR*) for all possible cut-off points (see Equations 5-6). The latter ratio signifies the *hit rate* and is often called the *sensitivity* or *recall* of the classifier.

$$FPR = \frac{FPs}{FPs + TNs} \tag{5}$$

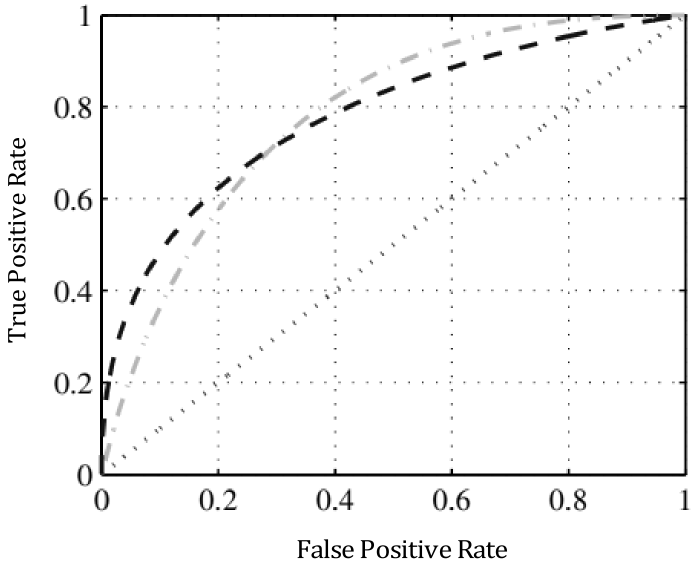
$$TPR = \frac{TPs}{TPs + FNs} \tag{6}$$

Figure 1 illustrates the main features of the ROC curve. The ROC curve is a two-dimensional plot of a model's ability to discriminate categorically between positive cases (bankruptcies) and negatives (survivors).

[32] Synonymously referred to as *confusion matrices* at times

[33] Cut-offs are the default probability thresholds (values) from which defaulters are classified and separated from non-defaulters

Figure 1: ROC Curves for Two Competing Models



It should come as no surprise that credit risk predictions may forecast low risk erroneously and vice versa. The first kind of inaccuracy may lead to a false negative classification (*Type II error*), and the second to a false positive (*Type I error*). The ROC curve can be thought of as the coordinates of the Type I and II errors, computed for every cut-off point contained in the estimation sample, and subsequently plotted in two-dimensional *TPR* and *FPR* space (Blochlinger and Leippold, 2006). A convenient way for summarizing the ROC curve comes in the form of its integral, i.e. the *area under the ROC curve* (*AUROC*, for short), which makes use of all quadrants in Table 1. A predictive model's AUROC has a highly intuitive reading; it is, to quote Stein (2007, p.83) "the probability that a randomly chosen default will be ranked worse than a randomly chose non-default". An AUROC of 0.5 – graphically represented as a dotted 45° line in Figure 1 – reproduces the classification power of a random model. 1.0, by extension, signifies a model with perfect power, ergo perfect discrimination ability between defaults and non-defaults (Stein, 2007). A model is said to dominate that of a rival in all situations (i.e. for all cut-off points) if it produces a significantly higher AUROC and its ROC curve never crosses that of the competing model. In Figure 1 although the AUROCs under each curve are equivalent, the dark line appears to be preferable over its grey rival for some cut-off values and vice versa. ROC curves and their self-contained AUROCs make them an appealing comparative measure of model power vis-à-vis, for instance, the cut-off dependent contingency tables (Flach and Kull, 2015). In fact, the observant reader may have picked up that the ROC curve is nothing other than a generalized graphical expression of every possible contingency table for every cut-off level[34].

The aforementioned CAP 'faction', as it were, sees the use of *accuracy ratios* (*AR*) as the equivalent to the ROC's AUROC summary term. Nevertheless, the CAP and the ROC convey the exact same information across slightly different mediums. Every point on the ROC and CAP can be transposed

[34] To see this, note that the ROC curve is plotted on TPR and FPR axes. Each ROC-coordinate, therefore, captures the entire information set contained in the contingency table for some cut-off point (see Table 2, Equations 5-6 and Figure 1)

to one another – so much so, that the AR and AUROC, Engelmann et al. (2003) derived, are simple linear transformations of one another[35]. For these reasons, we shelve the CAP and AR going forward, opting for the academically more elevated use of the ROC and AUROC in the fields of medicine, engineering and, more recently, default detection.

A related validation concept to power is that of model *calibration*. Collectively, these can be thought of as jointly revealing the goodness of fit and classification ability of the model. Calibration indicates the degree of fit between predicted probabilities and actual outcomes within the estimation sample (Stein, 2007). Predicted default probabilities that better match actuals provide for more accurate default rates. While discriminative tests and statistics, such as the AUROC, measure power, tests of calibration are tests of levels that may be distorted by "highly correlated default events or if the data represent only a portion of an economic cycle" (Stein, 2007, p.86). During model estimation using maximum likelihood, the log-likelihood can *de facto* be read as a measure of in-sample calibration, with higher (i.e. less negative) values suggesting greater calibration. For ease of review, however, a simple comparison of predicted mean default rates vs. the actual mean rate of bankruptcy sufficiently tells the same narrative. In sum, the ROC and AUROC evaluate a model's prowess in separating good credits from bad credits, while calibration appraises the model's capacity for achieving accurate probabilities of default. Both are inherently important validation dimensions in, e.g., deciding between model specifications or evaluating the economic benefit of each thereof. Nonetheless, as Stein (2007, p.87) would heed: "it is generally far easier to calibrate a powerful model to true default rates than it is to make a weak but well calibrated model more powerful".

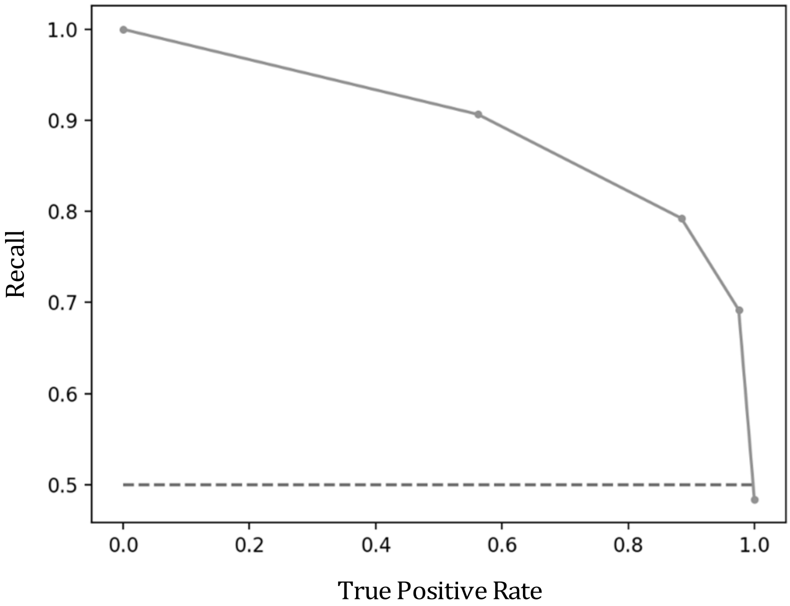
A related concept to that of TPR and FPR is the *positive predictive value* (*PPV*) or *precision (*see Equation 7).

$$PPV = \frac{TPs}{TPs + FPs} \tag{7}$$

Precision, as opposed to the true and false positive rate, measures the frequency with which a positive class is correctly classified, i.e. predicted. Phrased differently, it captures the fraction of correct predictions among all positive predictions, or the probability of bankruptcy given a predicted bankruptcy. When used together with recall (i.e. the TPR), the two may serve to comparatively indicate the degree of (mis)calibration at work in the estimation sample (Oommen et al., 2011). Represented graphically, this two-dimensional space is called a Precision-Recall (*PR*) curve. In spite of its merits, there is a blatant disregard for PR-analyses of this sort in the existing default literature, even when applications relating to model calibration (and not just power) are investigated. We are unable to identify a single published paper in the extant literature on public and private companies alike that, at the very least, includes a PR curve or similar discussion. Such is not the case in the medical academia and field of engineering, however. Perhaps – akin to the history of the ROC in the default literature – the PR curve is awaiting its milestone use in bankruptcy studies. Figure 2 shows a PR curve and its most essential features.

[35] Specifically, $AR = 2 \cdot (AUROC - 0.5)$

Figure 2: A Simple PR Curve



Here, the PR curve has the TPR (recall) on the x-axis – equivalent to the ROCs y-axis. Its interpretation, like the ROC, can be succinctly summarized into an *area under the PR curve* (AUPRC). It too has an intuitive reading as the average precision across all cut-off thresholds. Unlike the ROC's no-better-than-random constant baseline (a 45° line), the PR curve's baseline is the proportion of positives in the estimation window (e.g. 0.5 for a perfectly balanced sample), and as such varies with the class distribution of the sample (horizontal dashed line in Figure 2). Therefore, in rare-event studies the AUPRC may appear as disquietingly low relative the AUROC, but this is just a natural consequence of the PR curve as an expression of precision relative the sample distribution (i.e. class imbalance). Any point above the PR baseline (i.e. sample proportion of positives) is interpretable as a cut-off coordinate for which a model performs better-than-random.

Saito and Rehmsmeier (2015) provide the most comprehensive review and assertions relating to precision-recall, that we know of. The authors contest that in the face of class imbalanced data where a small minority class (e.g. defaults) is dominated by a larger majority class (e.g. non-defaults), ROC plots and interpretations can be deceptive and misleading. PR curves being evaluations of the fraction of true positives among positive predictions, however, can afford the researcher a view less contaminated by class imbalance. Quintessentially, to quote the authors, "the PR curve plot is more informative than ROC when evaluating binary classifiers on imbalanced datasets" (Saito and Rehmsmeier, 2015, p.12). The AUROC is insensitive to the estimation sample's degree of imbalance, while the AUPRC is not, wherefore the latter is more readily able to expose differences in e.g. early-retrieval performance (i.e. cut-off points for which there is relatively low 'costs' of misclassification). Moreover, the AUPRC awards the researcher no benefit for correctly identified negatives (survivors), why it is less prone to exaggerate model performance in imbalanced datasets where negatives dominate positives (Sofaer et al., 2019).
Unfortunately, the PR curve does not exhibit all the favourable econometric characteristics of the ROC (e.g. some coordinates on the PR curve are unobtainable, and points are non-linearly

interpolated). For an excellent walkthrough of these and the qualities of ROC and PR curves, the reader is referred to Flach and Kull (2015) who suggest employing Precision-Recall-*Gain* curves in place of regular PR curves.

2.6 Class Imbalance and Sampling Bias

Rare-event studies including the likes of bankruptcy prediction suffer inherently from *class imbalance* where the number of defaults (the minority class) are but a fraction of the number of survivors at any point in time (the majority class). A source of bias, in this regard, occurs when the estimation sample fails to mimic the proportions of the population minority and majority class (*sampling bias*). All matched-pair studies (e.g. Altman, 1968; Beaver, 1966) therefore suffer from sampling bias.

Econometric research has focused its exploits towards understanding, measuring and amending the impact of sampling bias and class imbalance on the predictive power of classification models (Oommen and Baise, 2011). Veganzones and Séverin (2018) empirically show that any class imbalance greater than an 80 / 20 split significantly deteriorates the power of a model. The performance loss (AUROC) thereof can be partially recovered by various sampling techniques including random oversampling examples (*ROSE*) and the synthetic minority oversampling technique (*SMOTE*)[36]. The higher the class imbalance and the smaller the training data, the greater the relative recovery rate relative the balanced dataset. Machine learning paradigms, expectedly, such as the *support vector machine* (SVM) are most capable in this regard. (Un)fortunately for those with estimation windows larger than 4 000 firms, the relative additional performance recovery flattens and quickly becomes marginal. Through these oversampling mechanisms, therefore, the performance recovery is largely capped at around 50% depending on the modelling technique used (*ibid.*). Through an extensive study on six variations of logistic regression, Ogundimu, (2019) echoes Veganzones and Séverin's (2018) result. He synthetically illustrates that SMOTE is the preferred choice when optimising model discriminative power (AUROC) while ROSE is desirable when maximizing the AUPRC.

What of sampling bias? As Oommen and Baise (2011) convey, no definitive conclusion is observable in the body of statistical literature jointly examining the effects of class imbalance and sampling bias that paradoxically trade-off each other[37]. Nevertheless, through a rigorous set of Monte Carlo simulations, the authors show that replicating the population distribution in the sample (i.e. limiting sampling bias) produces the most calibrated predicted probabilities in the maximum likelihood setting. They also suggest that neither class imbalance nor sampling bias significantly distort a model's power - i.e. its ability to separate between classes[38].

With this in mind, it seems the prudent researcher should always (data size allowing) strive to reproduce the population distribution in sample – i.e. prioritize concerns of *sampling bias* ahead of *class imbalance*. Rather than re-fitting a model on an unbiased sample mimicking the

[36] Oommen and Baise (2011) report an average recovery rate of 43.9% across several types of models, including logistic regression (the binary response model using the logit link function)
[37] A more class balanced sample may, in rare-event studies, cause a higher deviation from the population proportions of the majority and minority class (somewhat oversimplified)
[38] Reducing sampling bias by over- and under-sampling techniques only provide slight improvements in the probability estimates. To quote the researchers themselves: "The goal of sampling should be to mimic the population class ratio in the sample" (Oommen and Baise, 2011, p.118)

population[39], Skogsvik and Skogsvik (2013) provide a poignant solution to sampling bias, usable at the *ex post* stage of modelling. They eloquently suggest adjusting the predicted probabilities *a posteriori* relative to the degree of experienced sampling bias as per Equation 8. As Skogsvik and Skogsvik, (2013) themselves concede, minimizing sampling bias in this manner does not improve the model's ability to discriminate between classes – but it does improve calibration.

$$P_{fail}^{(adj)} = \left[1 + \left(\frac{1-\pi}{\pi}\right) \cdot \left(\frac{prop}{1-prop}\right) \cdot \left(\frac{1-P_{fail}^{(prop)}}{P_{fail}^{(prop)}}\right)\right]^{-1} \tag{8}$$

Briefly, π is the population proportion of the minority class; $prop$ is the estimation sample's distribution of defaults to non-defaults and $P_{fail}^{(prop)}$ is the (biased) sample-generated predicted probability of default. These culminate in $P_{fail}^{(adj)}$: the sampling bias-adjusted (i.e. population calibrated) probability of default. Gruszczyński (2019) corroborates the usefulness of the *Skogsvik adjusted* bankruptcy probabilities whose formulation coincides with previous logit-related probability corrections. Nonetheless, the vast majority of corporate default research is silent on sampling bias and, more importantly, what measures if any are implemented in rectifying these.

2.7 Economic Cost of (mis)classification and Basel III

Inaccurately classifying a company as a defaulter (i.e. committing a Type I error) comes with a certain cost. For a lender, it could mean losing out on interest income. Private equity investors might step away from a profitable investment, and a vendor might forgo revenue by overestimating counterparty risk. These are costs borne from (over-)prudence - *opportunity costs* from conservatism. More often than not, however, the loss incurred from a Type II error is considerably higher than that of a Type I error: Type II errors often entail significant economic losses for lenders and/or equity-holders. Several studies, (see e.g. Stein, 2005; Blöchlinger and Leippold, 2005; Agarwal and Taffler, 2007; Agarwal and Taffler, 2008a; Bauer and Agarwal, 2014) have consequently assessed the cost of (Type I and Type II) misclassifications by developing evaluation methods that abridge the relationship between default predictions and lending profits - bringing an appreciated degree of practicability to the literature.

Stein and Jordão (2003) and Stein (2005) are quite theoretically inclined in this endeavour. As Blöchinger and Leippold (2005) later explain, Stein and Jordão (2003) and Stein (2005) actually provide empirical evidence for how different default models create different economic impacts when applied for a lender. Stein (2005) simulate lending revenues, losses and profits for banks based on historical loan performance and the financial statements of middle-market firms. Blöchlinger and Leippold (2005) advance the misclassification literature further by deriving the profit-maximizing cut-off and the pricing curve, which in turn make it much easier to analytically explore the economic benefits of new default models. They, essentially, develop and derive a

[39] Which may force the researcher to forego some data

simple equation for the credit spread of a bank, later applied by Agarwal and Taffler (2008a) and Bauer and Agarwal (2014) among others.

Blöchlinger and Leippold (2005) also thoroughly explore the topics of *cut-off regimes*, *pricing regimes* and *mixed regimes*. A *cut-off regime* is a world in which lenders have strict rules for when to and not to lend. This is usually as simple as granting loans to all firms above a certain credit score, and rejecting all those with scores below the same threshold. Loan grants per this approach all have the same, *fixed* pricing. The most basic use of ROC analysis in this regard provides guidance for setting these lending cut-offs, and as such, there is a logical connection between ROC curves and the cut-off regime made use of. In reality, though, risk-adjusted pricing is the most common practice, meaning a strict cut-off regime oversimplifies how the lending and borrowing markets work. A practicing bank is much more likely to set a loan price *continuously* and *relatively* according to loan applicant's credit score, the latter of which is determined by the probability of default. In this setup, any applicant accepting the proposed price is granted a loan. These forces make for a dynamic, competitive market that operates under the banner of a *pricing regime*. Clearly, determining the price that maximizes the risk-adjusted profits is the be-all and end-all for a bank in this setting. As Blöchlinger and Leippold (2005) explain, however, credit specialists question the realism of both the cut-off and the pricing regime. There are two reasons for this. One is that a higher (interest) rate quoted by a bank might actually *cause* a firm to default: that is, the loan risk premium suffers from some level of endogeneity. Another inconsistency in the realism behind the pricing regime specifically is that minute differences in quoted rates are unlikely to entice a borrower to switch bank. This likely stems from some degree of client-bank customer 'stickiness': a by-product of relationship banking. Blöchlinger and Leippold (2005) conclude that some *mixed regime* (a combination of the cut-off and pricing regimes) best meet these critiques. Under this premise, riskier credits demand higher lender credit spreads and banks restrain from lending to some portion of the market's riskiest customers.

Two years after Blöchlinger and Leippold's (2005) paper made public headway, Agarwal and Taffler (2007) examine the long-term predictive ability of the Z-score (Taffler, 1983) in the UK public space. Specifically, they compare how two banks using different default prediction models - one using the Z-score covariates, the other using only a directional dummy on profit before tax - perform in a zero-sum, competitive environment[40]. A simple yet elegant test, this marks an important methodological milestone in which lenders' financial performance (in terms of market share, share of defaulters, revenues, credit losses, profits and return on capital and return on risk-adjusted capital) can be evaluated on an aggregate basis for an entire evaluation period.

One year later, the same authors Agarwal and Taffler (2008a) compare the profitability of two hypothetical UK banks that use different bankruptcy classifiers in making credit decisions. The first makes use of Z-score (Taffler, 1983) and the second employs the distance-to-default contingent-claims model (Merton, 1974) respectively. They find that the Z-score approach outperforms the contingent-claims model in a competitive pricing regime for listed companies[41]. The authors estimate bank profitability using the Basel requirements for exposure at default

[40] In fairness and for the sake of completeness, a third bank using a proportional chance model was also examined. It randomly classified firms as failures or non-failures solely based on the *ex post* failure rate of the population

[41] Since Black-Scholes-Merton- and Z-score-produced probabilities are inconsistent in the discrete hazard setting when used as independent variables, Agarwal and Taffler (2007) have to use the method developed by Hillegeist (2004) in converting these into non-parametric ranks or ''scores'' using the inverse logistic function

(EAD), loss given default (LGD), and risk-weighted assets, applying the internal ratings based approach (F-IRB) in which the probability of default determines capital requirements. A credit spread floor – the spread that the bank earns on its most creditworthy customer – is assumed constant for the entire period, and a minimum probability of default is set by Basel regulation[42].

Finally, Bauer and Agarwal (2014) later complement their previous research by extending the model-scope to include hazard models. Unsurprisingly, they find that said models subsume all other models both econometrically (i.e. in ROC space in terms of information content tests) and economically (i.e. regardless of the costs of misclassification assumed for bank loans to evaluated firms). The hazard models used were Shumway's (2001) and Campbell et al.'s (2008) covariate specifications, where Shumway (2001) was clearly the stronger of the two. The usefulness of Shumway's (2001) model was particularly evident in relation to the Z-score and contingent-claims models when evaluated on risk-weighted assets computed under Basel III: the, as per the authors, most important metric to measure and compare when evaluating a bank's capacity to deliver economic, risk-adjusted returns. Risk-weighted assets is the preferred credit benchmark on account of its direct connection with capital requirements. E.g. lower risk-weighted assets implies lower capital requirements which, *ceteris paribus*, increases banks' return on equity.

Although Bauer and Agarwal (2014) and Agarwal and Taffler (2008a) use all available data for 30 years in a large domicile (the UK), the number of defaults in their studies stack up to just 274. The yearly count of defaults is, additionally, highly volatile, ranging from 20 to just 1 (i.e. 0.1% to 2.34% of the sample observations). These descriptive details highlight that significant empirical improvements to their testing and benchmarking exercise can be realised by virtue of, simply, a greater sample size as afforded by non-listed private companies.

Using Korean data on 1,759 default events from a sample of 29,894 Korean firms, Duan et al. (2018) are, to the best of our knowledge, the only researchers that use the previously discussed techniques (setting a level of lending cut-off that minimize the cost and maximizes profits) developed by Stein and Jordão (2003) and Stein (2005) to estimate the different misclassification costs of bankruptcy models in the private firm space. The methodology differs from that used by Agarwal and Taffler (2007), Agarwal and Taffler (2008a), and Bauer and Agarwal (2014) insofar as Duan et al. (2018) do not consider a competitive banking environment. Nor do they consider how Basel requirements on risk-weighted assets affect return on assets and return on risk-weighted assets for the lenders. Rather, they focus on the increased predictive power arising from the integration of a public-firm equivalent distance-to-default variable in the private space (and the national savings thereof).

Consequently, we find no published research evaluating default models by examining the economic cost of misclassification from a lender's perspective acting in a competitive environment[43] in the private company default literature. Moreover, the research on public companies and the default prediction models tailored to them are evaluated by assuming (i) a constant size of the loan market, (ii) a constant spread for the most creditworthy firm, (iii) that each loan is of equal size and constant across years, and (iv) that all firms are willing to raise debt.

[42] Moreover, revenues, losses, profits and market shares are calculated by simply assuming the total size of the UK loan market is £100bn. Each loan is also assumed constant across time in the evaluation period, and of equal size

[43] i.e. applying Basel III requirements and quoting an interest rate (i.e. the price of the loan) which the prospective borrower compares to other rates offered by other banks at the same point in time

3. Data and Methodology

3.1 The Databases - Serrano, PATlink and Others

The bulk of the underlying data for the empirical analysis herein is derived from the Serrano Dataset (*Serrano*), obtained from the Swedish House of Finance. Serrano is a non-exhaustive combination of legal company data from several public and non-public sources including the Swedish Companies Registration Office, Statistics Sweden and Bisnode. The first two are Swedish governmental agencies, while the latter is a Swedish private research company and the direct beneficiary and proprietor of Serrano. As such, Bisnode continually updates, compiles and imputes, where necessary, Serrano bi-annually. Serrano is non-exhaustive insofar as it does not by default include all data entries from said suppliers of data. Instead, it also makes available the sources' raw data collections should the researcher be interested in additional information intentionally kept out of Serrano. Serrano is directly accessible to us through the Stockholm School of Economics and its close affiliation to the Swedish House of Finance.

Financial and non-financial accounting data make for the majority of Serrano and principally originate from the Swedish Companies Registration Office and Statistics Sweden. The latter organisation provides information on mergers, liquidations, restructurings and bankruptcy as per their legal definitions in Swedish Company Law. Other exogenous group data is made available via Bisnode's internal group register. Serrano is neatly organized such that every row of information is uniquely identifying for a calendar-year. On the organisational level, this means a corporate ID-number and a calendar-year are together uniquely identifying for the entire dataset: ergo, panel data with firm-year observations.

Serrano adheres to an explicit hierarchy of rules in compiling, imputing and updating data where the order of operations becomes important for the integrity of the database. This governing framework is especially necessary in Serrano's ambition to accurately reflect calendarised financials for companies. Bisnode explains: *"An important part of the framework handles the necessity to, to a certain degree, transform and modify the underlying register data into comparable calendar year values. That primarily applies to the income statement and balance sheet included in the financial statements, but other register data is also translated into calendar year values as needed.* (Bisnode, 2015)

Bisnode, hence, adjusts and corrects data entries for phenomena including (i) broken accounting periods[44], (ii) accounting periods shorter / longer than twelve months, (iii) omissions and gaps in a company's time-series of submitted financial statements, (iv) imputation for the most recent year's calendar-year values, (v) conversion to calendar-year values for balance sheet stock data and income statement flow data, (vi) rules for determining whether a business is active or not, and (vii) rules for what a newly started company is (*ibid.*).

Several variables of interest captured in company annual reports, such as whether an AGM[45] record was produced, the nominal value of pledged (collateralized) assets, and the CEO salary, have been excluded from Serrano. These are retrieved from the data-suppliers' raw data

[44] i.e. fiscal years that do not coincide with the Gregorian calendar
[45] Denotes "Annual General Meeting"

catalogues and matched into the Serrano dataset. Duplicates are few, easily identifiable and deleted.

Several other variables examined in this study are not included in either the Serrano dataset or its input data collections. These include annual Swedish macroeconomic variables and are effortlessly retrievable electronically from Statistics Sweden and the Swedish Central Bank (*Riksbanken*). Additionally, we have access to data on the company count of patent applications through a separate channel (PAtlink) also originating from the Swedish House of Finance. PAtlink is updated annually and contains a comprehensive information set of patent information organized by corporate-ID and calendar-year, wherefore integrating its contents with Serrano is a relatively painless exercise. Patent applications provide insights into the innovative churn of private companies.

Lastly, we have received access to board-level data through the faculty of the Stockholm School of Economics. This dataset's unique identifier is an individual's anonymized ID by company-position-year. As such each row contains information about a specific individual's main, second, third and fourth function on the board (where applicable) for a specific company for a specific calendar-year. Although a more onerous task, we transform this dataset into unique firm-years with variables for specific positions on the board[46] containing individuals' unique IDs, in turn allowing for assimilation into Serrano. Doing so gives us the option to create an array of various corporate governance-related covariates, including but not limited to the discrete (and cumulative) count of auditor changes, the CEO's salary relative non-executive employees, multiple directorships, board track-record[47], board composition and count, to name a few.

3.2 The Population of Private Firms

Data (un)availability has meant that most research on default prediction does not make use of the *de facto* unit population in its study. In order to be as permitting (and generalizable) as possible, this paper applies the widest possible definition of the private company population that is economically meaningful and utilizable from both a theoretical and practical standpoint.

First and foremost, financial and real estate companies are *excluded* as they are often analysed using completely different model specifications as a result of their disparate accounting standards, financial reporting requirements and risk evaluation profile[48].

Secondly, we examine exclusively Swedish *limited liability* companies (so-called *aktiebolag*) and thus discard shipping partnerships, mutual funds, partnerships / limited partnerships, joint-stock banks, insurance companies, cooperatives, public entities, primary municipalities, local federations, country councils, social insurance offices, public cooperatives and institutions, mortgage associations, regional state agencies, savings banks, and foreign legal entities entirely.

Thirdly, we restrict the business ownership to that of the *private* sort, i.e. limited liability companies that are not under state, regional, or municipal control or ownership on the premise

[46] Specifically, the positions of CEO, Chairman, Vice President and Auditor
[47] e.g. participation in previous defaults
[48] Chava and Jarrow (2004) is the primary exception of note to this stylized fact

that organisations of these types have an innate tendency to stay solvent and avoid bankruptcy in even the most misfortunate of circumstances.

Fourthly, only *independent* companies are analysed. Serrano collects only the annual accounts of *legally reporting* entities and makes no effort to consolidate these into group accounts in the event of a parent-subsidiary relationship. Consequently, it would constitute poor economic practice to include these in the estimation sample when it is well-known that company groups regularly micro-manage the accounts of parent companies and subsidiaries alike on tax, operative or other grounds. Companies with overseas enterprises, moreover, would be irreconcilable to any consolidation efforts as their foreign branches' accounts are not available to us - Serrano being an exclusively Swedish database. Therefore, we are forced to eliminate any corporate-IDs that have membership in a wider group structure. We maintain, however, that the accounting for a Swedish independent limited liability company would have the same economical interpretation as the consolidated financial statements of a group. As such, there should be no reason why the power and calibration of the models discussed in this paper do not also hold water if applied predictively to Swedish group accounts[49].

Finally, we drop company-year observations in which companies are *inactive*[50]. After this, we also remove firm-year data for which any company has been active for just one year during the entire 1998-2017 period[51]. A firm is considered active if it has more than SEK 10k in turnover and more than one employee, or - if it meets the first but not the second criterion - if it has other operating income / financial income in excess of SEK 10k, financial expenses below SEK -10k, total dividends greater than SEK 10k or total assets above SEK 500k. Allowing for companies to have less than SEK 10k, or approximately EUR 1k in turnover is clearly a very broad definition of what can be categorised as an active company.

With the above restraints in place, *any model fitted on this refined dataset is applicable to any active Swedish, non-financial, independent limited liability company with at least one year of data.*

3.3 Data Handling and The Order of Operations

Managing and sifting the above large datasets to (i) avoid losing "future" information, (ii) maintain accounting and calendar-year consistency between datasets, and (iii) retain consistency on what information is deemed non-missing or irrelevant, requires some degree of attention. As for Bisnode, when manipulating and transforming the data our imposed order of operations becomes essential in preserving the cohesion of the information set.

For illustrative purposes, consider a company that is active, enters into a two-year period of inactivity, and defaults immediately thereafter in the forthcoming year. Crucially, prior to cutting out inactive years, this company's default is captured in a 5-year, 4-year, 3-year, 2-year and 1-year leading binary default dummy. Removing inactive years directly without this foresight would have disseminated this company's default information completely[52]. Executed on the aggregate

[49] With the small caveat that Swedish conglomerates with a substantial portion of their operations or asset base located abroad may - by differences in local market conditions - be less predictable if subjected to models estimated on wholly Swedish data

[50] Inactive firms make no reports to any governmental authority and become, in the context of Serrano, missing

[51] An econometric *comme il faut*, since we are unable to predict into the future any company's survivability and default probability when limited to a single firm-year count for a corporate-ID

[52] Bankruptcies occur overwhelmingly on years without data. Therefore, they, too, constitute inactive years

level, and the majority of company defaults would be completely foregone while the sample default rate as determined by total firm-years would be heavily contorted vis-à-vis the population. After removing inactive years, growth rates in covariates, average balance sheet items and cash flows built from changes in balance sheets are calculated only for company-years exhibiting two or more consecutively non-missing (i.e. active) calendar-years.

On the backbone of our concise literary review, financial, non-financial and macroeconomic covariates are considered for a parsimonious model specification. These can be loosely grouped into three broad variable categories: (i) financial ratios (ii) non-financial information, and (iii) baseline hazard rate variables. In line with previous research, accounting ratios in different subcategories are included to allow for a parsimonious model to be built by including less *a priori* correlated risk characteristics. In total, we review in excess of 200 covariates[53] (see Table A1) for these purposes.

Financial ratios [(i)] can be further subdivided into fifteen separate groups: Profitability, Activity, Coverage, Efficiency, Growth, Leverage, Liquidity, Cash Flow Metrics, Accounting Quality, Changes in Ratios, Relative Performance, Returns, Dividends, Size & Age and Other. Combined, these make up 192 covariates. Following Altman and Sabato (2007), most ratios - where applicable - are also tested in log-form, both separately and (in the final model) in combination with non-logarithmised variables. Metrics related to returns such as return on capital employed and activity metrics such as the cash conversion cycle and capital employed turnover are calculated using averages of beginning and ending balance sheet items, except when only an ending balance sheet is available for newly formed companies (*left truncation*) or the year is 1998 - the first year of the study. Since Serrano does not contain cash flow accounting data[54], change in operating working capital and tangible fixed assets are used in conjunction with earnings to estimate cash flow from operations and free cash flow among others.

Non-financial information [(ii)] covariates refer to four groupings: Contagion, Reporting, Corporate Governance, and Patents. Several of these non-financial variables have so far received limited attention, or none altogether, in conjunction with financial variables. Examples of these include the patent application count, patent stock, industry-municipality interaction of defaults, company default track-record of key board members, records from AGMs, loans to related parties, severance pay mechanisms, shareholder and other contingent contributions, granted overdraft facilities, and many others. Several non-financial statement based variables examined in this study have, on the other hand, been used and selected for the final model in previous research (e.g. Altman et al., 2015; Altman et al., 2012).

The baseline hazard rates [(iii)] under review include a set of nine macroeconomic variables from Statistics Sweden and the Swedish Central Bank. These are all derivations of interest rates, currency exchange rates, GDP[55] and inflation. We also investigate the yearly change in these variables separately (e.g. Nam et al., 2008). Finally, the trailing 12 month realized default rate in the appropriate estimation period is examined as a potential baseline hazard variable (e.g. Hellegeist et al., 2001).

53 Not counting explored baseline hazard rates
54 Swedish private firms are required by law to report an income statement and balance sheet, but not the final cash flow statement (Bolagsverket, 2018)
55 Denotes "Gross Domestic Product"

Certain variables included in this paper are designed to approximate public bankruptcy prediction models' variables used in influential papers. For example, Shumway (2001) include idiosyncratic standard deviation of each firm's stock returns as well as stock market performance vs. the market. In the absence of market prices, we use historical company-level standard deviation of certain key ratios as well as the coefficient of variation of turnover, EBIT and net income as a proxy for idiosyncratic risk. Furthermore, to evaluate stock market out- or underperformance vs. the market, we use the yearly change in relative return on capital employed for a specific company vs. the entire sample. Moreover, we use historical average operating leverage[56] for the company and its industry to evaluate the unlevered risk (volatility) of earnings.

An extensive list of covariates notwithstanding, there are certain variables that have been successfully used in private company default domain which we are unable to study. A notable example includes the *county court judgment* (CCJ), often used in UK studies (Altman et al., 2012), and arises from a claim made to the court in the event of non-payment of unsecured debt. In the case of a creditor's claim being seconded by the court, a CCJ is issued and the debt (normally trade credit) must be settled accordingly. Company payment history and the private payment history of board members are two further explanatory variables we did not have access too[57].

All variables investigated[58], their respective definitions and example papers where they are used are located in Table A1 in the appendix.

Given the sheer size of the data and number of covariates (for the sake of maximum likelihood convergence and other econometric concerns) the latter, where appropriate, are winsorised (truncated) at the 1st / 99th percentile prior to any univariate or multivariate analyses.

3.4 Selecting the Baseline Hazard Rate

The DTDDH model necessitates a baseline hazard rate. Nineteen regressors are univariately examined in a DTDDH setting during the 1998-2013 estimation window. We account for shared frailty and recurrent events (multiple failures) using cluster-corrected standard errors, where the firm is the cluster. Directly incorporating a frailty term is problematic in survival analysis when (i) data-gaps exist between active / inactive years and (ii) a firm's earliest observational measurement does not correspond to its date of incorporation (*delayed entry*), but cluster-corrected standard errors achieve the same effect in controlling for intragroup event-time correlation. To save on computational time, we limit maximum likelihood iterations to fifty but use the stricter *Efron* method of handling tied events (defaults occurring at the same point in discrete time)[59]. The baseline hazard variables are then ranked and ordered on their *pseudo R-squared*[60].

The baseline hazard rate is investigated only once, given its naturally occurring temporal dependence. The selected baseline hazard rate is then used consistently and without replacement

[56] The mean of historical changes in EBIT in relation to changes in sales
[57] An important tool for practitioners (e.g. the private Swedish credit reference agency UC) and academics alike (e.g. Altman et al., 2015)
[58] Save for some logarithmised variations thereof
[59] A more accurate approximation of the marginal likelihood in the risk set than that of Breslow (1975)
[60] Not to be mistaken for a measure of the goodness-of-fit comparable to the variance-minimizing R-squared of Ordinary Least Squares regression. Rather, a higher pseudo R-squared is intelligible as a higher model fit in the broadest sense of the word

for all six competing models across all five default prediction horizons. Henceforth, our own one-to-five-year model specifications are referred to as the *Composite Model* (CM).

3.5 Selecting a Vector of Predictors - the Composite Model

The model selection process outlined below is done separately for different horizons of default prediction, i.e. the covariates are evaluated five times, once for one-year default predictions, once for two-year in advance default predictions and so on. Moreover, all regressors are selected on a subset of the total number of companies under observation (*sampled firms*) during a set period of time (*in-time*). Sample splitting techniques such as these allow for hold-out testing of models as a means of cross-validation. Collectively, the sampled firms in-time make up the *estimation sample*[61]. The training data represent 60% of the total population of firms, randomly selected and stratified (as closely as possible) on the population annual default rate. In narrowly maintaining the population distribution and proportions, we limit the distortive miscalibration effects imposed by sampling bias. This random firm-selection never changes across forecast horizons.

In the one-year forecast horizon, the in-time time-period is 1998-2013, leaving a three-year out-of-time interval of 2014-2016[62]. For each of the four remaining forecast horizons, the out-of-time window is consecutively kept as a three-year duration of time by shifting said interval backwards as appropriate. When doing so, the estimation sample's in-time scale is reduced accordingly. Consider the one-year forecast relative to its two-year sibling to understand this dynamic. The first, as we know, has an in-time / out-of-time period of 1998-2013 / 2014-2016 respectively. The second, by construction, has an in-time / out-of-time period of 1998-2012 / 2013-2015 instead. The final year being 2015 should hopefully come as no surprise to the reader, given that defaults occurring 2017 (the end of our data time-series) are detectable earliest in 2015 in the two-year forecast setting.

With the above in mind, there are four mutually exclusive sets of firm observations that constitute the entirety of our 1998-2017 panel data. These are *sampled firms in-time* (training data), *hold-out firms in-time*, *sampled firms out-of-time*, and *hold-out firms out-of-time*. Table 3 graphically represents these different sample splits for the one-year forecast, and Table 4 shows the absolute number of defaults in each sample split over time across all five prediction horizons.

Table 3: Firm-Year Observations (Thousands) by Sample Split for the One-Year Variable Selection

	Time Period																						
	In-Time																	Out-of-Time					Grand
	98'	*99'*	*00'*	*01'*	*02'*	*03'*	*04'*	*05'*	*06'*	*07'*	*08'*	*09'*	*10'*	*11'*	*12'*	*13'*	***Total***	*14'*	*15'*	*16'*	*17'*	**Total**	**Total**
Firm type																							
Sampled	48	51	52	51	51	51	52	52	52	54	54	54	55	56	56	57	*846*	57	58	57	41	213	***1 058***
Hold-Out	32	34	34	34	34	34	34	34	35	35	36	36	37	37	37	38	*560*	38	38	38	27	141	***701***
Total	79	85	86	86	86	86	86	86	87	89	90	90	92	92	93	94	***1 405***	95	96	95	68	**354**	***1 759***

The table shows the number of firm-year observations by sample split. Variable selection occurs on the sampled firms (845,578 in total) in-time (1998-2013). The two-year model's in-time window is 1998-2012 and its out-of-time window is 2013-2016. The out-of-time model fitting is made for the period 2013-2015 only since 2015 is the earliest instance for detecting defaults two years in advance. The three-to-five year predictions are adjusted in the same vein.

[61] Used interchangeably with *estimation window*, *estimation period*, *training data* and similar

[62] Although the data extends to 2017, leading defaults one period in the predictive setting implies no observable defaults in the final year

Table 4: Defaults by Sample Split and Forecast Horizon

	Sampled Firm, In-Time	Hold-Out Firm, In-Time	Sampled Firm, Out-of-Time	Hold-Out Firm, Out-of-Time
Default Horizon				
One-Year	6 027	4 003	814	558
Two-Year	13 946	9 448	2 354	1 521
Three-Year	16 408	11 118	2 808	1 904
Four-Year	14 647	9 896	2 647	1 818
Five-Year	12 454	8 367	2 449	1 650

This table reports the number of defaults by sample split, for all forecast horizons.

For lack of a unifying theory of private defaults and no empirical consensus on its determinants, we adapt an atheoretical standpoint in selecting the vector of (economically meaningful) covariates for the Composite Model specifications. Two parallel selection procedures are made use of. We dub these the *PR2* and *AME* methods of variable selection, respectively.

The *AME* methodology stems from Hosmer Jr. et al. (2013) and Gupta et al. (2017). It sees the implementation of the DTDDH model in univariately describing the *average marginal effect* (AME[63]) of each covariate at its mean relative the binary dependent. Predictors with higher absolute AME are thought to have higher explanatory (predictive) power in the multivariate default setting, and as such are *ceteris paribus* preferred over those with lower AMEs. In avoiding similarities between short-listed covariates in the multivariate context[64], correlations between AME-selected covariates are examined. If a high-correlation pair is identified, the covariate member therein with the smallest AME is dropped and replaced with the next-best AME-ranked covariate not yet short-listed. The process is repeated iteratively until no high-correlation pairs remain. In allowing for heterogeneity among firms (frailty), we cluster the DTDDH model-fitting on the organisational unit. To save on computational time, we limit maximum likelihood iterations to fifty for the one-to-three year forecasts, three-hundred for the four-year horizon and four-hundred for the final prediction length[65]. The number of integration points are set to seven[66] using the adaptive (penalizing) Gauss-Hermite quadrature regardless of the forecast horizon.

The *PR2* selection scheme re-visits the model fitting delineated in the selection of the Composite Model's baseline hazard rate, with one modification: univariate tests of the PH assumption by scaled Schoenfeld residuals are supplemented to the hazard model. Iterations are also limited to fifty for all forecast horizons[67]. Covariates are then ranked on pseudo R-squared twice. The first rank is limited to those regressors that meet the test for proportionality. The second rank is indifferent to whether the PH assumption is met or not. In this fashion, we are able to survey and prioritize covariates that univariately meet the underlying semi-parametric Cox proportional

[63] Univariately, AME can be read as the change in the probability of a positive (default) for a unit change of the predictor

[64] i.e. preventing selected covariates from capturing the same cross-sectional variation

[65] And in so doing, most - but not all - univariate regressions converge

[66] Specific to logit models using random effects (i.e. our DTDDH model), Lesaffre & Spiessens (2001) find that ten quadrature points using the adaptive Gauss-Hermite approach are largely sufficient for multivariate maximum likelihood estimation. For our univariate investigative purposes, however, we settle for seven on the grounds of computational time saved

[67] In so doing, the vast majority - but not all - predictors converge

hazards distribution in discrete time. Thereafter, a second sweep of the univariate results is made - this time disregarding proportionality - for economically meaningful explanatory variables with high pseudo R-squared. Finally, as in the AME-method above, high correlation-pairs are scanned for and replaced iteratively as necessary until none remain. The reader is reminded of Allison's (2010) *sic*: violations of the PH must be delved into synchronously as meaningful predictors are selected.

As a methodological control, a two-pronged lasso[68] is also run on all covariates for the one-year default forecast using (i) cross-validation minimization (ii) adaptive / penalized two-step cross-validation minimization (iii) the minimized Bayesian Information Criterion. The selected model variables common to all three of (i)-(iii) are largely similar to those ranked highly in the AME and, particularly, the PR2 designs. For this reason, and on the grounds of severe computational time, we run no further lassos for other default models with longer prediction horizons.

Naturally, all univariate regressions are run for each forecast horizon's respective estimation window. We limit the Composite Model's final form to no more than ten variables[69] in order to keep its specification parsimonious and (hopefully) generalizable.

In composing the multivariate Composite Model for each horizon, the selected covariates determined by AME and PR2 are subjected to a final review. *Ex ante* believed-to-be-economically-similar variables in each forecast short-list of each selection method are randomly exchanged for a neighbouring (by rank) variable located in the same category. E.g. return on average assets in the one-year PR2-defined default prediction model may be swapped for the next-best pseudo R-squared covariate of its category, say return on average capital employed, and the multivariate regression re-run. No covariates using this approach are found to improve the model's power or calibration.

Finally, all predicted probabilities for each default horizon are *Skogsvik adjusted* to minimize the occurrence of any sampling bias. Additionally, all models for all horizons are fitted with and without random effects (i.e. contagion) on the firm-level.

3.6 Selected Competing Model Specifications

Each model specification of Altman's (1968) Z-score, Zmijewski (1984), Shumway (2001), Altman and Sabato (2007) and Dakovic, et al. (2010) is fit as a DTDDH model as per Equation 9 below[70].

$$\Pr\left(y_{i,t} = 1\right) = \frac{1}{1 + e^{-(\alpha_t + x_{i,t} \cdot \beta + c)}} \tag{9}$$

[68] Although lassos are more commonly employed in discovering multivariate relationships akin to some methods of machine learning, it can also be used to good effect as a means for variable selection

[69] Excluding the baseline hazard rate requisite

[70] All rival models are of course fit exactly as our own in all other econometric regards (e.g. with and without random effects etc.)

Where α_t is the baseline hazard rate. To compare apples with apples, the baseline hazard rate selected for the Composite Model is also used for these rival prediction models irrespective of forecast horizon[71].

Altman's (1968) $x_{i,t}$-vector of predictors for each firm and year is composed of:

$$\frac{WC}{TA} \leftrightarrow \frac{EBIT}{TA} \leftrightarrow \frac{BEQ}{TL} \leftrightarrow \frac{S}{TA}$$

The market value of equity is replaced here for the book-value of equity BEQ in 'privatising' the otherwise public model.

Zmijewski's (1984) *ditto*:

$$\frac{NI}{TA} \leftrightarrow \frac{TL}{TA} \leftrightarrow \frac{CA}{CL}$$

Shumway's (2001) *ditto*:

$$\frac{NI}{TA} \leftrightarrow \frac{TL}{TA} \leftrightarrow LNRELSIZE \leftrightarrow NICV \leftrightarrow \Delta RELROACE$$

Where; $LNRELSIZE$ is the relative size of the firm to the population of Swedish private firms in terms of sales (as opposed to market capitalisation); $NICV$ is the company's coefficient of variation of net income (used to operationalize SIGMA); $\Delta RELROACE$ is the period-to-period change in the firm's return on average capital employed relative the population (used in place of ABSRETURN) .

Altman and Sabato's (2007) *ditto*:

$$\frac{STD}{EQ} \leftrightarrow \frac{CASH}{TA} \leftrightarrow \frac{EBITDA}{ATA} \leftrightarrow \frac{RE}{TA} \leftrightarrow \frac{EBITDA}{IE}$$

Lastly, Dakovic, Czado and Berg's (2010) *ditto*:

71 Even though, incorrectly, not all of these authors themselves incorporate a baseline hazard rate in their own works

$$(REVANM > 0) \leftrightarrow AGE \leftrightarrow \sum_{J=1}^{10} IND \leftrightarrow DIV \leftrightarrow \left(\frac{BEQ}{TA} \mid \frac{BEQ}{TA} \geq 0\right) \leftrightarrow \left(\left(\frac{BEQ}{TA}\right)^2 \mid \frac{BEQ}{TA} \geq 0\right)$$
$$\leftrightarrow \left(\frac{BEQ}{TA} \mid \frac{BEQ}{TA} < 0\right) \leftrightarrow \left(\left(\frac{BEQ}{TA}\right)^2 \mid \frac{BEQ}{TA} < 0\right) \leftrightarrow LNSIZE \leftrightarrow LNSIZE^2 \leftrightarrow e^{-\frac{CASH}{CL}}$$
$$\leftrightarrow \left(\frac{EBIE}{ATA} \mid \frac{EBIE}{ATA} \geq 0\right) \leftrightarrow \left(\left(\frac{EBIE}{ATA}\right)^2 \mid \frac{EBIE}{ATA} \geq 0\right) \leftrightarrow \left(\frac{EBIE}{ATA} \mid \frac{EBIE}{ATA} < 0\right)$$
$$\leftrightarrow \left(\left(\frac{EBIE}{ATA}\right)^2 \mid \frac{EBIE}{ATA} < 0\right) \leftrightarrow \left(\left(\frac{EBIE}{ATA}\right)^3 \mid \frac{EBIE}{ATA} < 0\right) \leftrightarrow \frac{CL}{TA}$$

Where; $\sum_{J=1}^{10} IND$ is a collection of dummy variables for each industry[72]; $\left(\frac{BEQ}{TA} \mid \frac{BEQ}{TA} \geq 0\right)$ ought to be read as $\frac{BEQ}{TA}$ conditioned on $\frac{BEQ}{TA}$ being greater or equal to zero, and so on. It is worth pausing to note that Dakovic et al. (2010) may contest the direct use of their Norwegian private model in the Swedish space – indeed, their principal learning is that examining functional relationships by Kernel densities may reveal non-linear relationships for different parts of the survival distribution. But we are not interested in testing their model *methodology* per se, but rather their actual covariate selection.

The reader should note that when income statement items are expressed relative to balance sheet items, the average of the latter's ending and beginning period is used where possible to the authors' (and their models') benefit (e.g. $\frac{EBIE}{ATA}$, where ATA is the average of beginning and ending total assets).

Seen as parts of a whole, these five contesting models provide for a challenging landscape for the Composite Model to compete in across all five forecast horizons. Combined, they come from different time-periods, geographies, sectors, are parsimonious and non-parsimonious, employ linear and non-linear predictors and are used in both the public, private and SME corporate default caucuses. Analogously to the Composite Model, each of the competing models' default predictions are *Skogsvik adjusted* for every forecast horizon.

3.7 Model Performance Evaluation

Model performance is gauged across several dimensions. Three separate hold-out samples are used to assess each model's external validity in each forecast horizon. The harshest, most testing of these takes the form of the *hold-out out-of-time* sample. It provides insights into the model's possible degree of overfitting while corroborating its external validity to observations never before seen both organisationally and temporally[73]. Across both the hold-out samples and the estimation window itself, each model's power and, to a lesser extent, calibration are compared using ROC and PR curves as well as their summary terms - the AUROC and AUPRC.

[72] Although the authors have fifteen (15) industry dummies in their study, we have twelve (12) as per a Serrano-condensed version of the most recently updated Swedish Standard Industrial Classification (SNI) standard. Of these, missing SNI codes and the catch-all group "Other" are excluded, leaving ten (10) industries

[73] As do, to a lesser extent, the two remaining modes of hold-out tests (i.e. *hold-out firms in-time* and *sampled firms out-of-time*)

Although there is no scientifically derived gold standard for the least acceptable level of AUROC or AUPRC, Hosmer Jr. et al.'s (2013) rule-of-thumb thresholds are sometimes used as benchmarks the literature (e.g. Gupta et al., 2017). A model's power, in terms of AUROC, is considered good if between 0.7-0.8, and excellent if above the 0.8 earmark (*ibid.*).

Finally, in holding ourselves accountable to a higher econometric standard, each model for each forecast is tested for violations of the PH assumption in the multivariate setting[74].

3.8 Economic Value of Model (Mis)classification Using the Simulated Approach

Similar to Stein and Jordão (2003), Stein (2005), Blöchlinger and Leippold (2005), Agarwal and Taffler (2008a), Bauer and Agarwal (2014) and Duan et al (2018), we argue that (in terms of associated costs) it is unreasonable to assume that lending to a firm that fails is the same as failing to lend to a firm that does not. Therefore, the economic value of model (mis)classification is examined by designing a loan market in line with Bauer and Agarwal (2014), where the lenders quote a credit spread for a specific company in a specific year in a mixed (pricing and cut-off) regime, and companies choose the lowest available spread quoted to them.

The credit spread quoted by the banks to the private companies is calculated using the intuitive formulation below (see Equation 10). The baseline credit spread for the most creditworthy customer is k,[75] and a bank will require an additional spread equal to the probability of default (PD) multiplied by the loss given default (LGD) (i.e. the expected loss)[76]. Economically, the interpretation of the formula is simple: a bank sets a minimum loan price for a borrower solely based on the predictions of its internal credit model. Furthermore, higher credit spreads cannot be quoted since the market is competitive enough to push the expected NPV of the loan to zero[77].

$$R = PD \cdot LGD + k \tag{10}$$

In order to retain some notion of conservatism in the banking system's corporate credit exposures, Basel III enforces a minimum estimated probability of default of 5 basis points (bps). This is incorporated as a probability of default floor: estimates lower than 5 basis points are replaced for 5 bps. As per Basel III, the loss given default is set to 40% of the exposure at default[78], the latter of which is equal to the loan value (Basel Committee on Banking Supervision, 2017).

Due to (i) a limited amount of data with high swings in default rates, and (ii) evaluation of different *types* of default models (i.e. Accounting Based Models, Accounting and Market Based Models vs. Contingent-Claims Models), researchers in previous papers have had to make several computational adjustments between model fitting and model evaluation. Firstly, they have had to convert probabilities of default from Z-score models and BSM-models (i.e. Black-Scholes Merton-models) into logit scores to fairly be able to compare discriminatory power between models. Secondly, they have had to winsorise the probability estimates themselves (not just the covariates producing them). Neither of these adjustments is necessary in this paper. We are data-privileged,

[74] More precisely, each regressor is tested individually (and the model globally) for violations of proportionality
[75] Which is set equal to 0.3%
[76] The relationship between the lender and borrower is assumed to have no value. The left-hand-side of Equation 10 would otherwise have to be restated to $R + C$, where C represents the value of said relationship
[77] The lender does not invest in negative NPV projects according to its own, best *a priori* appraisal
[78] Prior research has made use of previous Basel requirements, e.g. a 3 basis points minimum PD and a 45% LGD floor (Basel Committee on Banking Supervision, 2017)

providing stable *Skogsvik adjusted* default patterns, and only hazard models are compared against each other. Bauer and Agarwal (2014) argue that since the credit spread is determined by both the likelihood of failure and non-failure, it is inherently dependent on both model power and calibration. In seeking only the merits of model power, the authors use bankruptcy probability percentiles (ranks) to compare competing model estimates. While we re-calibrate predicted probabilities towards the population distribution of class imbalance using the Skogsvik adjustment, all six competing models are equally treated in this regard: and so, differences in calibration between models remain, centred on the population default rate. This is important, as overall bank performance in a competitive setting is a joint-product of both model power *and* calibration - why we do not wish to isolate model power or, equivalently, erode differences in calibration.

One further detail with regards to how we model bank-quoted credit spreads must be noted. Each year's 5th percentile of least credit worthy firms (i.e. the 5% of companies with the highest probabilities of default in each bank's yearly evaluation) are precluded from any loans[79]. This decision rule stems from previous research and represents the *cut-off regime* component of the model. The rest of the model operates as a *pricing regime*, i.e. the probabilities of default affect the credit spread in a continuous fashion. Taken together, a *mixed regime* setting is at work.

With the bedrock for the competitive market in place, the competing models (i.e. Altman, 1968; Shumway, 2001; Zmijewski, 1984; Altman and Sabato, 2007; Dakovic et al., 2010; and the Composite Model) represent six hypothetical banks applying their respective default models to decide on what credit spread to offer each of the 95% most creditworthy firms of a certain year.

Following the method developed by Agarwal and Taffler (2008a), the bank's required assets and risk-weighted assets (RWAs) are calculated using the Basel III formulas for required capital (see Equation 11-15). The simultaneous simplicity and real-world accuracy of this calculation is appealing. First, the only input needed is the bank's internally estimated probability of default (which is why this approach to capital requirement estimation in banking regulation is called the *Internal* Ratings-Based (IRB) approach), and a simple adjustment is made for companies with a turnover below EUR 50m (see the last term of Equation 12). Second, all licenced banks that do not want to use an externally prescribed schedule of risk-weights for calculating RWAs will need to apply said formulae for all companies which they lend to (Basel III will be in full-effect by the 1st of January 2022[80]). In practice, there are two means that banks can make use of in calculating its risk-weighted assets: the Advanced IRB (A-IRB) and the Foundation IRB (F-IRB) approach. The main difference between the two being that, criteria allowing, a bank that uses the Advanced IRB approach will use its own LGD, and thus a potentially different RWA. These criteria are in part based on the size of the borrower[81] but more so on the bank's ability to demonstrate certain statistical qualities and more complicated risk methodologies[82] in their evaluation of PDs.

[79] In the nominal credit exposure approach outlined in the next segment of the methodology, this 5% cut-off assumption is relaxed for reasons that will later become apparent

[80] The members of the Basel Committee on Banking Supervision agreed upon Basel II November of 2010. The initial introduction period was scheduled between 2013 and 2015, but after a series of extensions, implementation is currently due 1st of January 2022. (Financial Stability Board, 2019)

[81] When Basel III is fully implemented, banks will not be able to use the Advanced IRB for companies with an annual turnover exceeding EUR 500m revenue

[82] E.g. that the minimum firm observation period should (ideally) cover one economic cycle. The use of the A-IRB approach often focuses on additional minimum requirements for off-balance sheet items: banks need established

Differentiating between or otherwise determining which of the competing, hypothetical banks may be interested in using the A-IRB approach is beyond the scope of this exercise. Instead, it is assumed all banks use the F-IRB approach.

The formulae for the Basel III requirements of RWAs are outlined in Equation 11 to 15.

$$Correlation_{Large\ Corporates} = R = 0.12 \cdot \frac{(1 - e^{-50 \cdot PD})}{(1 - e^{-50})} + 0.24 \cdot \left(1 - \frac{(1 - e^{-50 \cdot PD})}{(1 - e^{-50})}\right) \tag{11}$$

$$Correlation_{SMEs} = R = 0.12 \cdot \frac{(1 - e^{-50 \cdot PD})}{(1 - e^{-50})} + 0.24 \cdot \left(1 - \frac{(1 - e^{-50 \cdot PD})}{(1 - e^{-50})}\right) - 0.04 \cdot \left(1 - \frac{(s - 5)}{45}\right) \tag{12}$$

$$Maturity\ Adjustment = b = \ [0.11852 - 0.05478 \cdot \ln{(PD)}]^2 \tag{13}$$

$$Capital\ requirement = K = \left[LGD \ \cdot N\left[\frac{G(PD)}{\sqrt{(1-R)}} + \sqrt{\frac{R}{1-R}} \cdot G(0.999)\right] - PD \cdot LGD\right] \cdot \frac{(1 + (M - 2.5) \cdot b)}{(1 - 1.5 \cdot b)} \tag{14}$$

$$RWA = K \ \cdot 12.5 \ \cdot EAD \tag{15}$$

As outlined in the literature review, extant research assumes a constant size of the loan market, a constant spread for the most creditworthy firm, that each loan size is of equal size and constant across years, and that all firms are willing to raise debt. We term this experimental set-up as a *simulated approach*. We apply the simulated approach in the evaluation of our six models, but choose to show key credit metrics on both a yearly and aggregate basis.

3.9 Economic Value of Model (Mis)classification Using a Notional Credit Exposure Method

Serrano includes the liabilities to credit institutions on a yearly basis for all observed companies. Therefore, it is possible to evaluate the economic value of model (mis)classification using the notional credit exposures recorded for each firm-year observation. That is, the loan market size is found by summing the yearly, actual balance sheet values of short term and long-term liabilities to credit institutions. Then, using the model-based predictions, the same key credit metrics (market shares, revenues, losses, RORWA, etc.) shown in the simulated approach are calculated based on *actual* loan sizes. Clearly, for some firm-years there are no loans outstanding (all firms do not borrow). Consequently, the underlying assumption of this method is that a bank can only offer a loan to a company that has *de facto* chosen to borrow. While assessing credit risk based on

procedures in place for the estimation of these items, including the effect of additional drawings. A-IRB adapting banks even need specific estimates for each credit facility, including correlations between these historical data and actual exposures is hardly new to credit modelling, the literature on the cost of default model (mis)classification in which banks compete in a loan market has not made use of this method.

Using notional credit exposures calculated for every firm-year observation has many more advantages than just dynamically capturing the varying size of the loan market. In fact, this suggested approach has - in our view - at least six clear advantages over the Bauer and Agarwal (2014) and Agarwal and Taffler (2008a) procedure.

Firstly, a company that has not borrowed from a credit institution during a specific year is not compelled to borrow for the purposes of model evaluation. Thus, there is no potential bias introduced into the model evaluation in which the capacity of banks to make profit maximizing decisions are evaluated on companies that for various reasons would not seek debt capital and pay interest to a bank in the first place.

Second, this model does not implicitly claim that the most competitive loan automatically leads to a company-assumed loan. Many firms operate with sub-optimal capital structures, and some firms (e.g. many family-owned companies) are decidedly against raising debt: even if, from a cost of capital perspective, it would be a preferable to do so. As statistics will indicate, far from all companies raise debt even in competitive and well-functioning financial markets such as Sweden's. In fact, less than half of all company-years in this paper's population of independent, limited liability firms have bank debt. The cost associated with these bank loans represent ca 95% of all financial costs[83]. That is, if a firm does not have bank debt, it does not have debt at all.

Third, loan losses are evaluated with actual nominal weights. Thus, notional credit exposures (by being notional) are free from any size-related bias concerns that the simulated approach may suffer from in treating small and large companies alike. In this regard, using notional credit exposures captures if a model consistently underappreciates the risk of, say, large corporates with large loans or small companies with small loans. We find this an especially strong reason to adopt the nominal credit exposure method, since risk-weighted assets (and its accompanying required capital buffer) is a *de facto* function of company size (see Equations 11 and 12).

Fourth, the default event is matched with the right type of contractual obligation. Akin to Bauer and Agarwal (2014) and Agarwal and Taffler (2008a), we define default as the occurrence of different types of failures. In our paper, specifically, we combine bankruptcy, liquidation and restructuring events as different modes of default. Similarly, Agarwal and Taffler (2008a) define default as any circumstance of liquidation, administration / receivership or a valueless company. Regardless of how it is operationalised, under the simulated approach, a liquidation or becoming a valueless company *unilaterally* equates to a loss given default for credit institutions. When replaced with notional credit exposures, only defaults where actual losses are incurred by credit institutions translate into a loss given default.

Fifth, notional credit exposures permit the researcher to capture any potential temporally-tied forms of misclassification, e.g. cyclicality-related biases in model separability and calibration.

[83] Moreover, ca 90% of the time, costs to financial institutions represent all financial costs. Other financial costs can include financial write-downs, profit/loss from share trading, or currency exchange-related losses

Finally, instead of assuming a constant credit spread (i.e. k, Equation 10) for the most creditworthy customer (as is done in the simulated approach), actual data allows us to use external interest costs in relation to bank debt when finding the most creditworthy customer. In the notional credit exposure methodology, the 5th percentile of the firm's interest rate on bank debt represents the most creditworthy borrowers for each year: their corresponding credit spreads range from 4.0% in 2001 to 1.3% in 2017.

In the rear-view mirror, the actual balance sheets represent the state of the world after a lending decision has been made. As such, there is no need to incorporate any cut-off regime. That is, there is no need to exclude certain risky loans from this exercise through an arbitrary threshold, since the firms that do not have debt on their balance sheets represent companies that did not seek a loan, as well as those that sought loans but were not granted any.

Other than the above outlined revisions vis-à-vis the simulated approach, the notional credit exposure method proposed in this paper does not alter the estimation of PDs, nor the calculations of market share, profit and loss, return on assets (ROA) or return on risk-weighted assets (RORWA).

4. Descriptive Statistics

4.1 Default Rates

Recall that we limit the companies under study to private, independent limited liability companies outside of the financial and real-estate sectors. Prior to removing inactive and, thereafter, single company-years from this population, we enjoy a total of 5.3 million unique firm-year observations across 1998-2017. After the fact, however, these diminish to ca 1.8 million observations corresponding to 245,844 unique corporate-IDs. Many of the companies included herein, however, experience a default event or leave the dataset (e.g. by merger) on inactive years (*censoring*), i.e. years where data is missing. Equally, many companies experience periods of operational inactivity and then become active again. Including all of these inactive entries increases the observational count back from 1.8 to ca 2.7 million for the same 245,844 unique firms. The population of 245,844 unique companies, therefore, can be understood as 1.8 million non-missing firm-years or equivalently 2.7 million firm-years including inactive / missing years.

Of the larger 2.7 million firm-year observations, a total of 55,411 defaults are observed, corresponding to a cumulative population default rate of 2.12%. This is the *true* population default rate, as the vast majority of defaults coincide with inactive years taxed by missing data[84]: so much so, that only 9.6% of our 55,411 defaults take place in a year in which complete or close-to-complete financial data is available (see Table 7). 11,402 of these defaults transpire within 12 months after the release date of the latest financial statements. An overview of the total default count by industry and year is included in Table 5. Due to a large aggregation of IT-related resources in the Swedish Companies Registration Office during the late 90's, the default rate is artificially low during 1998 and most of 1999[85].

[84] Dakovic et al. (2010) made the same observation in Norway: 75% of their sampled defaults occurred at least one year after the last available financial statements

[85] We are informed of this unfortunate circumstance through private correspondence with Serrano's proprietor, Bisnode

The reader should heed that no model is fitted on the population including inactive years. She is also reminded that prior to removing these[86], coincident defaults are stored through leading default dummy variables. As such, all hazard models are regressed on the reduced-form 1.8 million firm-years without any loss of (bankruptcy) information[87].

There is a significantly higher amount of defaults caught in two- (27,269) and three-year (32,238) lead-time, vis-à-vis the one-year horizon count of 11,402. This occurs for several reasons. Firstly, there are just five months out of a potential twelve in a one-year default prediction model where one can safely ascertain financial statements are publicly available[88]. Secondly, we conservatively manoeuvre any data imputed by Bisnode. To see this, consider a company that has its fiscal year between September and August. Suppose now that it reports its annual accounts for 2014 and 2015, and later defaults on the 1st of December 2016. Since the firm has a broken fiscal year, for the calendar-year ending 2015 Bisnode will impute the firm's 2015 and 2016 annual statements such that the calendarised data entry for 2015 includes one-third of the company's 2016 annual accounts, corresponding to the four missing months of September-December of 2015. Consequently, there is information included in the 2015 imputed calendar-year from the next year's annual statement. That very same statement, in turn, can be reported up to seven months after 2016's fiscal year has ended - i.e. at latest in April of 2017 - which is after the default has transpired. To account for this eventuality, these observations are not considered by the one-year model. Both of these prudent efforts attempt to limit the potential for any inadvertent look-ahead bias. Any potential over-conservatism in this approach to one-year defaults can potentially be regained in the two-year prediction model, where a portion of the defaults are, in fact, closer to one-year-away than two.

In any case, the one-year horizon's (relatively) low number of defaults (11,402) produces a default rate of just 0.67% when compared to the population observational count - why the Skogsvik adjustment becomes particularly useful for the one-year prediction. For further summary descriptives on one-year default data, see Table 6. After the three-year default horizon, there is another slump in the number of defaults in four- and five-year lead-time, as the look-ahead window within the 1998-2017 period becomes shorter.

It is worth reiterating that although there are financial statements readily available up to and including 2017, the one-year model is not tested on 2017 data, since there are no leading defaults that year. Consequently, the aggregate of total defaults and non-defaults in Table 6 (1,698,731) is lower than the population observations excluding inactive years of (precisely 1,759,105 in number). By the same rationale, the two-year prediction has a lower observational count yet, and so on.

Of 245,844 unique companies, 190,822 never fail in the study (*right-censoring*), 54,633 companies default once, and 389 fail twice (*recurrent events*), meaning the 55,411 defaults in fact represent 55,022 uniquely defaulting companies. As such, 22.4% of population companies fail at some point in time during 1998-2017 (see Table 7).

[86] On the basis of either missing data or inactivity as per the criteria outlined in *3.2 Defining the Population*
[87] *Censoring*, as in most survival analyses, is deemed to be non-informative
[88] Swedish Private Company Law dictates annual reports must be recorded with the Swedish Companies Registration Office within seven months of the end of the fiscal year

4.2 Basic Company and Dataset Characteristics

The average revenue in SEKk[89] as well as the average company age in years are reported in Table 8. As one might expect, the average company size in the population has increased over the course of the last twenty years save for during the financial crisis, which temporarily decreased the average private company size. This observation is especially apparent in more cyclical industries, such as the industrial goods and construction industries, while the convenience goods and energy and environment sectors remained relatively unaffected during the same period - at least on aggregate. The IT and electronics industry seems to have experienced a downwards trend in its average turnover post the burst of the IT bubble. Average company age has, overall, increased slightly - perhaps firms are more attuned to the ebb and flow of the economy today than before.

Several variables examined in this paper consider geographical patterns, including the likes of contagion by industry and region. Noting the differences in default rates across Swedish counties, the latter were considered for the Composite Model's final form as both random effects and dummy variables. The average default rate by county can be retrieved in the appendix, under Figure A1. A peculiarity worth remarking is that the county of Dalarna exhibits a substantially higher default rate than the rest of Sweden's many counties.

The rate of new entrants (*left-truncation*) to the population averages ca 5% per year, climbing slightly higher in the aftermath of the financial crisis. Both in relative and absolute terms, more firms are seen entering the IT and electronics, construction and health and education industries when compared to the materials, industrial goods and energy and environment industries.

Variable Statistics

The quality of the compiled and imputed Serrano database makes for a unique opportunity in revealing descriptive insights into the Swedish population of private, independent limited liability companies. We believe these may be of great value to practitioners, academia and society at large. On these, if nothing else, altruistic grounds we include the mean values of over thirty-five variables in Tables 10.1-10.10 located in the appendix. These should provide a comprehensive view on the state of health of all active Swedish private companies that report financial statements (audited or not). This is presented by industry and year, making it possible to quickly find answers to questions such as:

- What is the average Gross-[90], EBITDA-, EBIT- and Net Income- margin of all private companies in Sweden today?
- Has the cash conversion cycle for all companies changed in Sweden the last twenty years, and if so, which of inventories, receivables and payables are driving this?
- Is there an industry that has consistently returned a higher average capital employed?
- How have the debt and net debt to EBITDA ratios changed during the past twenty years? Are companies becoming more or less leveraged?
- Do asset-light industries show the highest consistent operating cash conversion?

[89] i.e. thousands of Swedish crowns

[90] The gross margins are not based on actual COGS as 98% of the population companies report their financial accounts 'by nature', and not 'by function'. Consequently, gross profits are estimated as sales less production costs, where production costs are defined as raw materials and consumables + goods for resale + other external costs

- How has the partial repeal of private companies' audit obligation impacted the number of companies that choose not to have an auditor? How has this impacted the number of "not recommended" audited accounts and other audit remarks?
- Do scalable business models in the IT and electronics space increase their value added per employee at a higher rate than other industries?
- Has the growth in sales been offset by an equal growth in assets, or is the average business less capital intensive today?
- How much cash and liquid assets do businesses hold as a percent of sales in different industries?

It is comforting to see that the average values of our handful of selected metrics[91] are generally very stable year-to-year, although some long-term trends seem to be at work. These descriptive statistics speak to the quality of the data, since (i) industries display similar characteristics in terms of their financials over time, and (ii) these characteristics are what one would expect when comparing industries. For example, the more cyclical industries show the highest margin and return on capital deterioration during cyclical downturns, whereas their less cyclical counterparts, such as health and education and shopping goods, remain largely unscathed.

While these statistics might seem superfluous in the context of default prediction for individual companies, we argue national default prediction research needs to be put into context of the characteristics of the population. Hopefully, as the corpus of literature on private default grows across Europe and internationally, researchers include detailed descriptives which can help compare and contrast not just default rates, but also the distribution of financial variables over time and by industry. This would, for instance, enable multi-level default models to be estimated where the organizational unit can be nested in an industry and country.

5. Results and Discussion

5.1 The Baseline Hazard Rate

Of nineteen pursued nominees, the *trailing realized rate of default* is selected as the baseline hazard rate for all models in all predictions. Although none of our candidates produce pseudo R-squared results worth reporting, the three most econometrically prominent (highly ranked) covariates are (i) the ten-year Swedish government bond rate (ii) the Krona Index (KIX) and (iii) the trailing realized default rate. The KIX is a geometric index of several currency denominations relative to the Swedish crown, weighted by flows of processed goods and commodities therein (Swedish Central Bank, 2018). As a measure of the baseline hazard rate, basket exchange rates are of particular salience to open-economies with large export markets or popular currencies (see e.g. Nam et al., 2008). In any case, pseudo R-squared differences being marginal at best, we opt for the aforementioned trailing realized rate of default as our final selection. It holds intuitive appeal and is interpretable as the unconditional baseline hazard rate of the foregoing period. Moreover, it has a foothold in the literature through the likes of Hillegeist et al. (2001).

5.2 The Composite Models' Atheoretical Vector of Predictors

[91] More correctly, average values winsorised at the 1st and 99th percentiles

The *AME* method used favourably by Gupta et al. (2017) and recommended by Hosmer Jr. et al. (2013) reproduce, in terms of AUROC and AUPRC, a satisfactory set of variables for multivariate use by the CM. The *PR2* approach, however, dominates the former across all forecast horizons. We are unable to reconcile this idiosyncrasy, but believe it to be an artefact of temporally changing intra-group hazard rates why the average marginal effect experienced at the covariate mean becomes less telling than the simpler, model-wide pseudo R-squared.

Table A1 in the appendix displays a gross list of all 200+ covariates investigated as per the AME and PR2 approach in determining the one-year model[92] CM_1.

Each of the Composite Model's vector of covariates as per the PR2 method is specified below. These amount to our final DTDDH models.

$$CM_1: \frac{EBIE}{ATA} \leftrightarrow \frac{RE}{TA} \leftrightarrow \frac{BEQ}{TA} \leftrightarrow \frac{IE}{TL} \leftrightarrow AUDNR \leftrightarrow LNA \leftrightarrow LNS \leftrightarrow DPO \leftrightarrow \frac{UTR}{EMP}$$

$$CM_2: \frac{AV}{EMP} \leftrightarrow \frac{IE}{TL} \leftrightarrow LATE \leftrightarrow \frac{CASH}{CA} \leftrightarrow TRDAYS \leftrightarrow \frac{CL}{TA} \leftrightarrow \frac{DIV}{EMP} \leftrightarrow LNAGE \leftrightarrow \frac{EBIE}{S} \leftrightarrow \frac{UTR}{EMP}$$

$$CM_3: \frac{AV}{EMP} \leftrightarrow ln\left(\frac{IE}{TL}\right) \leftrightarrow LATE \leftrightarrow \frac{CASH}{CA} \leftrightarrow TRDAYS \leftrightarrow \frac{CL}{TA} \leftrightarrow DIVD \leftrightarrow LNAGE \leftrightarrow AUDNR \leftrightarrow \frac{UTR}{EMP}$$

$$CM_4: \frac{AV}{EMP} \leftrightarrow ln\left(\frac{IE}{TL}\right) \leftrightarrow LATE \leftrightarrow LNAGE \leftrightarrow LNA \leftrightarrow \frac{CL}{TA} \leftrightarrow DIVD \leftrightarrow INDDEFR \leftrightarrow AUDANY \leftrightarrow \frac{UTR}{EMP}$$

$$CM_5: \frac{AV}{EMP} \leftrightarrow \frac{STD}{EQ} \leftrightarrow LATE \leftrightarrow LNAGE \leftrightarrow LNA \leftrightarrow \frac{CL}{TA} \leftrightarrow INDDEFR \leftrightarrow \frac{TAX}{TA} \leftrightarrow AUDANY \leftrightarrow \frac{UTR}{EMP}$$

Where; $\frac{IE}{TL}$ is the interest expense per total liability[93]; $AUDNR$ is a dummy for the audit remark "not recommended"; LNA is the natural logarithm of assets; LNS is the natural logarithm of sales; DPO are the days payable outstanding; $\frac{UTR}{EMP}$ are the untaxed reserves per employee; $\frac{AV}{EMP}$ is the added value per employee[94]; $LATE$ is a conservative dummy indicator if accounts are submitted late to the Swedish Companies Registration Office[95]; $TRDAYS$ are the days sales outstanding less the days payables outstanding; $\frac{DIV}{EMP}$ are the dividends paid per employee; $LNAGE$ is the natural logarithm of the firm's age since incorporation; $\frac{EBIE}{S}$ are the earnings before interest expenses per revenue ; $DIVD$ is a dummy indicator for dividends paid; $INDDEFR$ is the trailing period's industry default rate; $AUDANY$ is a binary variable for the audit remark "not recommended" or "recommended with notation"; $\frac{STD}{EQ}$ is the short term debt to equity; $\frac{TAX}{TA}$ is the income tax per asset. The reader will recognize outstanding variables not listed from the competing model specifications located in section *3.5: Selected Competing Model Specifications*.

---

[92] For brevity, no other univariate gross-lists are shown for the remaining forecast horizons

[93] Adjusted for financial expenses affecting comparability. Liabilities are adjusted for provisions and deferred tax

[94] Value captures operating profit (adjusted for financial income affecting comparability), plus labour costs and depreciation and amortization

[95] Estimated by examining whether annual statements are received 248 days after the end of the fiscal year. Smoothed to avoid reading manual corrections made by Bisnode as the time-stamp when annual statements were received

Allowing for different model specifications over different time-periods allows us to detect covariates' relative importance to different parts of the survival curve. Altman et al. (2015) and Duan et al. (2018) both highlight this empirically by showing the relative importance to AUROC of their selected covariates over different horizons.

5.3 Summary Statistics of the Composite Models' Covariates

The reader would do well to take a moment and review Tables 11-12. The first summarises the vector sets that produce each of the Composite Models, while presenting their mean development over time. The second examines the one-year look-ahead mean of these across defaults and non-defaults, also over time[96]. The latter shows an evident and persistent difference between the majority and minority class.

5.4 Model Performance Across Horizons

Armed with five Composite Models, five additional competing models and five forecasts, Table 13 presents a summary of our findings for the DTDDH models with random effects (frailty) on the organizational level[97]. The grey quadrant encapsulates the six competing models' one-year predictions, and the bold text therein highlights our own design: the Composite Model.

For conciseness, we streamline our efforts towards the one-year quadrant, its results and interpretations. Thereafter, we revisit Table 13's results for a wider discussion of all forecast periods' results.

5.5 The Composite Model's One-Year Prediction

5.5.1 Sample Estimation and Proportionality

Table 14 shows the model fitting of the one-year CM on the estimation window. All coefficients including the baseline hazard rate are significant on the 1% level. For the specification incorporating a frailty-term by random effects, these too are found to be significant. The AIC[98] and BIC[99] information criteria - as measures of model fit by prediction error - also seem to suggest that for the same number of observations, the CM with contagion is preferred.
The coefficients are also, fortunately, of the *a priori* expected sign save for LNS. $\frac{EBIE}{ATA}, \frac{RE}{TA}, \frac{BEQ}{TA}, LNA, \frac{UTR}{EMP}$ all reduce the propensity for bankruptcy while $\frac{IE}{TL}$, $AUDNR, LNS, DPO$ increase the risk of default. The days payables outstanding covariate, we believe, is to be read as a measure of outgoing, liquidity (cash) in the near-term as opposed to a reduction of the working capital burden to the corporation. For longer forecasts, however, this becomes more dubious – is DPO capturing short-term outflows to claimants or less operating capital requirements?

Untaxed reserves per employee as a significant predictor of default strikes us as curios result: indeed, what corporate fundamentals make this regressor economically meaningful when faced

[96] Once more, for brevity the corresponding tables for the two-to-five year forecasts are not presented. Their covariate constituents, however, are included in the one-year horizon for the reader to get an appreciation of their variance between defaults and non-defaults
[97] For a summary of AUROC for both hazard models (including and excluding random effects), see Table A5 in the appendix
[98] Denotes "Akaika Information Criterion"
[99] Denotes "Bayesian Information Criterion"

with dissolution? The answer, we postulate, lies in Swedish Private Company Law and its application to *independent limited liability companies* specifically. These are allowed tax-allowances on profits made today as an offset for possible losses in the future. For these reasons, untaxed reserves make up an equity and liability (deferred tax) component, why they are reported as a balance sheet line-item between equity and liabilities. Examining the *reported* equity or retained earnings of our population, therefore, without adjusting for equities located in untaxed reserves make for artificially low values.

When making deposits to untaxed reserves, the tax burden for the current period is lessoned, and future losses may be cushioned by past equity 'savings'. This is as the lawmaker intended untaxed reserves to be used - as an apparatus for smoothing earnings over time. There is an additional fiscal motivation, however, to make use of the option to defer tax. Any Swedish company that has done so since the 1980's will likely attest to this. These have, namely, not only delayed their tax incidence, but also lessened its nominal amount as the Swedish corporate tax rate has steadily and consistently declined over time.

There is further reason to believe untaxed reserves (scaled by employees) are important in the default context. Qualitatively, making untaxed reserves is a corporate governance decision made by management. Perhaps, therefore, untaxed reserves signal some degree of executive know-how or represent the workings of a frugal, more risk-averse management e.g. less prone to moral hazard. Or untaxed reserves simply capture management's inclination to postpone tax outflows when positive NPV projects are identifiable.

In any case, company members of a group, however, must split untaxed reserves into its equity and liability line items in their financial statements. For these reasons, we are inclined to suggest that untaxed reserves are economically meaningful for default studies on Swedish limited liability companies[100] on financial and possibly non-financial footings.

Economic worth and significance notwithstanding, only LNA meets the PH assumption (see Table 15). All other antecedents violate proportionality to varying degrees, wherefore the global test too is not upheld. We have, as Allison (2010) would lament, traded some degree of theoretical integrity for model fit.

5.5.2 Model Power and Calibration

Figures 3.1-3.4 show the ROC curves of each of our one-year sample splits: (i) the estimation window (ii) sampled firms out-of-time (iii) hold-out firms in-time, and (iv) hold-out firms out-of-time. For (i)-(iii), the DTDDH CM with random effects outperforms its equal without random effects in terms of AUROC. Note, however, that it does not dominate the former for every cut-off point and the simple logit does a (marginally) better job of discriminating between classes in our harshest test environment for external validity (iv). All AUROCs exceed the 0.8 mark, why the one-year CM can be considered an excellent default classifier (Hosmer Jr. et al., 2013). For all sample subsets, therefore, the CM is able to assign a higher probability of default to defaulting firms than survivors more than 80% of the time.

[100] Or any other company sort in any other geography, for that matter, that enjoys the same accounting privilege

In the precision-recall space (see Figures 3.5-3.8), we see that the hazard model excluding frailty fares better for sample splits (i)-(iii) while the hazard model including contagion performs in (iv). *Ceteris paribus*, one would indicatively expect the DTDDH model with random effects to exhibit better calibration in (iv) and (relatively) worse calibration in (i)-(iii) although such cannot be definitively maintained. All PR curves unilaterally showcase good early-retrieval prospects (i.e. adequate precision for low levels of recall). Finally, the reader is directed to the sampled out-of-time PR plot. Notice that in the far right end of the graph, both DTDDH model variants fall below the proportion of one-year defaults in the population. For very high levels of recall, therefore, you are better off guessing positive cases at the one-year default rate than you are using either of the CMs.

5.6 All Models' One-Year Prediction

Figures 4.1-4.4 express all competing model specifications' one-year discriminatory power by ROC curves. For brevity, only the estimation sample and strictest of hold-out samples are presented for the theoretically preferred logits with random effects. The Composite Model dominates all its rival hazard models, save for Dakovic et al. (2010) in the estimation period where the latter are able to produce an AUROC of 0.7612 relative our 0.7487. Notice, however, that Dakovic et al.'s (2010) covariate vector does not dominate the CM. For low and high levels of false positive rates, the CM is tangent or higher to the former.

In the hold-out out-of-time ROC comparison, the CM starkly outperforms its peers, dominating all other ROC curves with quite some margin. Dakovic et al.'s (2010) model power declines substantially to 0.7435 - below that of both Altman's (1968) Z-score and Altman and Sabato (2007) - while the CM improves its classification ability to 0.8297. The significant deterioration of Dakovic et al.'s (2010) power is likely a consequence of some moment of model overfitting.
Figures 4.5-4.8 present an analogous comparison through PR curves. Herein, the CM reigns supreme in terms of average precision (AUPRC) in both sample cuts and, in expectation, comparative calibration. It is also considerably more skilled in early-retrieval relative its competition.

In summary, the one-year Composite Model with random effects performs well against its rival hazard models. In terms of AUROC, ergo discriminatory power, it dominates all others in the hold-out out-of-time cross-validation test. In terms of AUPRC, ergo average precision, the CM outperforms its contestants across both the estimation window and the hold-out out-of-time sample. *Ex ante*, one would *ceteris paribus* expect the CM to showcase better calibration (empirically confirmed in Table 13), in part from its higher early-retrieval performance.

5.7 All Models' Two-to-Five Year Predictions

Returning our sights to Table 13, the CM is most calibrated (i.e. has the lowest absolute deviation from the population default rate) in the two-year prediction horizon. Only for the five-year forecast does it rank lower than second place. In terms of direct measures of model power (AUROC), the CM is *always* superior to *and* dominates Altman's (1968) Z-score, Zmijewski (1984), Shumway (2001) and Altman and Sabato (2007) regardless of the look-ahead forecast duration or the sample split under consideration (see appendix Table A2 for comparative ROC curve plots). In the three-to-five year horizons, the CM's AUROCs are also consistently higher than that of Dakovic et al. (2010) - significant at the 1% level. For the two-year horizon, the latter outperform the CM's power in the estimation window as well as the hold-out in-time control group - but not the stricter hold-out out-of-time sample.

In the domain of PR curves, the CM's AUPRC never falls below that of a competing model irrespective of prediction horizon, estimation window or cross-validation sample (see appendix Table A2 for comparative PR curve plots). Additionally, the CM is always more apt at early-retrieval, save for one instance: in the five-year sampled firm out-of-time window.

Unfortunately, none of the models meet the global PH test for any prediction period[101] (see Table A3 for tests of proportionality on the one-year horizon).

From the Composite Model's five different specifications (see Table 11) a couple of general observations can be made. Save for the baseline hazard rate, only one predictor is present across all of the CM's: untaxed reserves per employee. Added value per employee, late filing estimates, current liabilities to total assets and the logarithm of age are all made use of on four of five forecast occasions. These alone provide a telling account.

Adding the dividends paid per employee (present in the one- and two-year CMs) to the previous list, we propose that the (literal) common denominator to three of these may be an important scaling factor: employees. Most of the bankruptcy literature has substituted age, sales or assets for size and - potentially - somewhat overlooked the employee. The latter may be a more accurate proxy for size-risk, for instance, for employee-light industries with high sales - e.g. the IT industry.

Added value may well foreshadow bankruptcy risk for the simple reason that it catches gross profitability prior to the distortionary effects of shareholder salaries and dividends on further-down income statement lines. Indeed, for small firms the impact of these can be particularly exaggerated. This ought to be a somewhat generalizable result: for private firms of lesser size, the more gross the margin the less prone it is to volatility induced by its ownership's choice of remuneration.

Collectively, two variants of audit remarks show up four times. As do late filing approximations. This corroborates the well-documented result that qualitative markers are important default predictors, especially in the private and SME space. Notice that for later period forecasts (four- and five-year), the industry default rate assumes a place in the CM. For these predictions, there is no discernible difference between hazard models in ROC- or PR-space with and without random effects *specified on the firm-level* (see Table A2). This synchronous phenomena provides an interesting piece of evidence for the existence of *latent* contagion / frailty at work intra-industry for longer prediction horizons. It appears as though firms behave heterogeneously in the short-term, while industries are more heterogeneous for longer forecasts.

Although the CMs 'self-select' many popular covariates regularly used in the literature (e.g. interest expenses to total liabilities, the logarithm of age and assets), we are particularly sympathetic to the overlap between our CMs and the Dakovic et al. (2010) formulation used in Norway. Notably, we find partial evidence of the empirical relevance of the dividend dummy and

[101] Although many individual covariates do not violate the PH assumption in the multivariate setting, the global 'sum of parts' does. Closest to meeting PH is Zmijewski (1984), whose global test is spoiled by the inclusion of the selected baseline hazard rate (the trailing realized rate of defaults)

the current liabilities to total assets variable purported therein - two less regularly seen bankruptcy predictors.

Finally, we record – to a limited extent – a general transition from antecedents of a shorter nature to those of a longer character the longer the look-ahead horizon. More so, however, it appears as though some predictors typically thought of as long-term - e.g. value added, untaxed reserves and the logarithm of age - are remarkably predictive even in the short-to-mid-term. In the next section, we practically test the CM's one-year empirical prowess detailed in *5.5 The Composite Model's One-Year Prediction* by applying it from the lender's perspective to simulated and actual loan data where banks - each using a separate hazard model - compete for customers. The best performing bank makes use of the model formulation that most optimally trades-off power, calibration and the costs of misclassification in quoting a competitive and (hopefully) profit-making interest rate to the market.

5.8 Economic Value of Default Misclassification

The results from the above segments seem to, generally, demonstrate the superiority of the Composite Model in the AUROC and AUPRC space. Evaluating models using these methods assumes equal misclassification costs, however. Table 16 and Table 17 present the economic results of a simulated competitive loan market (i.e. the *simulated approach*) in which banks apply the one-year CM and its five contending models for quoting credit spreads. Revenues, market share, profits, ROA and RORWA along with other metrics are shown. These results consider that the cost of granting credit to a firm that consequently fails (a false negative) may be significantly higher than forgoing interest income. Since the Composite Model fared well in the out-of-time samples vis-à-vis the other models' AUROC and AUPRC, we examine the same time period (2014-2016) in this section to highlight any potential weaknesses of the CM model. This simultaneously evaluates all models on out-of-time data not before seen. All analysis regarding the economic value of misclassification is conducted using one-year predictions.

As the analysis has shifted to the perspective of banks and their economics, the "sampled firm, out-of-time" sample can be nicely interpreted as "same clients, another time period". That is, how well do the banks perform against each other when competing for credit market share raised by the same companies as their respective models were specified on? The same logic for the "hold-out firm, out-of-time" split of the data tells us we can interpret the banks' performance on these as the models' ability to attract new firms, i.e. companies different from those the banks' models were fit on. Since previous research has made no distinction between different time-periods and hold-out samples when applying the simulated approach, this is the first time evaluation of default models has been separated in this economically meaningful and intuitive way[102].

Tables 16-17 show that Dakovic et al. (2010) outperform the other models in terms of market share, gaining almost 70% of the granted credits every year both for existing and new clients. Shumway (2001), followed by the Composite Model, grab approximately 22% and 7% respectively for both existing and new clients. Dakovic et al. (2010) manage to take a considerable share of granted credits while assuming a (relatively speaking) smaller share of defaults: 47% and 51% for existing and new clients respectively. Shumway (2001), on the other hand, gained

[102] Note that each year resets the competition for clients, such that many of the clients for any specific bank can be a new clients for that year, even if the evaluation is made on sampled firms in-time

its market share of credits at the cost of shouldering a high share of defaulters: 43% and 40% for existing and new clients. The bank using the Composite Model, while not gaining an especially high market share, very effectively avoided losses, evident from its ca 2% share of defaulters. The Composite Model can be compared with the well-established Altman (1968) model, which was burdened by a slightly higher share of defaulters (2.0% for current clients and 4.7% for new clients), but was only granted 0.3% and 1.2% of the credits for existing and new clients during for the same period. Note that some differences in the total count of loans granted vs. the actual number of loans granted can arise if all banks overlap in their exclusion of companies they deem to be among the least (5%) creditworthy for a given year.

Spreads (interest rates) charged to the clients in combination with credit loss events result in revenues, losses and profits for the banks. Dividing profits with loan sizes, return on assets is examined. Ranging from 1.4% for Shumway (2001) to 1.9% for Altman (1968) and Zmijewski (1984) for existing clients, and from 1.4% for Shumway (2001) to 2.3% for Zmijewski (1984) for new clients, there seems to be limited read-across between their aforementioned share of credits / losses and ROA metrics. In a competitive marketplace, ROA should vary primarily with the risk taken (i.e. increase with higher risk and higher interest rates and decrease with lower risk and lower interest rates). RORWA, on the other hand, incorporates the risk of credit loss by reducing the asset value for safe assets (increasing returns when lowering risk, and decreasing returns when increasing risk), and is consequently a more appropriate total return measure. As can be seen, the Composite Model delivers the highest RORWA in all years, ranging from 1.07 (1.06) to 3.53 (3.39) times higher than its competing models for existing (new) clients.

Before understanding *why* RORWA is higher for the Composite Model vs. its competition, some recap of previous discussions on calibration and separability might be valuable for the reader. A bank can have an excellent internal model for differentiating between prospective clients' probabilities of default (captured by the AUROC and AUPRC), but if it quotes spreads to the market that are, say, 70% off-the-market (captured by principally the AUPRC), clients will choose another bank since interest rates are too high / miscalibrated. Conversely, a bank might not be able to differentiate between the credit risks of clients at all, but quoting spreads at exactly the market-average interest rate will undoubtedly allow it to sell its services to both defaulters and survivors. In a competitive setting with many banks, both separability and calibration clearly matter, and the interactions are highly complex the more competing banks there are. Previous research along the lines of Agarwal and Taffler (2008a) has handled this issue by purposely rinsing away any and all differences due to miscalibration through a process of logit transformations, bankruptcy probability percentiles, and ROC smoothing. We, on the other hand, are interested in the miscalibration produced by the different hazard models, allowing for evaluation and discussion of the joint-impact of separability and miscalibration on bank profitability.

We gauge the potential impact of these two model dimensions with reference to Table 18, which presents another version of the one-year prediction table from the summary table of results, adjusted such that it only displays the *out-of-time* (2014-2016) calibration and model power metrics (AUROC and AUPRC). The degree of calibration is operationalized as the absolute value of the delta (Δ) found in Table 18, while also noting that the precision-recall curve contains information both for power *and* calibration. Since Altman (1968) and Zmijewski (1984) are in fact equally (mis)calibrated as the Composite Model, it can be deduced from the differences in

AUROC that one of the primary reasons for outperforming Altman (1968) and Zmijewski (1984) in terms of both profits and RORWA is due to higher model power, ergo separability. The Composite Model's much higher RORWA vis-à-vis Shumway (2001) can partly be attributed to poor calibration of the latter model. High miscalibration is causing Shumway's (2001) model to estimate probabilities of default (and by extension credit spreads) at an arbitrarily low rate relative its competition[103]. Consequently, we observe - as one would expect - higher market shares of granted credits and a higher share of defaulters for Shumway's (2001) model.

In light of the CM's higher RORWA vs. Dakovic et al. (2010), strong forces pertaining to both power and calibration seem to be at work. That is, the Composite Model's RORWA-performance for existing clients (ca 13%) and new customers (ca 12%) stems from both superior calibration and superior classification. The CM being particularly more capable than its competitors in PR space (AUPRC and early-retrieval performance) makes its materially higher RORWA especially interesting. For example, Dakovic et al.'s (2010) relatively low AUPRCs of 1.7% and 1.5% for existing and new clients are markedly lower than those of the Composite Model's 2.5% and 3.2%. This might be why Dakovic et al.'s (2010) returns are lower than the Composite Model: identifying a large number of defaulters relative its default predictions (precision) should allow for better earnings power. On the other hand, Altman (1968) and Altman and Sabato (2007) outperform Dakovic et al. (2010) in terms of AUPRC and AUROC for new clients, but report a lower RORWA. Perhaps this is due to the complex relationship between power and calibration when multiple models compete, exacerbated by a PD-floor of 5 bps. Clearly, there is ample information in both ROC and PR curves that translate into measurable economic returns, but whether these are exhaustive as model dimensions determining lender performance remains uncertain.

Note that while the difference in RORWA between models is indicative of the differences in economic value that can be created, the absolute *level* of the RORWA is not fundamentally interesting in this exercise. Decreasing the loss given default, for example, would increase the RORWA for all models[104].

Thus far, it should be evident that the Composite Model (i) generally outperforms all competing models when assuming equal misclassification costs (AUROC and AUPRC space), and (ii) produces higher RORWAs than other banks for both existing and new clients under the F-IRB *simulated approach*. We have argued that there are several benefits to using notional credit exposures in place of a simple simulation when the aim of model evaluation is to re-create a realistic test-setting. In-so-doing, potential biases and unrealistic assumptions can be avoided.

The notional exposure method is by design not a simulation. Banks are only assumed to be able to earn interest and experience credit losses on loans that have actually been granted and assumed in the Swedish private corporate loan market. If the purpose is to evaluate the value of (mis)classification from the lender's perspective, this ought to be the researcher's preferred route: less than half of both firm-year observations and one-year defaults occur for companies that actually have bank loans on their balance sheets. Defaults in the private space also often arise

103 Shifting the range of credit spreads downwards also increases the probability of estimating probabilities of default below the Basel III floor of 5 bps. Some capacity of the model to separate firms' probability of default is therefore also lost, causing inefficient pricing and a lower RORWA

104 Bauer and Agarwal's (2014) ROAs and RORWAs range from between one-tenth and one-fifth of the nominal values reported in this paper, for instance. On the other hand, the ratios between ROA and RORWA are approximately similar

from a combination of poor earnings power and leverage, as well as shifting power dynamics in the value-chain.

In Tables 19 and 20, the results of the economic value of default (mis)classification using the notional credit exposure method are shown. Many of the trends apparent in the simulated approach can be identified here too. In terms of market share, for example, Dakovic et al. (2010) and Shumway (2001) dominate in both new and existing client segments, reaping large revenues and profits, while Altman (1968), Zmijewski (1984) and Altman and Sabato (2007) retain just 0.2% and 0.8% market share. Interestingly, Altman and Sabato (2007) run at a loss when lending to new clients as the two defaults generate *very* large credit losses in relation to revenues. Good calibration and model power for the CM were already exhibited in the simulated approach through its low count of experienced default events: 16 and 10 for existing and new clients respectively. Using notional credit exposures, the CM assumes just 1 default for existing customers - none for new clients - out of a combined 5,934 granted credits, which is rather remarkable. One final observation is that differences in RORWA between models grow larger as notional credit exposures are used, when comparing to the simulated approach.

Prior to understanding why the difference in RORWA becomes relatively larger using notional credits, a quick comment on how to interpret aggregate numbers might be of value to the reader. The right-most column in the Table 19-20 shows the aggregated interest income, i.e. the sum of all interest income on all the income statements for the relevant companies. To clarify, this is *not* the sum of charged interest as estimated through the default models and their internally estimated probabilities of default and credit spread formulae, but rather a summary of actual external interest expenses contained in the profit and loss statements for the companies included in the sample split. Clearly, for both existing and new clients, this aggregate number (SEK 6.41bn / SEK 3.95bn for new / existing clients) is larger than the total estimated (predicted) interest charged by all the banks in the competitive experimental design. Since the assets and interest expense are known quantities from the data, we can comfortably claim that the level of ROA is in the 5.2%-5.5% region for credit institutions lending to the private firms in the population[105]. Whether or not these are large banks or other types of credit institutions, and whether or not they use the F-IRB approach is impossible for us to tell. Consequently, it is not possible to estimate a RORWA directly from these interest expenses and asset values.

Returning to *why* differences in RORWAs become larger when using notional credit exposures rather than simulated values (see Figure 5), we wish to highlight several possible explanations. Shumway (2001), Dakovic et al. (2010) and the Composite Model have clearly seen the highest percentage-point increases in RORWA. Some proportion of this improvement could in fact be random: as the number of evaluated credits granted and defaults decreases, the likelihood of more outliers in terms of both revenues and losses increases. While this is a possible account of what has transpired, the 65,592 and 43,228 granted credits for existing and new customers dwarf the *entire* sample sizes of notable papers such as Bauer and Agarwal (2014) who have but 28,804 total firm-years. Random noise, in this respect, is an unlikely culprit. A more likely rendition is that the models are fit on *actual* training data, wherefore they are better positioned to show their prowess when applied in a realistic, actual context. Simulations smoothen-out models' divergent abilities in capturing the marketplace's idiosyncrasies and the causes of firm default, why differences in model specifications become more pronounced when applied to actual data

[105] This range does not vary substantially if the observer keeps to opening balance sheets

instead. While simplifying assumptions were helpful in the simulated approach, they hinder practical model evaluation in real-world settings demarcated by notional credit exposures.

Coming back to the Table 19, notice how the CM earns SEK 75k per loan[106] (almost double that of Shumway (2001), the second best bank ranked by the same metric). All while the CM loses just SEK 33.6k for the single defaulter it assumed. While loss per default cannot be evaluated for the new customer segment (since there were none), the revenues per granted loan of SEK 58.3k on new firms are also significantly higher than the corresponding amounts for other banks.

Revenues per granted credit and loss per default are credit ratios that have not been reported in previous research, for the simple reason that all loans are of the same size under the *simulated approach*. With reference to the simulated approach's results displayed in Table 16-17, losses per loan are constant, and the revenues per loan are quite stable across models. Perhaps of greater interest, using simulations the Composite Model's revenues per loan are significantly lower than the average bank. From this, we draw the conclusion that compared to its competing banks, the CM can quote competitive spreads to primarily large corporations that are able to deliver on larger financial obligations in part due to their size, trade-credit position and untaxed reserves. Note that Shumway (2001) and Dakovic et al. (2010) also make use of predictors relating to size and/or age. These models substantially improve their RORWA when applying notional credit exposures, a pattern across models which arguably highlights the removal of a potential size bias when using notional credit exposures instead of equally weighted, constant loan sizes. Again, Zmijewski (1984), and especially Altman (1968) and Altman and Sabato (2007) experience unchanged, even deteriorating performance *without* including any predictive covariates capturing company size.

The latter authors also (i) include covariates after financial expenses (e.g. net income to total assets), (ii) focus on debt (short-term debt to equity, and interest coverage ratios) and (iii) do not separate between trade working capital and all working capital (current assets and current liabilities, which include financial items). It is possible these models fail to capture the reasons companies fail (in the short term) that are unrelated to financial indebtedness. The difference in RORWA using notional credit exposures and the simulated approach arguably highlights these intricacies.

As a final comment on the relative performance of the different models, it is worth noting how well Altman's (1968) Z-score stacks up against its competition when offering loans to new clients. Fashioned for public manufacturing firms in the U.S. more than 50 years ago, it fares better than the more recent Altman and Sabato (2007) specification sculpted with U.S. SMEs in mind. As such, Altman (1968) continues to be highly empirically influential to the default prediction literature. The success of the Z-score across private and public markets in different geographies, industries and time-periods using a parsimonious vector of predictors is something we encourage and hope to see more of.

6. Limitations of Research

6.1. Variable Selection

[106] Which is also in the whereabouts of actual revenues per loan, see the right-most column in Tables x-y

While the number of covariates tested in this paper - to the best of our knowledge - constitutes the most exhaustive examination of regressors in the literature of default prediction, there are several interesting areas outstanding which we have been unable to properly delve into and explore.

One of these includes innovation. While a comprehensive list of all patent applications by Swedish corporations have been included in the analysis, other and better measures of innovation could have been incorporated. Whether analysing the proportions of patents granted, the relative number of patents vs. competitors, or patents per employee will prove economically and statistically meaningful, we do not know. Regardless of the list of innovation-covariates left unexamined, incorporating just two variables based on patent application count, we believe, does not provide a representative view of the relationship between default and inventiveness.

Another set of variables of which we know carries significant predictive power relates to payment behaviour. We expect significant model improvements could be achieved should (i) delayed firm payments on unsecured debt (similar to the county court judgement in the UK), and (ii) the firm's board members private payment history later be made available.

Auditor changes were included in this study both as a dummy for the current-period change and as a continuous variable capturing the cumulative number of times the firm has changed its appointed auditor. We had no preconceived notion of the direction said covariates would impact the default event. While neither variable proved insightful, perhaps a more economically meaningful approach would have been to exclude the legally required auditor rotations taking place every fourth year (Bolagsverket, 2016) from these variables.

With regards to variables that were in fact included in the Composite Model, are there any limitations which could have impacted model performance (positively or negatively)? The only one, as far as we can see, is the late filing predictor. While the dummy in question proved to be statistically significant in this study across several forecast horizons, it does have the caveat of being an approximation. The variable is computed by comparing the date of the underlying financial statement (plus the allowed time before the statement is due) to the time-stamp of the most recent change recorded for said financial statement. As explained by correspondence with Bisnode (the proprietor of the Serrano database), financial accounts' time of last change are not necessarily the same dates and times that the statements were received by the Swedish Company Registration Office: some later, manual corrections to the data have been made by Bisnode themselves for a portion of the companies. Although the data has been re-adjusted to correct for this on dates with a strikingly high number of coincident last-change-time-stamps (i.e. Bisnode modifications), the late filing dummy remains an imperfect estimate. Consequently, the explanatory power of the binary indicator could be revised and likely improved if a late filing register of sorts could be retrieved directly from the Swedish Company Registration Office instead.

A final point worth mentioning with respect to the variables made use of is the methodology in which they are chosen. While we recognize our implemented variable selection procedures are econometrically defendable, they are highly atheoretical in nature. If a unifying theory existed, and there were empirical consensus on predictors of private default, it is entirely possible other variable selection processes would have been carried out.

6.2 Economic Cost of (Mis)classification

While we certainly see several advantages of incorporating firm-year specific notional credit exposures into default model performance evaluation, balance sheet data used herein still fails to incorporate details with regards to the credit institution or institutions that granted the loans(s). Consequently, the loan balance could constitute several loans from several institutions, provided at different fixed and / or floating interest rates. Because of this, it is impossible to estimate the price sensitivity of firms or, put differently, what sub-par interest rates they are willing to accept before changing lender. Such important dynamics could increase the realism of the notional credit exposure method by incorporating an element of client 'stickiness'. While applying minimum required differences in quoted spreads for firms to change banks is a step in the right direction in this regard, such a solution would fail to incorporate the extent of client-firm relationships across banks. That is, if an *equal* minimum difference in quoted spread towards clients, i.e. stickiness, is assumed across banks, the expected value of bank profits would remain unchanged (since all banks customers are equally sticky), and the increased model realism would not carry any value.

Furthermore, if individual loan data similar to the data used by Duan et al. (2018) would be incorporated into a large dataset such as Serrano, the calibration and power of banks' internal bankruptcy prediction models could be effectively estimated by examining their share of credits, share of defaulters, and loan pricing. Theoretical models could then be tested against actual models, at least in terms of market share, default shares and ROA. RORWA, on the other hand, would require admission to internally estimated probabilities of default.

Another monition to the model evaluation and performance of banks is that the actual loss given default is assumed constant. This is true for both the simulated and notional credit exposure approaches. Certain models should logically be able to outperform others in estimating a bank's ability to recover value in the event of default. In fact, if only one model aims to capture the highest possible NPV of a loan, it is entirely possible that while it underperforms in terms of AUROC and AUPRC, its losses in event of a default are low enough to compensate for lower model power.

7. Conclusion

The literature on corporate bankruptcies spans the better part of eight decades of research (Bellovary et al., 2007). In spite of this, there exists no unifying microeconomic theory nor empirical consensus on the predictors of bankruptcy. While most have focused their efforts towards the public market for shorter predictive horizons, we acknowledge the deficit of private company attention and longer forecasting periods. Even so, the private default body of research is becoming increasingly un-parsimonious and un-private through complex model building and private-public read-across (e.g. Dakovic et al., 2010; Duan et al., 2018). We mean-revert these two recent trends by comprehensively reviewing an information set of more than 200 unique financial and non-financial private default predictors in search of a parsimonious specification for detecting defaults in the population of Swedish independent limited liability companies outside of the financial and real-estate industries. One- to five-year predictions ahead-in-time are made between 1998-2017 for an aggregate of 55,411 bankruptcy events corresponding to 2.6 million firm-year observations and 245,844 unique companies. Two atheoretical methods for model variable selection inspired by Hosmer Jr. et al.'s (2013) *average marginal effects* empirical approach are designed and implemented. The selection procedure most favoured of the two relies on ranking covariates by their univariate pseudo R-squared using survival analysis. The resulting covariate vectors make up the multivariate *Composite Models* and are modelled as discrete-time duration-dependent hazard models with and without random effects. Five hazard models of the same kind are introduced, each with its own set of extant covariates: Altman's (1968) Z-score, Zmijewski (1984), Shumway (2001), Altman and Sabato (2007) and Dakovic et al. (2010). Together, these make up rival model formulations from different epochs, geographies and markets for the Composite Models to compete against.

Stand alone, the Composite Model's one-year prediction achieves excellent discriminatory ability as a classifier in the most demanding of cross-validation tests: the *hold-out-firms out-of-time* sample. Comparatively, the Composite Models' power for separability between defaults and non-defaults dominates all other models across all hold-out tests and forecast horizons (significant at the 1% level), save for Dakovic et al. (2010) on occasion. The Composite Model's average precision never falls below that of a competing model, and its early-retrieval performance is consistently higher than its peers except for one instance. Model calibration is also highest for one-year and two-year predictions; second highest for the three- and four-year horizons. In practically assessing the Composite Model's merits, we backtest all six models' one-year hold-out forecasts on simulated and actual loan data for the period 2014-2016 from the lender's perspective using the Basel III F-IRB approach. Under the simulated (equally-weighted) loan market, banks using the Composite Model capture a lacklustre market share, but significantly outperform all other bankruptcy-prediction models on RORWA. When replaced with actual market data (i.e. nominal credit exposures), the Composite Model's relative RORWA-outperformance increases further yet.

Our contributions to the corporate bankruptcy literature are multiple-fold. First and foremost, to the best of our knowledge this study constitutes the most comprehensive chronicle of reviewed default predictors and is second only to Altman et al. (2012) in the count of defaults under study in the private market. On aggregate, we are able to outperform other well-regarded model formulations across dimensions of power, calibration, early-retrieval performance and RORWA, suggesting there is still much work to be done in locating economically and empirically important antecedents to bankruptcy. On the latter topic, we make four additions. We find some evidence of latent contagion (frailty) operating on the intra-industry level for longer predictions. The number of employees, we posit, is a relevant denomination of size for industries where sales, assets and / or age may be misleading. Thirdly, possibly unique to the Swedish legal environment surrounding independent limited liability firms, we discover that untaxed reserves are highly foretelling in private default prediction. Several possible explanations for this are offered and discussed. Fourthly, through a variable capturing earnings before labour costs, we suggest that the more 'gross' an income statement line is, the less prone it is to distortion from the salary and dividend remuneration preferences of its ownership. We also make some minor methodological improvements to the literature, including adjustments for sampling bias (Skogsvik and Skogsvik, 2013), global tests for proportional hazards and the use of precision-recall (PR) curves as complements to receiver operating characteristic (ROC) curves in measuring model power and calibration sensitized to class-imbalance. In our lender application, by subjecting our various models to nominal historical credit risk exposures we show that the results obtained from a simple yet naive simulation need not coincide with the former: a model skilled in identifying size-weighted risk may be unfairly penalized in an equally-weighted simulation. Finally, combining hold-out and sampled firms when backtesting prediction models allows for a simple yet intuitive interpretation of model performance within new and existing customer segments.

8. Cited References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance*, *32*(8), 1541–1551. Retrieved from https://econpapers.repec.org/RePEc:eee:jbfina:v:32:y:2008:i:8:p:1541-1551

Agarwal, V., & Taffler, R. J. (2007). Twenty-five years of the Taffler z-score model: Does it really have predictive ability? *Accounting and Business Research*, *37*(4), 285–300. https://doi.org/10.1080/00014788.2007.9663313

Alley,Richard B. (2019). No Title. *Science*. Retrieved from https://science.sciencemag.org/content/sci/363/6425/342.full.pdf

Allison, P. D. (2010). *Survival Analysis Using SAS: A Practical Guide, Second Edition* (2nd ed.). SAS Publishing.

Altman, E., Iwanicz-Drozdowska, M., Laitinen, E., & Suvas, A. (2016). Financial and nonfinancial variables as long-horizon predictors of bankruptcy. *The Journal of Credit Risk*, *12*, 49–78. https://doi.org/10.21314/JCR.2016.216

Altman, E., Sabato, G., & Wilson, N. (2012). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk*, *6*.

Altman, E I. (1968). FINANCIAL RATIOS, DISCRIMINANT ANALYSIS AND THE PREDICTION OF CORPORATE BANKRUPTCY. *The Journal of Finance*, *23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, Edward I, & Sabato, G. (2007). Modelling Credit Risk for SMEs: Evidence from the U.S. Market. *Abacus*, *43*(3), 332–357. https://doi.org/10.1111/j.1467-6281.2007.00234.x

Andrikopoulos, P., & Khorasgani, A. (2018). Predicting unlisted SMEs' default: Incorporating market information on accounting-based models for improved accuracy. *British Accounting Review*, *50*(5), 559–573. https://doi.org/10.1016/j.bar.2018.02.003

Audretsch, D. (1995). Innovation, growth and survival. *International Journal of Industrial Organization*, *13*(4), 441–457. Retrieved from https://econpapers.repec.org/RePEc:eee:indorg:v:13:y:1995:i:4:p:441-457

Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance*, *40*(C), 432–442. Retrieved from https://econpapers.repec.org/RePEc:eee:jbfina:v:40:y:2014:i:c:p:432-442

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, *4*, 71–111. https://doi.org/10.2307/2490171

Beck, N., Katz, J. N., & Tucker, R. (1998). Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science*, *42*(4), 1260–1288. https://doi.org/10.2307/2991857

Bellovary, J., Giacomino, D., & Akers, M. (2006). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Accounting Faculty Research and Publications*, *33*.

Bisnode. (2015). *No Title*. Retrieved from https://data.houseoffinance.se/serrano/index

Blochlinger, A., & Leippold, M. (2006). Economic benefit of powerful credit scoring. *Journal of Banking & Finance*, *30*(3), 851–873. Retrieved from https://econpapers.repec.org/RePEc:eee:jbfina:v:30:y:2006:i:3:p:851-873

Board, F. S. (2019). No Title. Retrieved December 1, 2019, from https://www.fsb.org/work-of-the-fsb/implementation-monitoring/monitoring-of-priority-areas/basel-iii/

Bolagsverket. (2013). No Title. Retrieved from https://bolagsverket.se/en/bus/business/limited/2.1147/auditor-limited-companies-1.8643

Bolagsverket. (2018). No Title. Retrieved December 9, 2019, from https://bolagsverket.se/en/bus/business/limited/annual-reports/annual-reports-and-financial-years-limited-companies-1.8677

Breslow, N. E. (1975). Analysis of Survival Data under the Proportional Hazards Model. *International Statistical Review / Revue Internationale de Statistique*, *43*(1), 45–57. https://doi.org/10.2307/1402659

Buddelmeyer, H., Jensen, P. H., & Webster, E. (2009). Innovation and the determinants of company survival. *Oxford Economic Papers*, *62*(2), 261–285. https://doi.org/10.1093/oep/gpp012

Cefis, E., & Marsili, O. (2005). A matter of life and death: innovation and firm survival. *Industrial and Corporate Change*, *14*(6), 1167–1192. Retrieved from https://econpapers.repec.org/RePEc:oup:indcch:v:14:y:2005:i:6:p:1167-1192

Chava, S., & Jarrow, R. (2004). Bankruptcy Prediction With Industry Effects. *Review of Finance*, *8*. https://doi.org/10.1093/rof/8.4.537

Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. *Journal of Business Research*, *12*. https://doi.org/10.1016/j.jbusres.2014.10.003

Ciampi, F., & Gordini, N. (2013). Small Enterprise Default Prediction Modeling Through Artificial Neural Networks: An Empirical Analysis of Italian Small Enterprises. *Journal of Small Business Management*. https://doi.org/10.1111/j.1540-627X.2012.00376.x

Com, U. . S. and E. C. (n.d.). No Title. Retrieved from https://www.sec.gov/smallbusiness/goingpublic/exchangeactreporting

Daily, C. M., & Dalton, D. R. (1994). Bankruptcy and Corporate Governance: The Impact of Board Composition and Structure. *The Academy of Management Journal*, *37*(6), 1603–1617. https://doi.org/10.2307/256801

Dakovic, R., Czado, C., & Berg, D. (2010). Bankruptcy prediction in Norway: a comparison study. *Applied Economics Letters*, *17*, 1739–1746. https://doi.org/10.1080/13504850903299594

Das, S. R., Duffie, D., Kapadia, N., & Saita, L. (2007). Common failings: How corporate defaults are correlated. *Journal of Finance*, *62*(1), 93–117. https://doi.org/10.1111/j.1540-6261.2007.01202.x

Duan, J.-C., Kim, B., Kim, W., & Shin, D. (2018). Default Probabilities of Privately Held Firms. *Journal of Banking & Finance*, *94*. https://doi.org/10.1016/j.jbankfin.2018.08.006

DUFFIE, D., ECKNER, A., HOREL, G., & SAITA, L. (2009). Frailty Correlated Default. *The Journal of Finance*, *64*(5), 2089–2123. https://doi.org/10.1111/j.1540-6261.2009.01495.x

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, *83*(3), 635–665. https://doi.org/https://doi.org/10.1016/j.jfineco.2005.10.011

Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *Journal of Financial and Quantitative Analysis*, *7*(2), 1477–1493. Retrieved from https://econpapers.repec.org/RePEc:cup:jfinqa:v:7:y:1972:i:02:p:1477-1493_01

Engelmann, B., Hayden, E., & Tasche, D. (2003). Testing Rating Accuracy. *Risk*, *16*, 82–86.

Flach, P., & Kull, M. (2015). Precision-Recall-Gain Curves: PR Analysis Done Right. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 28* (pp. 838–846). Retrieved from http://papers.nips.cc/paper/5867-precision-recall-gain-curves-pr-analysis-done-right.pdf

Gepp, A., & Kumar, K. (2008). The Role of Survival Analysis in Financial Distress Prediction. *Business Papers*, *16*.

Gruszczyński, M. (2019). On Unbalanced Sampling in Bankruptcy Prediction. *International Journal of Financial Studies* , Vol. 7. https://doi.org/10.3390/ijfs7020028

Gupta, J., Gregoriou, A., & Ebrahimi, T. (2017). Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance*, *18*, 1–30. https://doi.org/10.1080/14697688.2017.1307514

Hillegeist, S., Keating, E., Cram, D., & Lundstedt, K. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, *9*, 5–34. https://doi.org/10.1023/B%3ARAST.0000013627.90884.b7

J. Acs, Z., & Audretsch, D. B. (1989). Patents as a Measure of Innovative Activity. *Kyklos*, *42*(2), 171–180. https://doi.org/10.1111/j.1467-6435.1989.tb00186.x

Jensen, P., Webster, E., & BUDDELMEYER, H. (2008). Innovation, Technological Conditions and New Firm Survival. *The Economic Record*, *84*, 434–448. https://doi.org/10.1111/j.1475-4932.2008.00509.x

Jo, H., Han, L., Lee., H. (1997) Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications .*, 13, 97-108. //doi.org/10.1016/S0957-4174(97)00011-0

Jr, D., Lemeshow, S., & Sturdivant, R. (2013). *The Multiple Logistic Regression Model*. https://doi.org/10.1002/9781118548387.ch2

Keasey, K., Pindado, J., & Rodrigues, L. (2014). The determinants of the costs of financial distress in SMEs. *International Small Business Journal*, 1–20.

Lando, D., & Nielsen, M. S. (2010). Correlation in corporate defaults: Contagion or conditional independence? *Journal of Financial Intermediation*, *19*(3), 355–372. Retrieved from https://econpapers.repec.org/RePEc:eee:jfinin:v:19:y:2010:i:3:p:355-372

Lang, L. H. P., & Stulz, R. (1992). Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis. *Journal of Financial Economics*, *32*(1), 45–60.

https://doi.org/https://doi.org/10.1016/0304-405X(92)90024-R

Lesaffre, E., & Spiessens, B. (2001). On the Effect of the Number of Quadrature Points in a Logistic Random-Effects Model: An Example. *Journal of the Royal Statistical Society Series C*, *50*, 325–335. https://doi.org/10.1111/1467-9876.00237

Merton, R. C. (1974). ON THE PRICING OF CORPORATE DEBT: THE RISK STRUCTURE OF INTEREST RATES*. *The Journal of Finance*, *29*(2), 449–470. https://doi.org/10.1111/j.1540-6261.1974.tb03058.x

Nam, C., Kim, T., Park, N., & Lee, H. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting*, *27*, 493–506. https://doi.org/10.1002/for.985

Ogundimu, E. O. (2019). Prediction of default probability by using statistical models for rare events. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(4), 1143–1162. https://doi.org/10.1111/rssa.12467

Ohlson, J. (1980). FINANCIAL RATIOS AND THE PROBABILISTIC PREDICTION OF BANKRUPTCY. *Journal of Accounting Research*, *18*(1), 109–131. Retrieved from https://econpapers.repec.org/RePEc:bla:joares:v:18:y:1980:i:1:p:109-131

Oommen, T., Baise, L. G., & Vogel, R. M. (2011). Sampling Bias and Class Imbalance in Maximum-likelihood Logistic Regression. *Mathematical Geosciences*, *43*(1), 99–120. https://doi.org/10.1007/s11004-010-9311-8

Pederzoli, C., & Torricelli, C. (2010). A parsimonious default prediction model for Italian SMEs. *Universita Di Modena e Reggio Emilia, Facoltà Di Economia "Marco Biagi", Centro Studi Di Banca e Finanza (CEFIN) (Center for Studies in Banking and Finance)*, *5*.

Peter Miller. (2013). No Title. Retrieved from https://www.nationalgeographic.com/news/2013/3/130307-weather-snowstorm-wrong-forecast-meteorology-world-europe-science/

Reisz, A., & Perlich, C. (2007). A Market-Based Framework for Bankruptcy Prediction. *Journal of Financial Stability*, *3*, 85–131. https://doi.org/10.1016/j.jfs.2007.02.001

Ribeiro, B., Silva, C., Chen, N., Vieira, A., & das Neves, J. (2012). Enhanced Default Risk Models with SVM+. *Expert Syst. Appl.*, *39*(11), 10140–10152. https://doi.org/10.1016/j.eswa.2012.02.142

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PloS One*, *10*, e0118432. https://doi.org/10.1371/journal.pone.0118432

Sasieni, P. (1992). Generalized additive models. T. J. Hastie and R. J. Tibshirani, Chapman and Hall, London, 1990. No. of Pages: xv + 335. Price: £25. ISBN: 0-412-34390-8. *Statistics in Medicine*, *11*(7), 981–982. https://doi.org/doi:10.1002/sim.4780110717

Senteney, D., Bazaz, M., & Ahmadpour, A. (2006). Tests of the incremental explanatory power of auditor qualified opinion and audit firm changes in predicting impending bankruptcy. *International Journal of Accounting, Auditing and Performance Evaluation*, *3*, 434–451. https://doi.org/10.1504/IJAAPE.2006.011205

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The*

Journal of Business, *74*(1), 101–124. Retrieved from https://econpapers.repec.org/RePEc:ucp:jnlbus:v:74:y:2001:i:1:p:101-24

Skogsvik, K., & Skogsvik, S. (2013). On the choice based sample bias in probabilistic bankruptcy Prediction. *Investment Management and Financial Innovations*, *10*, 29–37.

Sofaer, H. R., Hoeting, J. A., & Jarnevich, C. S. (2019). The area under the precision-recall curve as a performance metric for rare binary events. *Methods in Ecology and Evolution*, *10*(4), 565–577. https://doi.org/10.1111/2041-210X.13140

Stein, R. (2007). Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation. *Journal of Risk Model Validation*, *1*. https://doi.org/10.21314/JRMV.2007.002

Stein, R. M. (2005). The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing. *Journal of Banking & Finance*, *29*(5), 1213–1236. Retrieved from https://econpapers.repec.org/RePEc:eee:jbfina:v:29:y:2005:i:5:p:1213-1236

Taffler, R. J. (1983). The Assessment of Company Solvency and Performance Using a Statistical Model. *Accounting and Business Research*, *13*(52), 295–308. https://doi.org/10.1080/00014788.1983.9729767

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, *112*. https://doi.org/10.1016/j.dss.2018.06.011

Zmijewski, M. (1984). METHODOLOGICAL ISSUES RELATED TO THE ESTIMATION OF FINANCIAL DISTRESS PREDICTION MODELS. *Journal of Accounting Research*, *22*, 59–82. Retrieved from https://econpapers.repec.org/RePEc:bla:joares:v:22:y:1984:i::p:59-82

Table 5: Industry Default by Year, 1998-2017

	Energy & Environment			Materials			Industrial Goods			Construction Industry			Shopping Goods			Convenience Goods		
Year	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate
1998	0	566	0,0%	0	1865	0,0%	3	12164	0,0%	2	13620	0,0%	5	26688	0,0%	0	5458	0,0%
1999	0	586	0,0%	0	1 988	0,0%	40	12 853	0,3%	21	14 345	0,1%	81	28 483	0,3%	19	5 893	0,3%
2000	1	602	0,2%	23	2 047	1,1%	126	13 154	0,9%	136	14 972	0,9%	374	29 638	1,2%	131	6 120	2,1%
2001	2	617	0,3%	37	2 120	1,7%	201	13 303	1,5%	209	15 577	1,3%	636	30 463	2,0%	183	6 305	2,8%
2002	4	647	0,6%	38	2 171	1,7%	257	13 431	1,9%	309	16 197	1,9%	829	31 471	2,6%	242	6 406	3,6%
2003	9	595	1,5%	40	2 251	1,7%	355	13 526	2,6%	408	16 583	2,4%	892	32 146	2,7%	289	6 511	4,3%
2004	6	609	1,0%	46	2 296	2,0%	321	13 601	2,3%	407	17 137	2,3%	994	32 913	2,9%	269	6 708	3,9%
2005	12	605	1,9%	39	2 370	1,6%	298	13 307	2,2%	399	17 569	2,2%	948	33 290	2,8%	271	6 645	3,9%
2006	7	636	1,1%	45	2 379	1,9%	266	13 120	2,0%	375	18 489	2,0%	1 012	33 639	2,9%	256	6 625	3,7%
2007	9	702	1,3%	41	2 457	1,6%	230	12 366	1,8%	369	19 777	1,8%	937	34 163	2,7%	239	6 562	3,5%
2008	13	718	1,8%	56	2 505	2,2%	272	12 156	2,2%	453	20 683	2,1%	1 154	34 496	3,2%	259	6 509	3,8%
2009	17	735	2,3%	52	2 524	2,0%	355	11 954	2,9%	518	21 156	2,4%	1 285	34 612	3,6%	296	6 447	4,4%
2010	14	720	1,9%	37	2 575	1,4%	270	11 814	2,2%	478	22 457	2,1%	1 106	35 146	3,1%	202	6 444	3,0%
2011	13	734	1,7%	60	2 662	2,2%	248	11 668	2,1%	523	24 189	2,1%	1 075	36 218	2,9%	225	6 561	3,3%
2012	20	717	2,7%	53	2 669	1,9%	267	11 407	2,3%	511	25 194	2,0%	1 104	36 608	2,9%	221	6 585	3,2%
2013	9	707	1,3%	54	2 682	2,0%	287	11 161	2,5%	575	25 892	2,2%	1 088	37 337	2,8%	235	6 544	3,5%
2014	16	684	2,3%	56	2 663	2,1%	238	10 835	2,1%	608	26 399	2,3%	1 099	37 507	2,8%	198	6 545	2,9%
2015	19	664	2,8%	44	2 624	1,6%	244	10 474	2,3%	502	26 845	1,8%	1 080	37 311	2,8%	250	6 360	3,8%
2016	12	628	1,9%	45	2 562	1,7%	213	10 128	2,1%	550	26 960	2,0%	974	37 035	2,6%	256	6 286	3,9%
2017	13	602	2,1%	41	2 490	1,6%	234	9 676	2,4%	667	26 042	2,5%	1 020	35 253	2,8%	187	5 989	3,0%
Total	*196,0*	*13 074*	*1,5%*	*807*	*47 900*	*1,7%*	*4 725*	*242 098*	*1,9%*	*8 020*	*410 083*	*1,9%*	*17 693*	*674 417*	*2,6%*	*4 228*	*127 503*	*3,2%*

	Health & Education			IT & Electronics			Telecom & Media			Corporate Services			Other & SNI107 Missing			Total		
Year	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate
1998	0	5 022	0,0%	1	3 442	0,0%	0	1 240	0,0%	0	24 477	0,0%	1	4 924	0,0%	12	99 466	0,0%
1999	7	5 428	0,1%	3	3 975	0,1%	2	1 366	0,1%	40	26 250	0,2%	5	5 195	0,1%	218	106 362	0,2%
2000	23	5 867	0,4%	25	4 555	0,5%	11	1 459	0,7%	244	27 839	0,9%	51	5 069	1,0%	1 145	111 322	1,0%
2001	40	6 175	0,6%	35	4 812	0,7%	27	1 480	1,8%	433	29 044	1,5%	129	5 101	2,5%	1 932	114 997	1,7%
2002	62	6 424	1,0%	92	4 922	1,8%	46	1 528	2,9%	647	30 049	2,1%	145	5 049	2,8%	2 671	118 295	2,2%
2003	88	6 640	1,3%	170	5 295	3,1%	47	1 582	2,9%	743	30 843	2,4%	77	4 966	1,5%	3 118	120 938	2,5%
2004	113	6 896	1,6%	157	5 513	2,8%	49	1 621	2,9%	782	31 677	2,4%	76	4 855	1,5%	3 220	123 826	2,5%
2005	124	6 963	1,7%	153	5 539	2,7%	41	1 629	2,5%	753	32 077	2,3%	72	4 834	1,5%	3 110	124 828	2,4%
2006	157	7 035	2,2%	148	5 752	2,5%	25	1 657	1,5%	827	32 391	2,5%	77	4 861	1,6%	3 195	126 584	2,5%
2007	90	7 572	1,2%	92	5 887	1,5%	28	1 844	1,5%	591	34 695	1,7%	75	5 092	1,5%	2 701	131 117	2,0%
2008	125	7 958	1,5%	128	6 066	2,1%	40	1 903	2,1%	761	35 725	2,1%	82	5 225	1,5%	3 343	133 944	2,4%
2009	148	8 211	1,8%	168	6 105	2,7%	41	1 933	2,1%	938	36 368	2,5%	133	5 278	2,5%	3 951	135 323	2,8%
2010	132	8 543	1,5%	130	6 295	2,0%	41	1 994	2,0%	782	37 394	2,0%	248	6 307	3,8%	3 440	139 689	2,4%
2011	149	8 971	1,6%	121	6 590	1,8%	41	2 147	1,9%	860	38 966	2,2%	75	5 465	1,4%	3 390	144 171	2,3%
2012	144	9 203	1,5%	129	6 729	1,9%	50	2 178	2,2%	783	39 749	1,9%	79	5 539	1,4%	3 361	146 578	2,2%
2013	141	9 499	1,5%	145	6 809	2,1%	42	2 225	1,9%	803	40 295	2,0%	87	5 640	1,5%	3 466	148 791	2,3%
2014	158	9 616	1,6%	112	6 931	1,6%	38	2 258	1,7%	786	40 667	1,9%	76	5 702	1,3%	3 385	149 807	2,2%
2015	145	9 804	1,5%	124	6 927	1,8%	46	2 237	2,0%	760	40 584	1,8%	68	5 758	1,2%	3 282	149 588	2,1%
2016	130	9 837	1,3%	118	6 921	1,7%	37	2 236	1,6%	775	40 389	1,9%	66	5 764	1,1%	3 176	148 746	2,1%
2017	150	9 452	1,6%	97	6 665	1,4%	45	2 180	2,0%	765	39 088	1,9%	76	5 634	1,3%	3 295	143 071	2,3%
Total	*2 126*	*155 116*	*1,4%*	*2 148*	*115 730*	*1,8%*	*697*	*36 697*	*1,9%*	*13 073*	*688 567*	*1,9%*	*1 698*	*106 258*	*1,6%*	*55 411*	*2 617 443*	*2,1%*

This table reports the number of defaults, non-defaults and the default rate for the population by industry and year. These observations include all observations for the population of 245,844 companies that make up the 2,672,854 firm-year observations (including missing entries).

Table 6: Industry Default Rates in the One-Year Model, 1998-2017

	Energy & Environment			Materials			Industrial Goods			Construction Industry			Shopping Goods			Convenience Goods		
Year	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate
1999	1	440	0,2%	4	1 535	0,3%	48	10 223	0,5%	33	11 242	0,3%	110	22 020	0,5%	31	4 674	0,7%
2000	0	467	0,0%	8	1 641	0,5%	68	10 772	0,6%	59	11 827	0,5%	160	23 402	0,7%	65	4 960	1,3%
2001	3	468	0,6%	8	1 643	0,5%	90	10 700	0,8%	72	11 997	0,6%	256	23 417	1,1%	75	4 841	1,5%
2002	0	470	0,0%	11	1 611	0,7%	89	10 541	0,8%	103	12 125	0,8%	236	23 134	1,0%	91	4 735	1,9%
2003	5	481	1,0%	12	1 637	0,7%	109	10 353	1,0%	107	12 216	0,9%	234	23 230	1,0%	88	4 654	1,9%
2004	2	437	0,5%	11	1 693	0,6%	86	10 237	0,8%	111	12 224	0,9%	254	23 366	1,1%	75	4 702	1,6%
2005	2	428	0,5%	11	1 740	0,6%	73	10 100	0,7%	88	12 490	0,7%	222	23 624	0,9%	78	4 750	1,6%
2006	0	446	0,0%	9	1 800	0,5%	54	9 803	0,5%	77	12 745	0,6%	227	23 423	1,0%	71	4 679	1,5%
2007	3	449	0,7%	5	1 780	0,3%	32	9 685	0,3%	75	13 535	0,6%	156	23 679	0,7%	49	4 644	1,0%
2008	1	492	0,2%	12	1 808	0,7%	68	9 040	0,7%	72	14 291	0,5%	243	23 790	1,0%	57	4 646	1,2%
2009	2	496	0,4%	5	1 818	0,3%	55	8 744	0,6%	79	14 654	0,5%	208	23 637	0,9%	53	4 588	1,1%
2010	1	496	0,2%	7	1 792	0,4%	41	8 516	0,5%	70	14 990	0,5%	164	23 371	0,7%	48	4 547	1,0%
2011	1	490	0,2%	5	1 836	0,3%	36	8 203	0,4%	75	15 868	0,5%	151	23 546	0,6%	58	4 481	1,3%
2012	1	479	0,2%	10	1 902	0,5%	52	7 927	0,7%	84	16 481	0,5%	148	23 726	0,6%	46	4 597	1,0%
2013	1	474	0,2%	7	1 888	0,4%	47	7 630	0,6%	92	17 010	0,5%	199	23 954	0,8%	38	4 563	0,8%
2014	2	453	0,4%	7	1 854	0,4%	37	7 375	0,5%	108	17 525	0,6%	198	24 457	0,8%	39	4 542	0,9%
2015	3	436	0,7%	4	1 846	0,2%	28	7 053	0,4%	107	17 953	0,6%	165	24 573	0,7%	56	4 452	1,2%
2016	2	422	0,5%	6	1 807	0,3%	27	6 840	0,4%	110	18 446	0,6%	178	24 766	0,7%	51	4 402	1,1%
2017	2	399	0,5%	5	1 763	0,3%	26	6 605	0,4%	75	18 765	0,4%	102	24 770	0,4%	24	4 361	0,5%
Total	*32*	*8 723*	*0,4%*	*147*	*33 394*	*0,4%*	*1 066*	*170 347*	*0,6%*	*1 597*	*276 384*	*0,6%*	*3 611*	*449 885*	*0,8%*	*1 093*	*87 818*	*1,2%*

	Health & Education			IT & Electronics			Telecom & Media			Corporate Services			Other & SNI107 Missing			Total		
Year	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate	Defaults	Non-Defaults	Default Rate
1999	9	3 950	0,2%	5	2 364	0,2%	3	855	0,3%	56	18 125	0,3%	9	3 585	0,3%	309	79 013	0,4%
2000	19	4 316	0,4%	16	2 785	0,6%	7	958	0,7%	118	19 661	0,6%	6	3 922	0,2%	526	84 711	0,6%
2001	22	4 487	0,5%	33	3 020	1,1%	13	946	1,4%	184	19 893	0,9%	20	3 665	0,5%	776	85 077	0,9%
2002	28	4 657	0,6%	48	3 040	1,6%	14	902	1,5%	210	19 887	1,0%	26	3 533	0,7%	856	84 635	1,0%
2003	34	4 744	0,7%	54	3 060	1,7%	12	908	1,3%	211	19 893	1,0%	25	3 442	0,7%	891	84 618	1,0%
2004	36	4 855	0,7%	40	3 135	1,3%	12	896	1,3%	191	19 704	1,0%	19	3 353	0,6%	837	84 602	1,0%
2005	54	4 950	1,1%	34	3 140	1,1%	7	923	0,8%	213	19 833	1,1%	16	3 213	0,5%	798	85 191	0,9%
2006	32	4 881	0,7%	19	3 103	0,6%	5	947	0,5%	133	19 898	0,7%	16	3 170	0,5%	643	84 895	0,8%
2007	24	4 929	0,5%	15	3 282	0,5%	3	960	0,3%	108	20 143	0,5%	7	3 168	0,2%	477	86 254	0,5%
2008	41	5 215	0,8%	17	3 315	0,5%	6	1 041	0,6%	142	21 478	0,7%	11	3 247	0,3%	670	88 363	0,8%
2009	16	5 419	0,3%	23	3 429	0,7%	10	1 102	0,9%	140	21 867	0,6%	20	3 272	0,6%	611	89 026	0,7%
2010	19	5 597	0,3%	14	3 441	0,4%	6	1 091	0,5%	120	21 875	0,5%	11	3 278	0,3%	501	88 994	0,6%
2011	29	5 758	0,5%	17	3 483	0,5%	9	1 111	0,8%	107	22 198	0,5%	17	3 942	0,4%	505	90 916	0,6%
2012	21	5 890	0,4%	16	3 604	0,4%	8	1 180	0,7%	111	22 664	0,5%	7	3 342	0,2%	504	91 792	0,5%
2013	36	6 026	0,6%	27	3 660	0,7%	7	1 140	0,6%	116	22 726	0,5%	12	3 378	0,4%	582	92 449	0,6%
2014	24	6 098	0,4%	7	3 740	0,2%	10	1 179	0,8%	99	22 995	0,4%	13	3 445	0,4%	544	93 663	0,6%
2015	32	6 205	0,5%	9	3 736	0,2%	7	1 222	0,6%	112	23 245	0,5%	6	3 458	0,2%	529	94 179	0,6%
2016	26	6 396	0,4%	12	3 831	0,3%	4	1 218	0,3%	90	23 336	0,4%	7	3 508	0,2%	513	94 972	0,5%
2017	16	6 375	0,3%	7	3 884	0,2%	8	1 241	0,6%	57	23 285	0,2%	8	3 531	0,2%	330	94 979	0,3%
Total	*518*	*100 748*	*0,5%*	*413*	*63 052*	*0,7%*	*151*	*19 820*	*0,8%*	*2 518*	*402 706*	*0,6%*	*256*	*65 452*	*0,4%*	*11 402*	*1 678 329*	*0,7%*

This table reports the default rates on which the one-year default models are specified i.e. default rates calculated for time t at time t-1. The reason for the lower one-year default rates when compared to the population default rates in Table 5 are two-fold. Firstly, there are merely 5 months out of a possible 12 months in a one-year default prediction model where it is safe to assume financial accounts are public. Secondly, conservatism on what can be deemed avalable information in light of imputed data lower default rates.

Table 7: The Incidence of Defaults in the Population of Companies

		Unique Companies	
		Nr.	%
Total Number of Unique Companies	*A = B + C + D*	**245 844**	**100,0%**
o/w does not default	*B*	*190 822*	*77,6%*
o/w defaults once	*C*	*54 633*	*22,2%*
o/w defaults twice	*D*	*389*	*0,2%*
Nr of Unique Defaulters	*E = C + D*	***55 022***	***22,4%***

		Defaults		o/w Happening on Active Years	
		Nr.	%	Nr.	%
Total Number of Defaults	*F = C + G*	**55 411**	**100,0%**	5 338	9,6%
From Companies That Default Once	*C*	54 633	98,6%	n.a.	n.a.
From Companies That Default Twice	*G = 2 * D*	778	1,4%	n.a.	n.a.

The table reports the incidence of defaults in the population of 245,844 companies

Table 8: Size and Age by Industry and Year, 1998-2017

	Average Revenue, SEKk											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	19 559	7 478	7 875	4 568	6 603	15 191	3 116	4 737	6 486	4 056	4 348	*6 135*
1999	20 849	7 754	8 268	5 110	6 995	15 522	3 350	4 831	6 824	4 291	4 434	*6 455*
2000	18 403	8 697	8 714	5 543	7 232	15 858	3 525	4 631	5 492	4 532	4 486	*6 692*
2001	24 569	9 133	9 007	5 818	7 279	16 418	3 874	4 576	6 059	4 742	4 794	*6 922*
2002	25 655	8 713	8 999	5 973	7 348	18 203	4 299	4 358	5 727	4 747	4 829	*7 063*
2003	32 088	8 592	8 928	6 130	7 405	18 083	4 482	4 580	5 711	4 803	4 950	*7 153*
2004	30 042	10 449	9 335	6 291	7 450	17 756	4 592	4 819	5 950	4 894	5 180	*7 292*
2005	27 474	10 720	9 827	6 646	7 686	18 211	4 725	4 974	4 957	5 236	5 290	*7 566*
2006	27 046	11 326	10 529	7 112	8 088	19 468	4 909	5 202	4 966	5 520	5 354	*7 964*
2007	23 394	11 682	11 482	7 614	8 368	19 643	4 925	5 103	5 587	5 996	5 870	*8 286*
2008	24 132	9 847	11 815	7 743	8 303	20 724	5 025	5 072	5 531	6 103	6 205	*8 343*
2009	24 726	9 017	10 434	7 152	7 900	21 228	5 291	4 977	4 995	5 627	6 037	*7 877*
2010	27 339	9 310	11 087	6 928	7 862	21 280	5 423	5 014	4 792	5 727	5 273	*7 856*
2011	26 857	12 009	11 902	7 101	7 766	20 694	5 452	4 993	5 355	5 865	6 288	*8 031*
2012	33 536	11 130	11 456	6 881	7 292	20 895	5 594	4 905	5 278	5 826	6 416	*7 823*
2013	33 892	9 784	11 375	6 741	7 070	20 914	5 689	4 763	5 051	5 823	6 421	*7 666*
2014	31 526	9 102	11 623	6 852	6 996	21 080	5 896	4 890	5 226	5 933	6 423	*7 678*
2015	32 014	9 380	12 389	7 194	7 143	21 178	5 825	5 146	5 145	6 038	6 587	*7 857*
2016	33 862	9 916	12 397	7 590	7 335	21 903	6 307	5 553	6 989	6 254	6 886	*8 148*
2017	42 870	12 635	14 100	8 479	8 226	21 447	6 872	6 318	6 687	7 100	7 586	*9 008*
Total	*27 592*	*9 837*	*10 298*	*6 761*	*7 509*	*19 184*	*5 057*	*4 978*	*5 614*	*5 481*	*5 682*	*7 595*

	Average Age, Years											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	15,1	12,0	14,0	12,5	12,9	12,1	10,4	8,0	11,8	11,3	10,8	*12,1*
1999	15,7	12,5	14,4	12,9	13,2	12,3	10,7	8,0	11,8	11,5	11,4	*12,4*
2000	16,2	12,8	14,7	13,2	13,3	12,3	10,9	7,6	11,6	11,6	11,8	*12,5*
2001	16,7	13,1	15,1	13,4	13,5	12,1	10,8	7,8	12,3	11,9	12,6	*12,8*
2002	17,1	13,4	15,5	13,6	13,5	12,3	11,1	8,1	12,4	12,1	13,0	*13,0*
2003	16,8	13,8	15,9	13,9	13,7	12,3	11,3	8,5	12,6	12,3	13,7	*13,3*
2004	17,4	14,1	16,2	14,0	13,7	12,1	11,7	8,7	12,8	12,5	14,2	*13,4*
2005	16,9	14,3	16,5	14,1	13,8	12,1	11,9	8,9	13,1	12,7	14,6	*13,6*
2006	16,8	14,4	16,7	14,0	13,9	12,2	12,1	9,1	12,9	12,7	15,0	*13,6*
2007	16,4	14,5	17,3	13,7	14,0	12,1	11,9	9,0	12,9	12,8	15,3	*13,6*
2008	16,2	14,8	17,7	13,6	14,0	12,4	11,9	9,2	12,6	12,9	15,5	*13,7*
2009	16,4	15,1	18,0	13,6	14,0	12,6	11,8	9,5	12,6	13,1	15,9	*13,8*
2010	16,0	15,0	18,4	13,1	13,9	13,0	11,7	9,5	12,6	12,9	13,5	*13,6*
2011	16,1	14,7	18,6	12,8	13,6	12,8	11,5	9,5	12,0	12,9	16,3	*13,5*
2012	16,0	14,9	18,9	12,5	13,5	13,0	11,3	9,6	12,2	12,8	16,3	*13,4*
2013	16,1	15,1	19,3	12,3	13,3	13,0	11,3	9,7	12,0	12,8	16,3	*13,3*
2014	17,1	15,4	19,8	12,3	13,2	13,1	11,2	9,9	12,2	12,8	16,4	*13,3*
2015	17,2	15,8	20,0	12,2	13,2	13,2	11,0	10,0	12,0	12,8	16,7	*13,3*
2016	17,8	16,3	20,3	12,1	13,1	13,4	11,1	10,2	11,4	12,8	16,8	*13,3*
2017	18,9	17,7	21,6	13,3	14,4	14,6	12,1	11,0	12,6	13,8	17,3	*14,4*
Total	*16,6*	*14,5*	*17,1*	*13,1*	*13,6*	*12,6*	*11,4*	*9,2*	*12,3*	*12,6*	*14,7*	*13,3*

The table reports average revenue in thousands of SEK, and average age, in years, for all active Swedish, non-financial, independent limited liability companies with at least two years of data. The results are presented by year and industry. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 9: New Active Companies Entering the Population, 1998-2017

	Energy & Environment			Materials			Industrial Goods			Construction Industry			Shopping Goods			Convenience Goods		
Year	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %
1998	548	18	3,2%	1 809	56	3,0%	11 818	349	2,9%	13 196	426	3,1%	25 569	1 124	4,2%	5 226	232	4,3%
1999	569	17	2,9%	1 903	85	4,3%	12 458	435	3,4%	13 796	570	4,0%	27 244	1 320	4,6%	5 619	293	5,0%
2000	580	23	3,8%	1 980	90	4,3%	12 783	497	3,7%	14 399	709	4,7%	28 527	1 485	4,9%	5 887	364	5,8%
2001	597	22	3,6%	2 071	86	4,0%	13 127	377	2,8%	15 138	648	4,1%	29 699	1 400	4,5%	6 160	328	5,1%
2002	621	30	4,6%	2 135	74	3,3%	13 289	399	2,9%	15 770	736	4,5%	30 767	1 533	4,7%	6 338	310	4,7%
2003	582	22	3,6%	2 207	84	3,7%	13 488	393	2,8%	16 329	662	3,9%	31 570	1 468	4,4%	6 429	371	5,5%
2004	592	23	3,7%	2 252	90	3,8%	13 505	417	3,0%	16 692	852	4,9%	32 197	1 710	5,0%	6 468	509	7,3%
2005	594	23	3,7%	2 309	100	4,2%	13 252	353	2,6%	17 092	876	4,9%	32 526	1 712	5,0%	6 566	350	5,1%
2006	605	38	5,9%	2 317	107	4,4%	13 004	382	2,9%	17 850	1 014	5,4%	32 951	1 700	4,9%	6 568	313	4,5%
2007	684	27	3,8%	2 377	121	4,8%	12 231	365	2,9%	18 805	1 341	6,7%	33 272	1 828	5,2%	6 438	363	5,3%
2008	699	32	4,4%	2 455	106	4,1%	12 093	335	2,7%	19 884	1 252	5,9%	33 842	1 808	5,1%	6 480	288	4,3%
2009	724	28	3,7%	2 495	81	3,1%	12 004	305	2,5%	20 591	1 083	5,0%	34 064	1 833	5,1%	6 413	330	4,9%
2010	703	31	4,2%	2 476	136	5,2%	11 696	388	3,2%	20 987	1 948	8,5%	33 937	2 315	6,4%	6 335	311	4,7%
2011	712	35	4,7%	2 576	146	5,4%	11 576	340	2,9%	22 609	2 103	8,5%	34 787	2 506	6,7%	6 340	446	6,6%
2012	715	22	3,0%	2 627	95	3,5%	11 398	276	2,4%	24 132	1 573	6,1%	35 558	2 154	5,7%	6 442	364	5,3%
2013	690	26	3,6%	2 643	93	3,4%	11 199	249	2,2%	25 040	1 427	5,4%	36 242	2 183	5,7%	6 450	329	4,9%
2014	682	18	2,6%	2 649	70	2,6%	10 890	183	1,7%	25 637	1 370	5,1%	36 642	1 964	5,1%	6 442	301	4,5%
2015	667	16	2,3%	2 614	54	2,0%	10 543	175	1,6%	26 025	1 322	4,8%	36 607	1 784	4,6%	6 337	273	4,1%
2016	632	8	1,3%	2 559	48	1,8%	10 189	152	1,5%	26 361	1 149	4,2%	36 503	1 506	4,0%	6 324	218	3,3%
Total	*12 196*	*459*	*3,6%*	*44 454*	*1 722*	*3,7%*	*230 543*	*6 370*	*2,7%*	*370 333*	*21 061*	*5,4%*	*622 504*	*33 333*	*5,1%*	*119 262*	*6 293*	*5,0%*

	Health & Education			IT & Electronics			Telecom & Media			Corporate Services			Other & SNI107 Missing			Total		
Year	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %	Existing Firm	New Firm	New Firms, %
1998	4 790	232	4,6%	3 139	304	8,8%	1 178	62	5,0%	23 444	1 033	4,2%	4 198	727	14,8%	94 915	4 563	4,6%
1999	5 120	315	5,8%	3 573	405	10,2%	1 275	93	6,8%	24 933	1 357	5,2%	4 252	948	18,2%	100 742	5 838	5,5%
2000	5 471	419	7,1%	3 973	607	13,3%	1 345	125	8,5%	26 332	1 751	6,2%	4 450	670	13,1%	105 727	6 740	6,0%
2001	5 849	366	5,9%	4 512	335	6,9%	1 452	55	3,6%	28 079	1 398	4,7%	4 739	491	9,4%	111 423	5 506	4,7%
2002	6 164	322	5,0%	4 728	286	5,7%	1 501	73	4,6%	29 270	1 426	4,6%	4 807	387	7,5%	115 390	5 576	4,6%
2003	6 442	286	4,3%	5 182	283	5,2%	1 554	75	4,6%	30 280	1 306	4,1%	4 773	270	5,4%	118 836	5 220	4,2%
2004	6 671	338	4,8%	5 315	355	6,3%	1 571	99	5,9%	30 955	1 504	4,6%	4 710	221	4,5%	120 928	6 118	4,8%
2005	6 736	351	5,0%	5 375	317	5,6%	1 600	70	4,2%	31 297	1 533	4,7%	4 733	173	3,5%	122 080	5 858	4,6%
2006	6 852	340	4,7%	5 528	372	6,3%	1 597	85	5,1%	31 541	1 677	5,0%	4 783	155	3,1%	123 596	6 183	4,8%
2007	7 161	501	6,5%	5 554	425	7,1%	1 764	108	5,8%	33 256	2 030	5,8%	5 002	165	3,2%	126 544	7 274	5,4%
2008	7 562	521	6,4%	5 837	357	5,8%	1 840	103	5,3%	34 546	1 940	5,3%	5 115	192	3,6%	130 353	6 934	5,1%
2009	7 855	504	6,0%	5 953	320	5,1%	1 874	100	5,1%	35 704	1 602	4,3%	5 253	158	2,9%	132 930	6 344	4,6%
2010	8 054	621	7,2%	5 951	474	7,4%	1 903	132	6,5%	35 881	2 295	6,0%	5 491	1 064	16,2%	133 414	9 715	6,8%
2011	8 495	625	6,9%	6 185	526	7,8%	2 005	183	8,4%	37 271	2 555	6,4%	5 289	251	4,5%	137 845	9 716	6,6%
2012	8 820	527	5,6%	6 482	376	5,5%	2 120	108	4,8%	38 542	1 990	4,9%	5 433	185	3,3%	142 269	7 670	5,1%
2013	9 117	523	5,4%	6 609	345	5,0%	2 155	112	4,9%	39 311	1 787	4,3%	5 535	192	3,4%	144 991	7 266	4,8%
2014	9 309	465	4,8%	6 713	330	4,7%	2 194	102	4,4%	39 709	1 744	4,2%	5 591	187	3,2%	146 458	6 734	4,4%
2015	9 427	522	5,2%	6 788	263	3,7%	2 215	68	3,0%	39 909	1 435	3,5%	5 655	171	2,9%	146 787	6 083	4,0%
2016	9 657	310	3,1%	6 803	236	3,4%	2 203	70	3,1%	40 031	1 133	2,8%	5 715	115	2,0%	146 977	4 945	3,3%
Total	*139 552*	*8 088*	*5,5%*	*104 200*	*6 916*	*6,2%*	*33 346*	*1 823*	*5,2%*	*630 291*	*31 496*	*4,8%*	*95 524*	*6 722*	*6,6%*	*2 402 205*	*124 283*	*4,9%*

The table shows the number of firms that existed in the in the population during the calendar year, and how many newly incorporated firms entered the population. The number of new firms is shown as a % of all firm observations during the same year as the firm entered the population. Population is defined in the same way as in Table 5, i.e. the observations belong to companies that at some point are deemed active.

Table 11: Average Values of Variables Used in the Composite Models, 1998-2017

	Horizon*					Year																				
	1	2	3	4	5	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	***Total***
Baseline Hazard	x	x	x	x	x	3,3%	2,6%	2,9%	3,4%	4,0%	4,2%	4,2%	3,9%	3,8%	3,2%	3,5%	3,9%	3,3%	3,1%	3,0%	2,9%	2,7%	2,6%	2,4%	2,5%	*3,3%*
ROAA	x					8,1%	8,9%	8,9%	7,7%	7,3%	6,8%	6,9%	7,7%	9,7%	10,0%	8,8%	8,1%	9,9%	10,6%	9,8%	10,4%	11,8%	13,3%	13,4%	13,9%	*9,6%*
RETA	x					9,0%	9,7%	10,8%	12,2%	12,8%	12,7%	12,7%	13,1%	13,7%	14,2%	14,9%	14,8%	14,2%	15,0%	15,8%	15,7%	16,3%	17,0%	17,6%	18,9%	*14,1%*
EQTA	x					25,7%	26,7%	27,3%	27,5%	27,8%	28,1%	28,7%	29,9%	30,8%	30,7%	30,5%	30,9%	31,2%	31,5%	31,3%	32,3%	34,0%	35,1%	35,5%	36,1%	*30,6%*
UTREMP	x	x	x	x	x	52,6	57,0	59,5	63,6	67,5	71,6	69,6	64,2	65,1	67,9	72,0	73,4	73,8	77,3	83,8	84,0	83,9	84,3	85,3	89,9	*72,5*
DPO	x					27,0	27,4	27,0	26,2	25,8	25,2	25,7	26,0	25,3	24,9	24,1	23,7	24,2	22,2	20,9	20,7	19,7	19,1	18,6	16,1	*23,4*
Interest / Liabilities	x	x	x			3,3%	3,0%	3,3%	3,7%	4,1%	3,1%	2,7%	2,4%	2,3%	2,7%	3,6%	2,4%	2,0%	2,5%	2,2%	1,9%	1,8%	1,7%	1,5%	1,5%	*2,6%*
Audit. - Not Recommend	x		x			1,6%	1,6%	1,7%	1,7%	1,7%	1,9%	2,1%	2,2%	2,2%	2,1%	2,1%	2,0%	1,8%	1,3%	1,2%	1,0%	0,9%	0,8%	0,7%	0,5%	*1,5%*
Logarithm of Assets	x			x	x	7,3	7,3	7,4	7,4	7,4	7,4	7,5	7,5	7,5	7,6	7,6	7,6	7,6	7,5	7,5	7,5	7,5	7,5	7,6	7,6	*7,5*
Logarithm of Sales	x					8,0	8,0	8,0	8,1	8,1	8,1	8,1	8,1	8,2	8,2	8,2	8,2	8,1	8,2	8,1	8,1	8,1	8,2	8,2	8,3	*8,1*
Added Value / Employee		x	x	x	x	313,2	326,6	338,0	347,6	355,9	364,0	373,7	386,2	406,0	426,5	430,8	426,1	437,5	452,5	454,2	459,3	472,3	492,7	525,4	570,6	*418,7*
Late Filing Estimate		x	x	x	x	22,1%	16,1%	14,8%	11,0%	9,7%	7,8%	12,0%	13,4%	12,5%	10,5%	7,9%	6,8%	8,9%	4,7%	3,6%	7,5%	3,8%	4,8%	4,3%	20,0%	*9,8%*
Cash / Current Assets		x	x			30,8%	31,8%	32,3%	32,6%	33,1%	33,3%	33,6%	34,5%	35,2%	35,8%	36,7%	37,4%	37,9%	38,5%	39,0%	40,0%	41,4%	42,8%	42,4%	42,2%	*36,7%*
Current Liab. / Assets		x	x	x	x	45,4%	45,4%	45,3%	45,0%	44,7%	44,7%	45,8%	46,3%	46,1%	46,3%	45,9%	45,6%	45,9%	45,6%	45,3%	44,7%	43,6%	43,4%	43,4%	43,3%	*45,1%*
Dividend / Employee		x	x			4,04	4,66	4,78	4,49	4,56	4,83	6,12	11,90	17,77	21,11	20,76	21,09	22,03	22,63	22,52	26,91	32,47	40,68	48,96	42,01	*19,48*
Logarithm of Age		x	x	x	x	2,12	2,14	2,14	2,15	2,16	2,18	2,19	2,19	2,19	2,17	2,17	2,18	2,13	2,11	2,09	2,09	2,10	2,10	2,10	2,26	*2,15*
EBIE Margin		x				5,2%	5,3%	5,3%	4,5%	4,1%	4,1%	4,3%	4,6%	5,7%	6,0%	5,4%	4,8%	5,7%	6,1%	5,7%	6,0%	6,5%	7,1%	7,3%	7,3%	*5,6%*
Log. Interest / Liabilities				x		3,2%	2,9%	3,2%	3,6%	3,9%	3,0%	2,6%	2,3%	2,2%	2,6%	3,4%	2,3%	2,0%	2,4%	2,1%	1,9%	1,7%	1,6%	1,4%	1,5%	*2,5%*
Trade Days (DSO - DPO)		x	x			9,5	9,6	10,6	10,4	10,9	10,8	10,6	11,5	12,7	13,2	12,9	13,7	16,3	15,6	15,6	15,8	15,9	16,4	16,1	16,2	*13,3*
Industry Default rate				x	x	3,1%	2,5%	2,7%	3,2%	3,7%	4,1%	4,1%	3,9%	3,8%	3,2%	3,4%	3,8%	3,2%	3,0%	3,0%	2,9%	2,8%	2,6%	2,5%	2,5%	*3,2%*
Any Auditor Remark				x	x	10,0%	10,5%	10,3%	11,1%	11,9%	12,8%	13,9%	14,7%	15,2%	15,6%	15,8%	15,9%	16,1%	13,3%	12,0%	11,2%	10,1%	9,1%	8,8%	8,0%	*12,3%*
Dividend Dummy				x		34,0%	33,0%	33,0%	31,0%	30,0%	29,0%	32,0%	41,0%	47,0%	47,0%	45,0%	45,0%	43,0%	41,0%	39,0%	40,0%	42,0%	45,0%	46,0%	42,0%	*39,0%*
STDEQ					x	10,0%	10,6%	11,5%	12,4%	13,6%	16,1%	20,4%	21,8%	21,2%	21,5%	21,9%	20,5%	19,7%	19,5%	18,6%	18,2%	16,1%	14,6%	14,1%	13,6%	*16,9%*
Tax Cost / Assets					x	1,8%	1,9%	2,0%	1,9%	1,8%	1,7%	1,8%	2,1%	2,4%	2,6%	2,5%	2,4%	2,5%	2,7%	2,5%	2,4%	2,6%	2,9%	3,0%	3,1%	*2,4%*

*The table reports average values for the variables included in the Composite Model for the one-to-five-year prediction horizons. Note that the baseline hazard is specified on the population before removing inactive years. * x = included in default model forecast horizon. The baseline hazard rate is represented by the trailing 12-month realized rate of default. ROAA is the return on average assets, or earnings before interest expenses to average total assets. RETA is retained earnings to total assets. EQTA is total shareholders equity to total assets. Interest to liabilities is (Financial costs - Financial expenses affecting comparability) / (Non-current liabilities + Current liabilities + Provisions + Deferred tax liability). Audit. - Not Recommend is a dummy variable for an auditor remark stating the company's financials are not recommended. DPO are the days payables outstanding, defined as (Trade Payables/Sales)*365. UTREMP are untaxed reserves per employee. EBIE margin is the earnings before interest expense margin. The logarithm of Interest / Liabilities is the natural logarithm of the above interest to liabilities ratio. Any auditor remark is any of the remarks which does not constitute a recommendation. STDEQ is the short-term debt to equity.*

Table 12: Defaulters and Non-Defaulters: One-Year Ahead Average Values of Composite Model Variables

		One-year model	Year																		
			1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Return on Average Assets	ND	x	8,1%	9,0%	9,0%	7,8%	7,4%	6,8%	7,0%	7,8%	9,8%	10,1%	9,0%	8,2%	10,0%	10,8%	9,9%	10,5%	11,9%	13,4%	13,5%
	D		-3,7%	1,9%	-0,6%	-1,5%	-1,2%	-0,9%	-0,8%	0,1%	0,8%	-2,9%	-6,8%	-6,9%	-6,2%	-9,0%	-8,7%	-8,2%	-8,5%	-4,6%	-10,7%
Retained Earnings to Total Assets	ND	x	9,1%	9,7%	10,8%	12,2%	12,9%	12,8%	12,8%	13,2%	13,7%	14,3%	15,0%	14,9%	14,3%	15,1%	15,9%	15,8%	16,4%	17,1%	17,7%
	D		-1,2%	3,2%	3,8%	6,9%	6,0%	8,8%	7,7%	5,1%	5,0%	4,6%	0,1%	0,2%	-7,3%	-1,2%	1,3%	-0,6%	2,3%	3,6%	1,7%
Equity to Total Assets	ND	x	25,8%	26,7%	27,3%	27,5%	27,8%	28,0%	28,7%	29,9%	30,8%	30,7%	30,6%	31,0%	31,4%	31,6%	31,5%	32,5%	34,1%	35,2%	35,6%
	D		5,4%	22,1%	22,9%	26,6%	26,7%	28,6%	29,4%	28,3%	30,1%	19,8%	12,9%	12,0%	10,6%	6,8%	10,8%	4,9%	13,4%	16,1%	5,9%
Untaxed Reserves per Employee	ND	x	52,8	57,3	59,9	64,0	68,0	72,1	70,0	64,6	65,4	68,4	72,4	73,8	74,2	77,7	84,3	84,5	84,4	84,8	85,6
	D		12,0	21,1	16,0	22,8	23,5	24,6	33,6	13,3	10,3	9,3	13,1	8,5	6,5	6,4	7,9	11,5	7,1	6,7	5,1
Days Payables Outstanding	ND	x	26,9	27,4	26,9	26,1	25,7	25,1	25,6	25,9	25,2	24,8	24,0	23,6	24,1	22,1	20,9	20,6	19,6	19,0	18,5
	D		43,0	33,2	34,1	32,2	29,8	29,9	29,4	28,8	31,9	33,7	35,4	31,4	33,6	36,7	32,1	33,4	30,5	30,3	28,5
Interest Cost to Total Liabilities	ND	x	3,3%	3,0%	3,3%	3,7%	4,1%	3,1%	2,7%	2,4%	2,3%	2,7%	3,6%	2,4%	2,0%	2,5%	2,2%	1,9%	1,8%	1,6%	1,5%
	D		4,8%	4,3%	5,0%	5,5%	6,8%	5,0%	4,0%	3,7%	4,0%	4,3%	5,5%	4,0%	3,7%	3,8%	3,9%	3,5%	3,2%	3,2%	3,0%
Auditor Does Not Recommend	ND	x	1,5%	1,6%	1,7%	1,6%	1,7%	1,9%	2,1%	2,1%	2,2%	2,1%	2,0%	1,9%	1,7%	1,3%	1,2%	1,0%	0,9%	0,8%	0,7%
	D		10,4%	9,1%	5,9%	5,6%	3,8%	6,2%	6,9%	7,9%	9,2%	9,0%	8,3%	10,4%	11,1%	6,3%	6,7%	5,5%	5,7%	3,9%	5,2%
Logarithm of Assets	ND	x	7,28	7,34	7,38	7,41	7,42	7,44	7,48	7,51	7,55	7,59	7,58	7,57	7,56	7,55	7,52	7,50	7,49	7,52	7,56
	D		7,32	6,91	7,05	6,90	6,88	6,88	6,93	6,78	6,84	7,07	7,12	6,99	7,02	7,08	6,92	6,81	6,74	6,78	6,80
Logarithm of Sales	ND	x	7,98	8,00	8,03	8,07	8,08	8,09	8,11	8,13	8,17	8,21	8,21	8,17	8,14	8,16	8,15	8,13	8,13	8,15	8,20
	D		8,18	7,75	7,78	7,66	7,63	7,68	7,58	7,62	7,66	7,90	8,03	7,94	7,98	8,04	7,97	7,91	7,83	7,92	8,03
Value Added per Employee	ND		313,5	327,1	338,8	348,6	357,0	365,0	374,7	387,2	406,9	427,4	431,8	426,9	438,3	453,4	455,2	460,2	473,3	493,8	526,1
	D		227,8	254,2	249,6	250,4	260,2	269,8	284,2	269,5	270,9	302,4	292,5	294,2	301,0	300,2	307,3	309,7	307,2	318,1	334,0
Late Filing Estimate	ND		22,1%	16,0%	14,8%	11,0%	9,7%	7,8%	12,0%	13,4%	12,4%	10,4%	7,9%	6,8%	8,9%	4,6%	3,6%	7,5%	3,8%	4,8%	4,3%
	D		37,5%	18,4%	20,7%	10,6%	9,0%	9,0%	11,7%	15,1%	19,3%	16,6%	14,7%	18,2%	16,0%	14,7%	6,2%	19,9%	8,7%	10,7%	7,6%
Cash to Current Assets	ND		30,9%	31,8%	32,2%	32,5%	33,0%	33,3%	33,5%	34,5%	35,2%	35,8%	36,7%	37,4%	37,9%	38,6%	39,1%	40,1%	41,4%	42,9%	42,5%
	D		17,3%	36,3%	35,7%	38,1%	40,7%	40,8%	42,3%	39,8%	41,3%	31,7%	31,0%	31,3%	27,6%	25,8%	26,6%	28,3%	30,8%	31,3%	25,1%
Current Liabilities to Total Assets	ND		45,3%	45,4%	45,3%	45,0%	44,6%	44,7%	45,8%	46,3%	46,1%	46,2%	45,8%	45,5%	45,8%	45,5%	45,2%	44,5%	43,5%	43,2%	43,3%
	D		58,9%	51,8%	49,8%	49,5%	49,3%	49,0%	50,9%	54,0%	54,4%	59,9%	63,7%	65,1%	67,3%	68,4%	67,9%	70,3%	66,3%	65,4%	69,1%
Dividend per Employee	ND		4,05	4,66	4,78	4,49	4,55	4,80	6,11	11,91	17,77	21,17	20,82	21,16	22,08	22,71	22,61	27,01	32,56	40,81	49,11
	D		1,68	4,98	4,23	4,26	5,72	7,08	6,17	12,01	17,17	13,95	13,54	10,27	13,17	8,49	8,89	12,04	17,69	19,29	10,25
Logarithm of Age	ND		2,13	2,14	2,14	2,15	2,16	2,18	2,19	2,19	2,19	2,17	2,17	2,18	2,13	2,11	2,09	2,09	2,10	2,10	2,10
	D		1,90	2,13	2,10	2,08	2,14	2,24	2,21	2,18	2,20	2,12	2,05	2,07	2,04	1,97	1,91	1,85	1,87	1,90	1,79
EBIE Margin	ND		5,2%	5,3%	5,3%	4,6%	4,1%	4,1%	4,3%	4,7%	5,7%	6,1%	5,4%	4,8%	5,8%	6,2%	5,8%	6,1%	6,6%	7,2%	7,4%
	D		-1,7%	1,6%	-0,3%	-0,9%	-0,6%	-1,0%	0,8%	0,7%	1,5%	-1,1%	-3,4%	-3,6%	-3,4%	-6,0%	-4,7%	-5,8%	-4,5%	-4,6%	-6,2%
Log of Interest to Liabilities	ND		3,2%	2,9%	3,2%	3,5%	3,8%	3,0%	2,6%	2,3%	2,2%	2,6%	3,3%	2,3%	2,0%	2,4%	2,1%	1,9%	1,7%	1,6%	1,4%
	D		4,6%	4,1%	4,8%	5,2%	6,2%	4,7%	3,9%	3,6%	3,8%	4,1%	5,0%	3,8%	3,6%	3,6%	3,7%	3,4%	3,1%	3,1%	2,9%
Trade Days (DSO - DPO)	ND		9,6	9,7	10,6	10,4	10,9	10,8	10,7	11,6	12,8	13,3	13,0	13,8	16,4	15,7	15,7	15,9	16,1	16,5	16,2
	D		-10,8	-1,0	0,8	3,3	6,8	4,0	5,7	3,0	3,7	-0,4	-4,6	2,7	1,1	-6,0	-3,1	-4,6	-2,9	-1,1	-1,7
Current Industry Default Rate	ND		3,1%	2,5%	2,7%	3,2%	3,7%	4,1%	4,1%	3,9%	3,7%	3,2%	3,4%	3,8%	3,2%	3,0%	3,0%	2,9%	2,8%	2,6%	2,5%
	D		3,3%	2,7%	2,9%	3,3%	3,9%	4,2%	4,2%	4,0%	3,9%	3,3%	3,6%	4,0%	3,3%	3,2%	3,1%	3,1%	3,0%	2,9%	2,6%
Any Auditor Remark	ND		9,9%	10,3%	10,2%	10,9%	11,7%	12,7%	13,7%	14,6%	15,1%	15,4%	15,6%	15,7%	15,9%	13,1%	11,8%	11,1%	10,0%	9,0%	8,7%
	D		47,6%	29,8%	28,2%	24,3%	24,2%	25,3%	28,1%	30,2%	31,4%	40,6%	43,5%	48,3%	49,5%	44,2%	39,7%	32,2%	24,8%	23,0%	21,8%
Dividend Paid Dummy	ND		34,0%	33,3%	32,8%	31,2%	30,2%	29,3%	32,0%	41,6%	47,0%	47,7%	45,2%	45,7%	43,2%	40,9%	39,2%	39,9%	42,0%	45,1%	46,5%
	D		11,3%	22,2%	20,4%	20,9%	22,9%	24,4%	23,8%	22,7%	24,3%	21,9%	19,8%	14,6%	14,9%	11,5%	10,8%	10,5%	15,1%	17,5%	10,6%
Short Term Debt To Equity	ND		10,0%	10,7%	11,5%	12,4%	13,7%	16,2%	20,6%	21,9%	21,2%	21,5%	21,9%	20,6%	19,7%	19,5%	18,6%	18,2%	16,1%	14,6%	14,0%
	D		16,4%	6,3%	8,0%	10,6%	7,7%	10,2%	10,0%	10,8%	14,7%	20,6%	26,7%	16,9%	21,8%	29,2%	28,9%	21,5%	16,7%	16,6%	17,4%
Tax Cost to Total Assets	ND		1,8%	1,9%	2,0%	1,9%	1,8%	1,7%	1,8%	2,1%	2,4%	2,6%	2,5%	2,4%	2,5%	2,7%	2,5%	2,4%	2,6%	2,9%	3,0%
	D		0,8%	2,1%	2,1%	2,0%	2,0%	2,0%	2,2%	2,4%	2,3%	2,0%	1,9%	1,6%	1,7%	1,3%	1,5%	1,6%	1,5%	1,8%	1,8%

This table shows the average values of the Composite Model variables for firms that defaulted as well as those that did not default in the one-year prediction horizon. Note that while the average values by defaulters and non-defaulters are reported for the variables included in longer-term Composite Models, these models are designed to predict longer term defaults. Still, there are clear differences between the means of defaulters and non-defaulters for the longer term variables also for one-year predictions. ND = non-defaulter; D = Defaulter.

Table 13: Summary Table of Results

One year Prediction

	Calibration						AUROC				AUPRC			
	Defaults	Non Defaults	Total	Default rate	Actual	Δ	$1^{\dagger\dagger\dagger}$	2***	$3^{\dagger\dagger\dagger}$	4***	1	2	3	4
Altman (1968)	13 590	1 674 684	1 688 274	0,8%	2,1%	-1,3%	73,8%	76,9%	74,2%	78,1%	2,3%	1,7%	2,4%	1,7%
Zmijewski (1984)	18 124	1 669 660	1 687 784	1,1%	2,1%	-1,0%	65,6%	68,8%	66,8%	72,5%	1,8%	1,3%	2,0%	1,4%
Shumway (2001)	3 546	1 364 692	1 368 238	0,3%	2,1%	-1,9%	65,9%	67,9%	66,2%	69,2%	1,8%	1,3%	1,9%	1,4%
Altman and Sabato (2007)	13 359	1 427 179	1 440 538	0,9%	2,1%	-1,2%	71,7%	74,3%	72,5%	75,6%	2,2%	1,9%	2,4%	2,4%
Dakovic et al. (2010)	7 263	1 681 397	1 688 660	0,4%	2,1%	-1,7%	76,1%	76,4%	76,1%	74,4%	2,5%	1,7%	2,5%	1,5%
Composite Model	**26 547**	**1 660 485**	**1 687 032**	**1,6%**	**2,1%**	**-0,5%**	**74,9%**	**80,6%**	**75,7%**	**83,0%**	**2,7%**	**2,5%**	**2,9%**	**3,2%**

Two Year Prediction

	Calibration						AUROC				AUPRC			
	Failures	Non Failures	Total	Default rate	Actual	Δ	$1^{\dagger\dagger\dagger}$	2***	$3^{\dagger\dagger\dagger}$	4***	1	2	3	4
Altman (1968)	16 367	1 577 881	1 594 248	1,0%	2,1%	-1,1%	72,2%	73,8%	72,1%	74,2%	4,9%	4,1%	5,1%	4,2%
Zmijewski (1984)	18 055	1 575 749	1 593 804	1,1%	2,1%	-1,0%	68,2%	69,4%	68,4%	69,8%	5,3%	4,4%	5,4%	4,3%
Shumway (2001)	10 645	1 278 707	1 289 352	0,8%	2,1%	-1,3%	67,8%	70,3%	67,9%	71,6%	4,5%	3,9%	4,8%	4,0%
Altman and Sabato (2007)	15 700	1 353 616	1 369 316	1,2%	2,1%	-1,0%	72,5%	73,0%	72,2%	73,9%	5,4%	4,6%	5,5%	4,7%
Dakovic et al. (2010)	20 164	1 574 467	1 594 631	1,3%	2,1%	-0,8%	76,7%	77,8%	76,5%	77,3%	5,8%	4,5%	5,9%	4,3%
Composite Model	**29 626**	**1 563 481**	**1 593 107**	**1,9%**	**2,1%**	**-0,2%**	**75,4%**	**79,5%**	**75,3%**	**79,6%**	**6,0%**	**5,8%**	**6,1%**	**5,7%**

Three year Prediction

	Calibration						AUROC				AUPRC			
	Failures	Non Failures	Total	Default rate	Actual	Δ	1***	2***	3***	4***	1	2	3	4
Altman (1968)	33 342	1 465 447	1 498 789	2,3%	2,1%	0,2%	67,7%	71,3%	67,9%	69,9%	4,8%	4,1%	4,9%	3,9%
Zmijewski (1984)	33 278	1 465 118	1 498 396	2,3%	2,1%	0,2%	65,0%	67,8%	65,1%	66,6%	4,8%	4,0%	5,0%	3,8%
Shumway (2001)	24 546	1 184 435	1 208 981	2,1%	2,1%	0,0%	66,3%	69,1%	66,2%	68,1%	4,4%	3,6%	4,4%	3,5%
Altman and Sabato (2007)	29 346	1 266 919	1 296 265	2,3%	2,1%	0,2%	68,3%	71,7%	68,4%	70,2%	5,0%	4,3%	5,1%	4,2%
Dakovic et al. (2010)	33 737	1 465 458	1 499 195	2,3%	2,1%	0,2%	72,2%	74,9%	72,2%	74,4%	5,7%	4,5%	5,8%	4,5%
Composite Model	**32 214**	**1 466 116**	**1 498 330**	**2,2%**	**2,1%**	**0,1%**	**72,7%**	**76,3%**	**72,5%**	**76,4%**	**6,7%**	**5,4%**	**7,1%**	**5,8%**

Four Year Prediction

	Calibration						AUROC				AUPRC			
	Failures	Non Failures	Total	Default rate	Actual	Δ	1***	2***	3***	4***	1	2	3	4
Altman (1968)	31 499	1 372 597	1 404 096	2,3%	2,1%	0,2%	65,1%	68,4%	65,2%	67,5%	3,9%	3,3%	3,9%	3,2%
Zmijewski (1984)	31 472	1 372 271	1 403 743	2,3%	2,1%	0,2%	61,9%	63,9%	62,3%	64,1%	3,8%	3,0%	3,9%	3,0%
Shumway (2001)	23 032	1 105 925	1 128 957	2,1%	2,1%	0,0%	63,9%	65,7%	64,1%	66,1%	3,5%	2,8%	3,5%	2,8%
Altman and Sabato (2007)	27 429	1 194 938	1 222 367	2,3%	2,1%	0,2%	65,4%	68,3%	65,2%	67,3%	3,9%	3,3%	4,0%	3,2%
Dakovic et al. (2010)	31 748	1 372 771	1 404 519	2,3%	2,1%	0,2%	69,2%	71,9%	69,1%	71,5%	4,5%	3,6%	4,6%	3,6%
Composite Model	**30 829**	**1 373 622**	**1 404 451**	**2,2%**	**2,1%**	**0,1%**	**70,2%**	**73,1%**	**70,0%**	**72,9%**	**5,1%**	**4,1%**	**5,2%**	**4,0%**

Five Year Prediction

	Calibration						AUROC				AUPRC			
	Failures	Non Failures	Total	Default rate	Actual	Δ	1***	2***	3***	4***	1	2	3	4
Altman (1968)	29 480	1 280 436	1 309 916	2,3%	2,1%	0,2%	63,0%	65,8%	62,7%	66,4%	3,2%	2,7%	3,2%	2,7%
Zmijewski (1984)	29 433	1 280 163	1 309 596	2,3%	2,1%	0,2%	59,3%	61,3%	58,8%	61,9%	3,0%	2,4%	2,9%	2,4%
Shumway (2001)	21 574	1 028 104	1 049 678	2,1%	2,1%	0,0%	62,1%	64,4%	61,9%	64,2%	2,9%	2,4%	2,9%	2,4%
Altman and Sabato (2007)	25 467	1 121 936	1 147 403	2,3%	2,1%	0,1%	62,8%	65,9%	62,3%	65,4%	3,2%	2,7%	3,1%	2,8%
Dakovic et al. (2010)	29 524	1 280 835	1 310 359	2,3%	2,1%	0,2%	66,9%	68,9%	66,4%	69,7%	3,7%	3,0%	3,6%	3,0%
Composite Model	**29 147**	**1 281 071**	**1 310 218**	**2,3%**	**2,1%**	**0,2%**	**67,5%**	**69,9%**	**67,2%**	**70,6%**	**3,9%**	**3,0%**	**3,8%**	**3,2%**

*This table shows the models' estimated number of defaults and non-defaults for the entire population (sampled firm and hold-out firm, in and out-of-time) after Skogsvik-adjusted probabilities, as well as the AUROC and AUPRC results for the four different sample splits. The results above are estimated with random effects on the firm-level for each default prediction horizon. This is reported for all 5 default prediction horizons. 1 = sampled firm, in-time; 2 = sampled firm, out-of-time; 3 = hold-out firm, in-time; 4 = hold-out firm, out-of-time. */ ** /*** Composite Model is significantly superior at the 10% / 5% / 1% level vis-à-vis all other models. † / †† / ††† Composite Model is significantly superior at the 10% / 5% / 1% level vis-à-vis all models except for the highest scoring AUROC/AUPRC model. All model specifications' AUROC and AURPC are statistically above their respective reference lines (i.e. better than random).*

Table 14: One-Year DTDDH Composite Model

	Simple Logit	Simple Logit with Random Effects
Baseline Hazard	15.192***	49.532***
	(2,75)	(4,03)
EBIE/ATA	-2.143***	-3.301***
	(0,08)	(0,15)
RE/TA	-0.641***	-1.665***
	(0,06)	(0,13)
BEQ/TA	0.519***	1.424***
	(0,07)	(0,14)
IE/TL	6.262***	10.907***
	(0,18)	(0,40)
AUDNR	0.847***	1.502***
	(0,06)	(0,10)
LNA	-0.240***	-0.464***
	(0,02)	(0,05)
LNS	0.163***	0.395***
	(0,02)	(0,04)
DPO	0.004***	0.005***
	0,00	0,00
UTR/EMP	-0.007***	-0.012***
	0,00	0,00
constant	-5.140***	-12.014***
	(0,15)	(0,36)
		20.648***
Firm Random Effects		(0,97)
Observations	843 461	843 461
Defaults	11 402	11 402
BIC	66 540	64 007
AIC	66 412	63 867
Chi-Sq	4 734	2 034

*This table reports results from the binary logit regression using Swedish, private, independent, non financial/real estate limited liabilitiy companies. The regression is specified on 843,461 observations representing the sampled firms, in-time (1998-2013). DTDDH = discrete-time duration-dependent hazard. Dependent variable: Failure is a 1 (0) if a firm defaulted (did not default) within one year, contingent on the financial information being released and avaliable. The baseline hazard is respresented by the trailing 12-month realized rate of default. EBIE/ATA is the return on average assets, or earnings before interest expenses to average total assets. RE/TA is retained earnings to total assets. BEQ/TA is total sharehodlers equity to total assets. IE/TL is (Financial costs - Financial expenses affecting comparability) / (Non-current liabilities + Current liabilities + Provisions + Deferred tax liability). AUDNR is a dummy variable for an auditor remark stating the company's financials are not recommended. LNA is the natural logarithm of total assets. LNS is the natural logarithm of sales. DPO are the days payables outstanding, defined as (Trade Payables/Sales)*365. UTR/EMP are untaxed reserves per employee. Variables are winsorised at the 1% level. Standard errors are in parenthases.*

Table 15: Test of Proportional Hazard

	One-Year Model		
	Rho	Chi-Squared	Prob>Chi-Squared
Baseline Hazard	0,056	21,29	0.000***
EBIE/ATA	0,016	3,83	0.051*
RE/TA	-0,046	29,10	0.000***
BEQ/TA	0,056	44,74	0.000***
IE/TL	-0,030	4,41	0.0356**
AUDNR	0,037	9,41	0.0022**
LNA	0,000	0,00	0,993
LNS	-0,065	48,27	0.000***
DPO	-0,031	5,76	0.0164**
UTR/EMP	0,032	35,83	0.000***
Global test		234,72	0.000***

Reports the individual and global test of survival curves' proportionality (the Cox semi-parametric PH assumption) for the one-year model using scaled Schoenfeld residuals.

Figure 3.1: Sampled Firm, In-Time ROC for the Composite Model

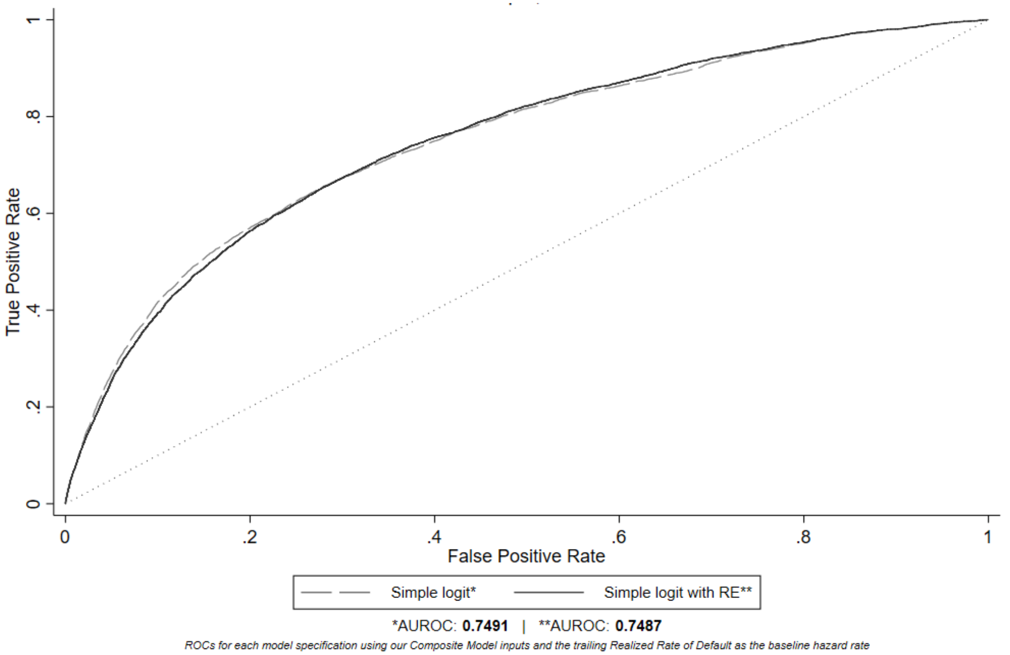


Figure 3.2: Sampled Firm, Out-of-Time ROC for the Composite Model

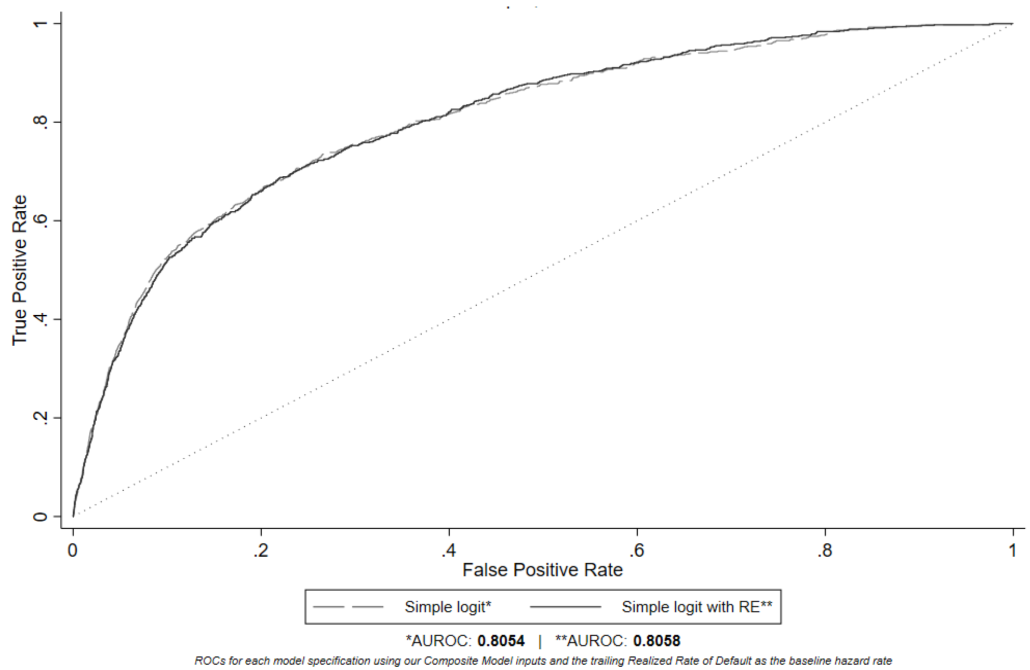


Figure 3.3: Hold-Out, In-Time ROC for the Composite Model

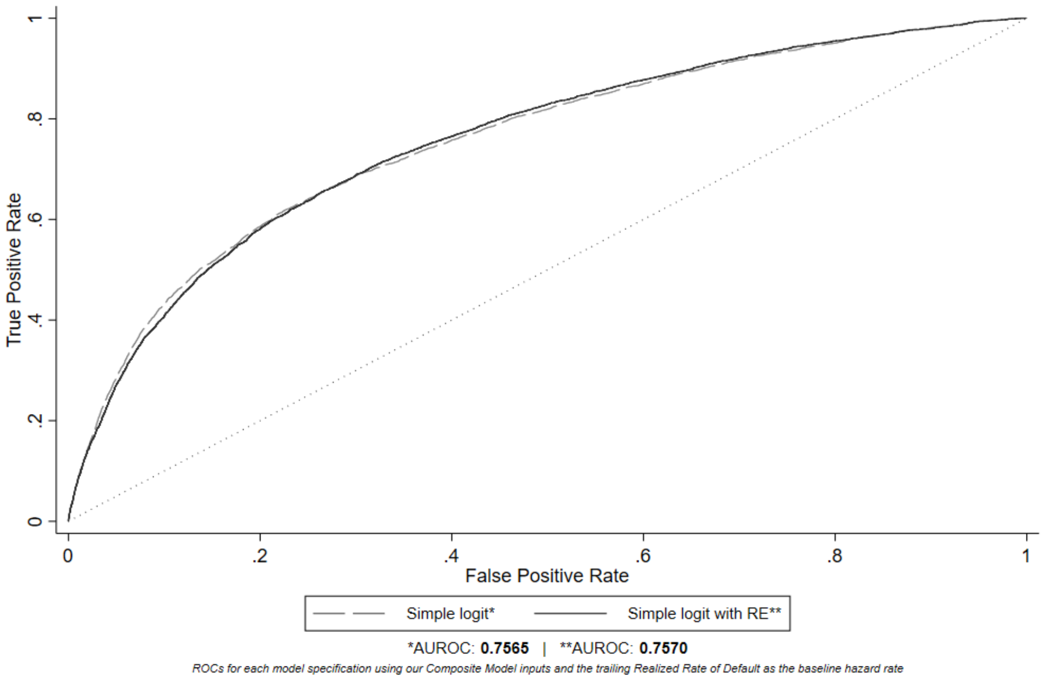


Figure 3.4: Hold-Out, Out-of-Time ROC for the Composite Model

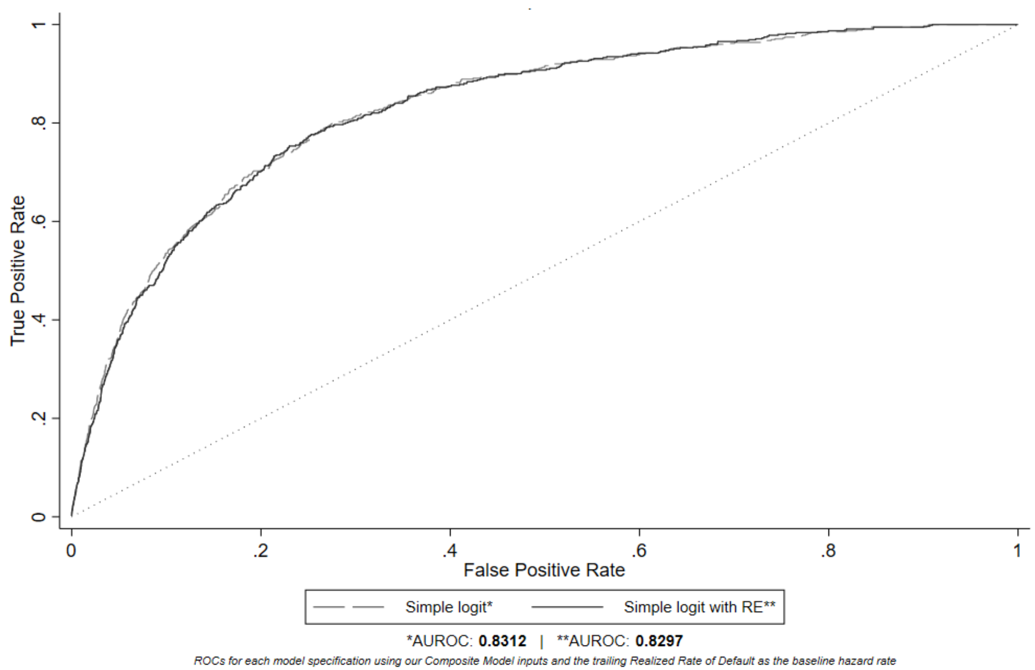
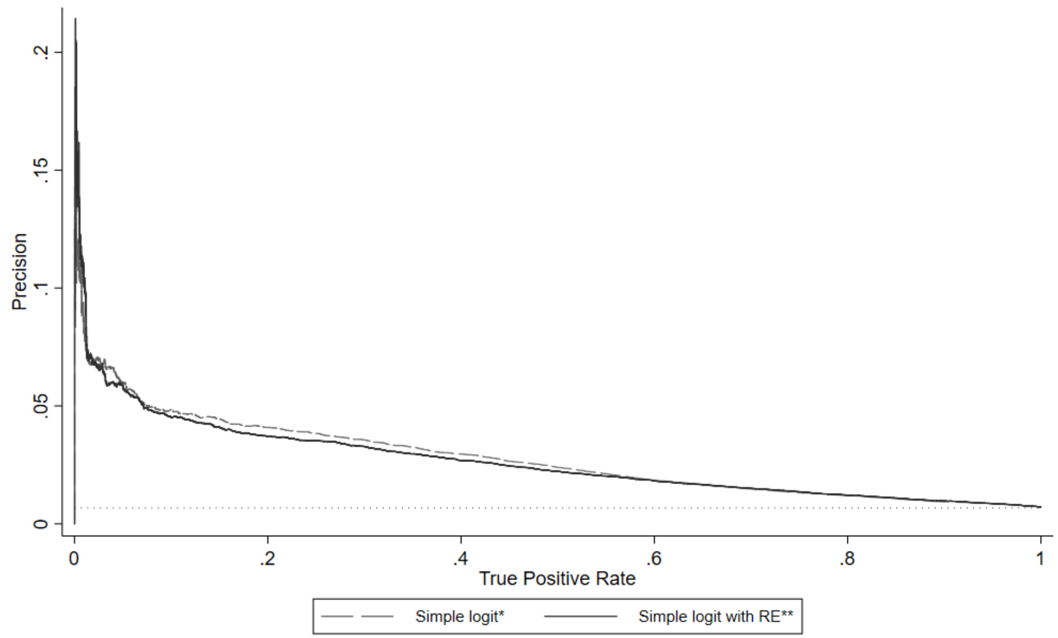


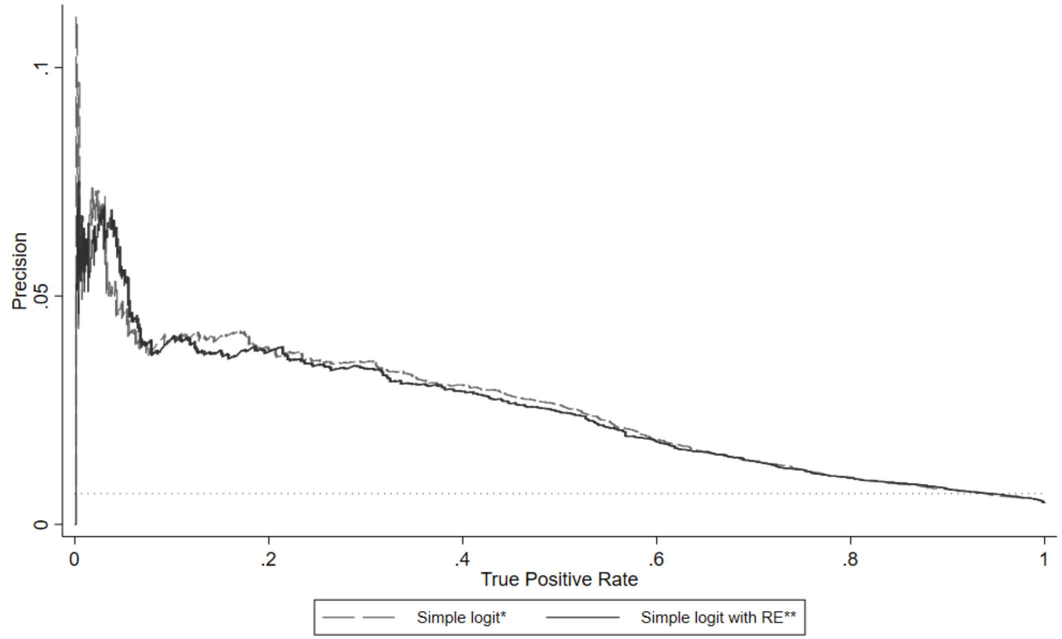
Figure 3.5: Sampled Firm, In-Time PR for the Composite Model



*AUPRC: **0.0275** | **AUPRC: **0.0265**

PRCs for each model specification using our Composite Model inputs and the trailing Realized Rate of Default as the baseline hazard rate

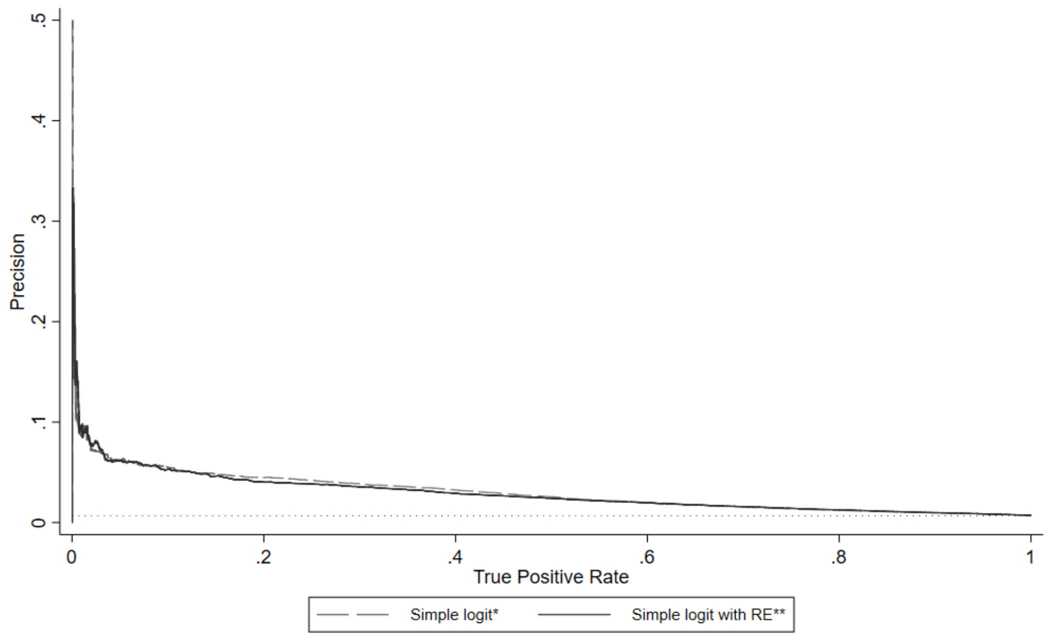
Figure 3.6: Sampled Firm, Out-of-Time PR for the Composite Model



*AUPRC: **0.0252** | **AUPRC: **0.0247**

PRCs for each model specification using our Composite Model inputs and the trailing Realized Rate of Default as the baseline hazard rate

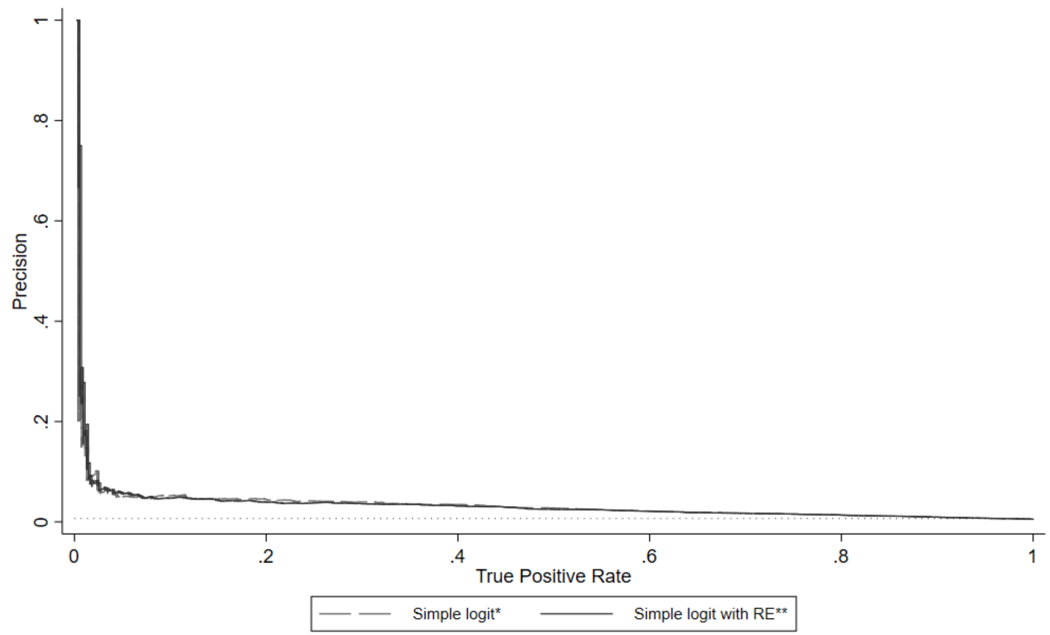
Figure 3.7: Hold-Out, In-Time PR for the Composite Model



*AUPRC: **0.0300** | **AUPRC: **0.0290**

PRCs for each model specification using our Composite Model inputs and the trailing Realized Rate of Default as the baseline hazard rate

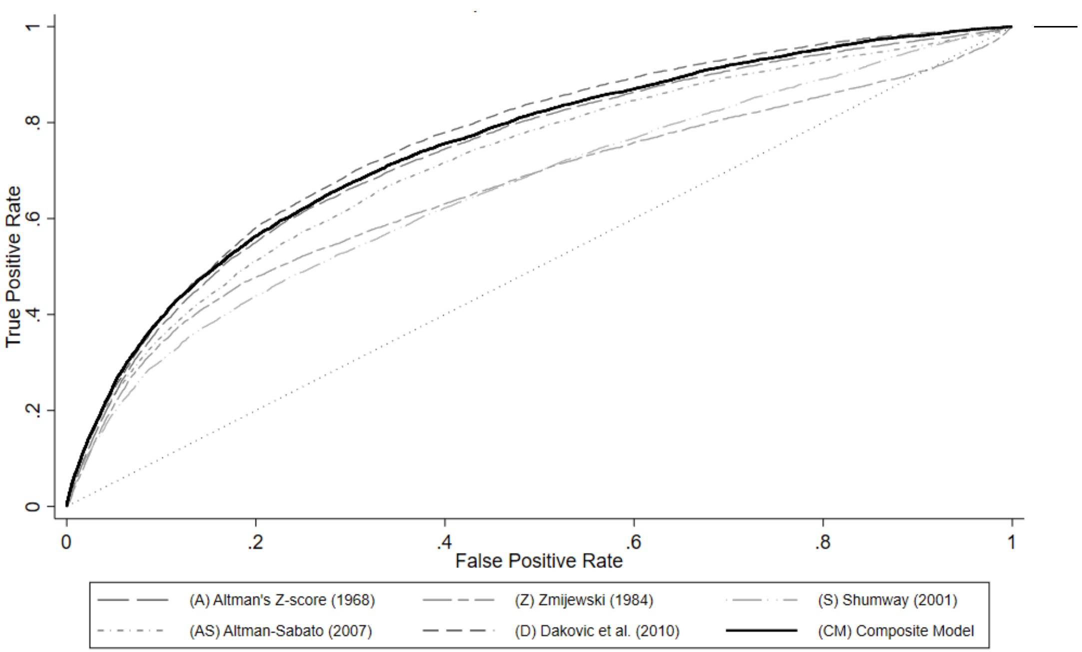
Figure 3.8: Hold-Out, Out-of-Time PR for the Composite Model



*AUPRC: **0.0312** | **AUPRC: **0.0316**

PRCs for each model specification using our Composite Model inputs and the trailing Realized Rate of Default as the baseline hazard rate

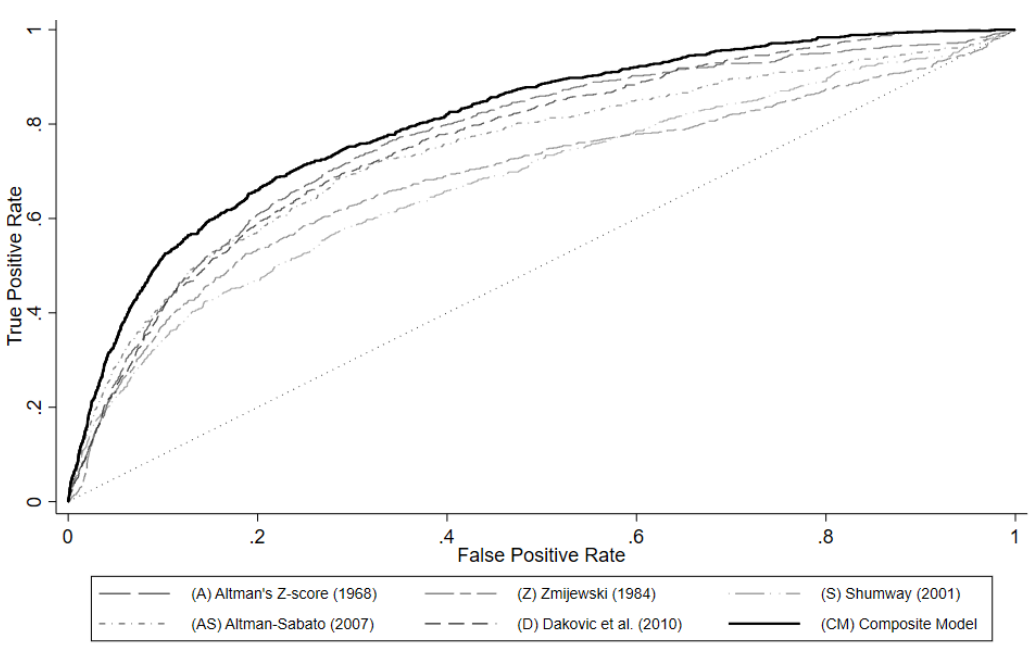
Figure 4.1: Sampled Firm, In-Time ROC Model Comparison



(A) AUROC: **0.7376** | (Z) AUROC: **0.6562** | (S) AUROC: **0.6585** | (AS) AUROC: **0.7172** | (D) AUROC: **0.7612** | (CM) AUROC: **0.7487**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

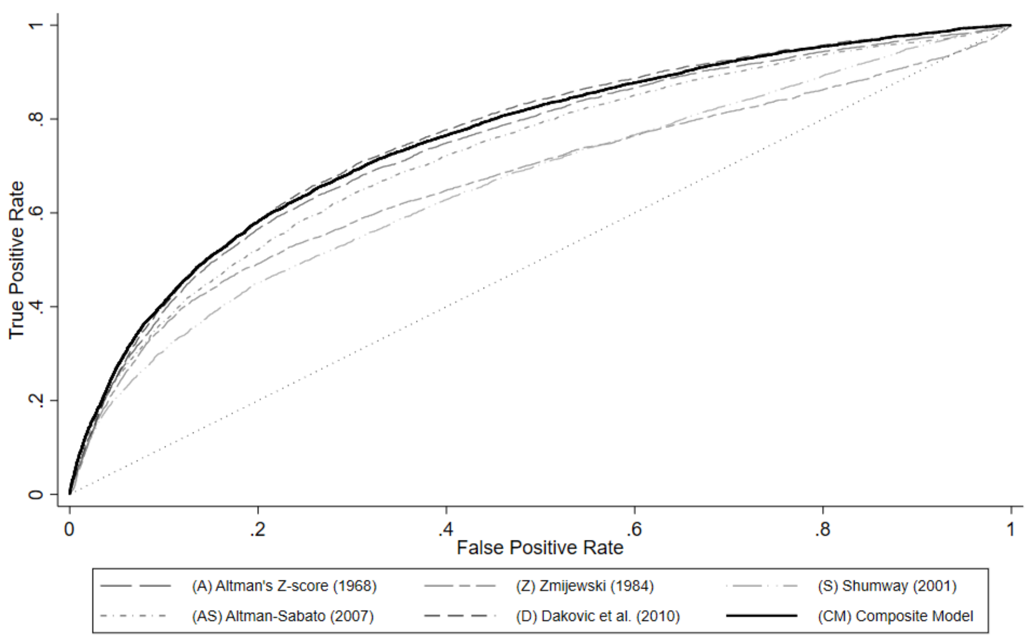
Figure 4.2: Sampled Firm, Out-of-Time ROC Model Comparison



(A) AUROC: **0.7691** | (Z) AUROC: **0.6875** | (S) AUROC: **0.6792** | (AS) AUROC: **0.7430** | (D) AUROC: **0.7639** | (CM) AUROC: **0.8058**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

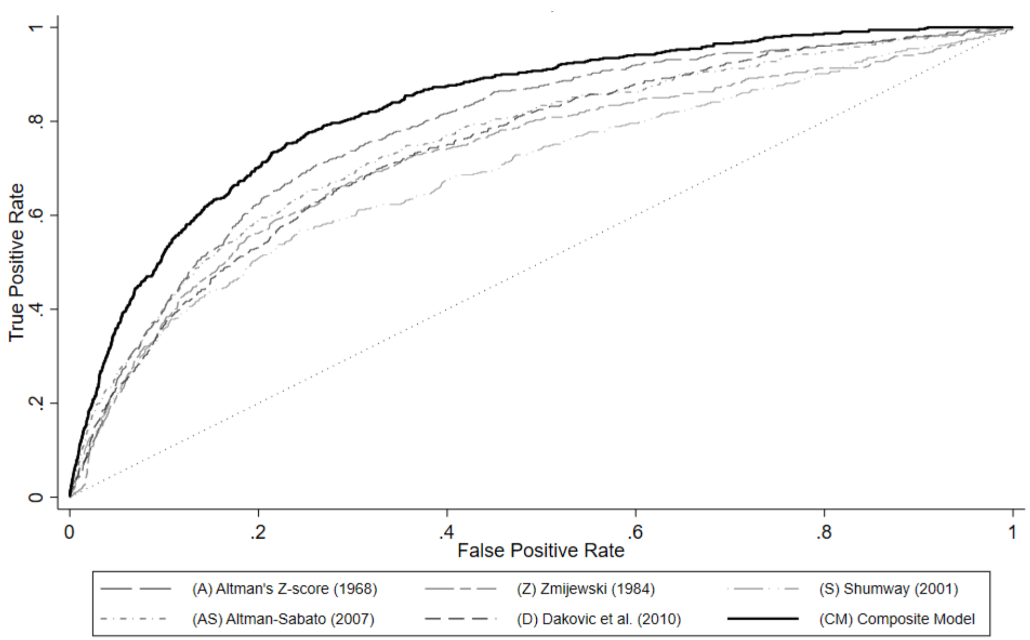
Figure 4.3: Hold-Out, In-Time ROC Model Comparison



(A) AUROC: **0.7424** | (Z) AUROC: **0.6684** | (S) AUROC: **0.6618** | (AS) AUROC: **0.7246** | (D) AUROC: **0.7610** | (CM) AUROC: **0.7570**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

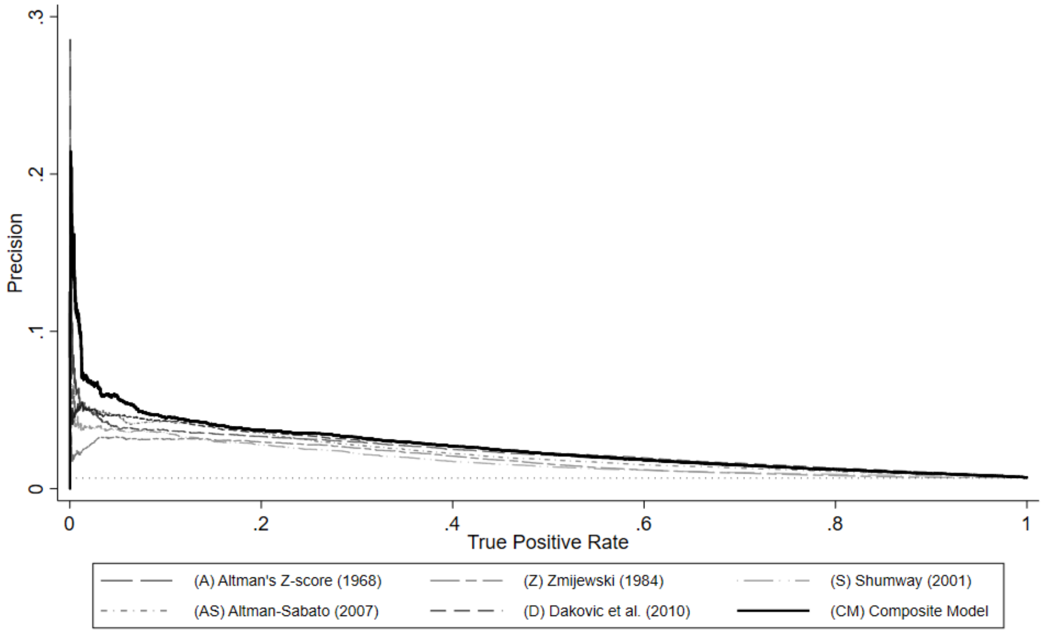
Figure 4.4: Hold-Out, Out-of-Time ROC Model Comparison



(A) AUROC: **0.7805** | (Z) AUROC: **0.7250** | (S) AUROC: **0.6920** | (AS) AUROC: **0.7555** | (D) AUROC: **0.7435** | (CM) AUROC: **0.8297**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

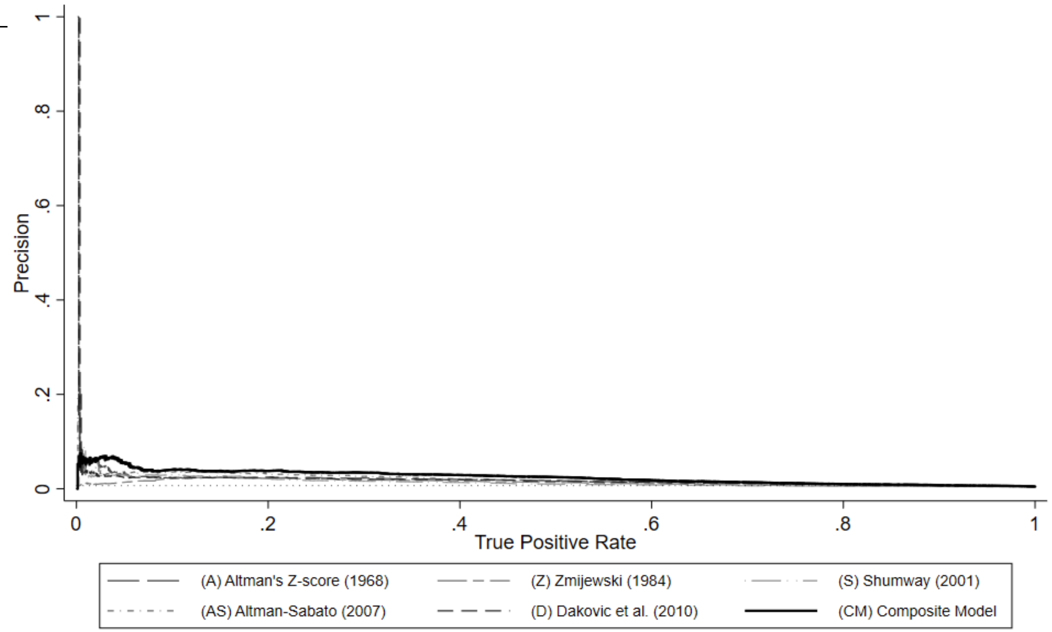
Figure 4.5: Sampled Firm, In-Time PR Model Comparison



(A) AUPRC: **0.0230** | (Z) AUPRC: **0.0179** | (S) AUPRC: **0.0179** | (AS) AUPRC: **0.0224** | (D) AUPRC: **0.0247** | (CM) AUPRC: **0.0265**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

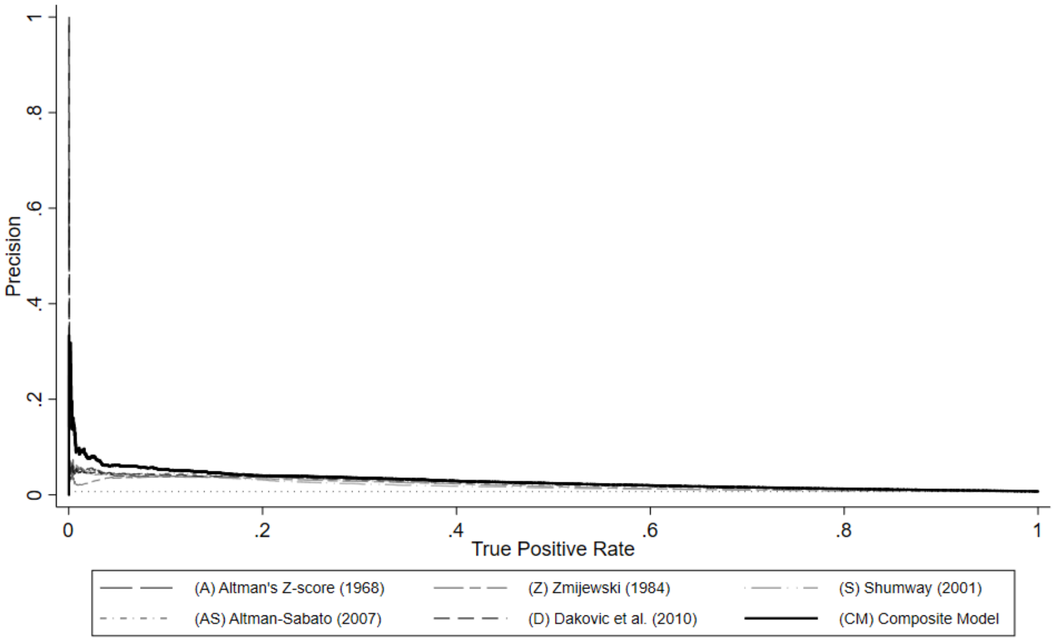
Figure 4.6: Sampled Firm, Out-of-Time PR Model Comparison



(A) AUPRC: **0.0174** | (Z) AUPRC: **0.0129** | (S) AUPRC: **0.0134** | (AS) AUPRC: **0.0190** | (D) AUPRC: **0.0170** | (CM) AUPRC: **0.0247**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

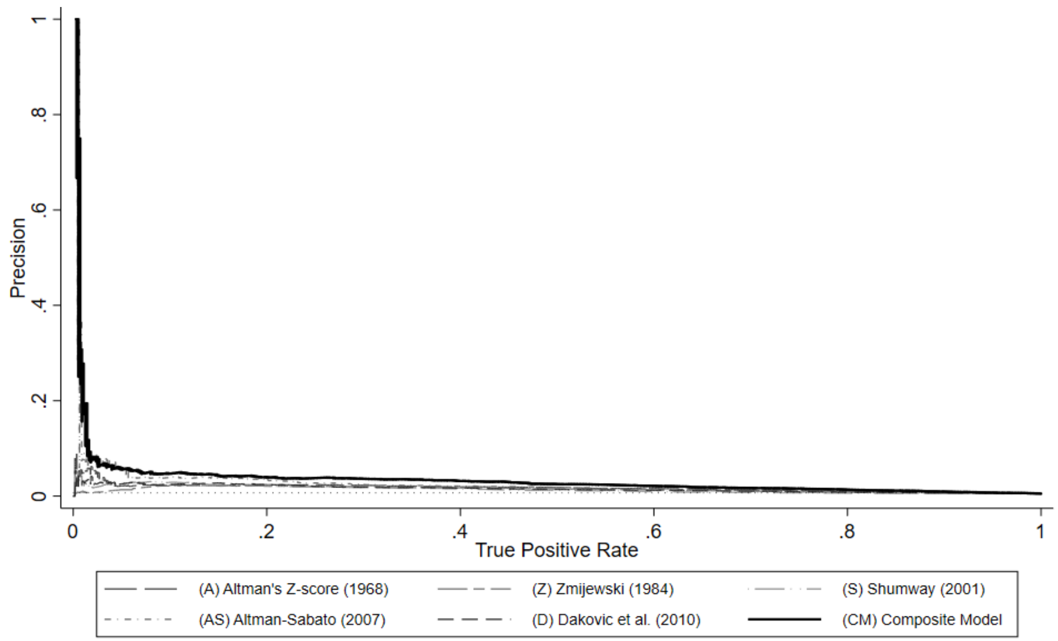
Figure 4.7: Hold-Out, In-Time PR Model Comparison



(A) AUPRC: **0.0239** | (Z) AUPRC: **0.0197** | (S) AUPRC: **0.0193** | (AS) AUPRC: **0.0235** | (D) AUPRC: **0.0252** | (CM) AUPRC: **0.0289**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

Figure 4.8: Hold-Out, Out-of-Time PR Model Comparison



(A) AUPRC: **0.0173** | (Z) AUPRC: **0.0136** | (S) AUPRC: **0.0135** | (AS) AUPRC: **0.0235** | (D) AUPRC: **0.0152** | (CM) AUPRC: **0.0316**

ROCs for each model specification with random effects using trailing Realized Rate of Default as the baseline hazard rate

Table 16: Economic Value of Default (Mis)classification using the Simulated Approach - Existing Clients

	Sampled firms (existing clients) : **2014**						Sampled firms (existing clients) : **2015**						Sampled firms (existing clients) : **2016**						Sampled firms (existing clients) : **2014-2016**					
Bank:	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**
Credits Granted	746	391	11 772	553	39 218	3 812	750	391	391	579	38 795	3 917	737	348	13 111	643	37 989	4 040	549	1 370	37 962	1 992	116 625	11 830
Defaults	18	6	138	6	153	7	14	6	6	9	129	5	7	3	76	5	89	5	16	24	351	23	384	16
Market Share Granted	*1,3%*	*0,7%*	*20,8%*	*1,0%*	*69,4%*	*6,7%*	*1,3%*	*0,7%*	*0,7%*	*1,0%*	*68,0%*	*6,9%*	*1,3%*	*0,6%*	*23,1%*	*1,1%*	*66,8%*	*7,1%*	*0,3%*	*0,8%*	*22,3%*	*1,2%*	*68,5%*	*6,9%*
Default Share	*5,5%*	*1,8%*	*42,1%*	*1,8%*	*46,6%*	*2,1%*	*4,7%*	*1,8%*	*1,8%*	*3,0%*	*42,9%*	*1,7%*	*3,8%*	*1,6%*	*41,1%*	*2,7%*	*48,1%*	*2,7%*	*2,0%*	*2,9%*	*43,1%*	*2,8%*	*47,2%*	*2,0%*
Defaults to Credits	2,4%	1,5%	1,2%	1,1%	0,4%	0,2%	1,9%	1,5%	1,5%	1,6%	0,3%	0,1%	0,9%	0,9%	0,6%	0,8%	0,2%	0,1%	2,9%	1,8%	0,9%	1,2%	0,3%	0,1%
Revenues (SEKk)	13 139	8 881	180 386	8 892	550 952	55 418	10 418	8 881	8 881	7 009	406 646	43 153	9 417	5 586	136 617	7 549	360 131	41 194	12 006	24 671	468 792	26 424	1 326 874	142 091
Losses (SEKk)	4 913	1 638	37 667	1 638	41 761	1 911	3 821	1 638	1 638	2 457	35 210	1 365	1 911	819	20 744	1 365	24 292	1 365	4 367	6 551	95 805	6 278	104 812	4 367
Profits (SEKk)	8 226	7 243	142 720	7 254	509 191	53 507	6 597	7 243	7 243	4 553	371 435	41 788	7 506	4 767	115 873	6 185	335 838	39 829	7 638	18 120	372 987	20 146	1 222 062	137 724
Avg. Rev to Avg. Loss	6,5%	8,0%	5,6%	5,9%	5,2%	5,3%	5,0%	8,0%	8,0%	4,5%	3,8%	4,0%	4,6%	5,7%	3,8%	4,3%	3,5%	3,7%	7,5%	6,4%	4,5%	4,8%	4,2%	4,4%
ROA	*1,6%*	*2,6%*	*1,7%*	*1,9%*	*1,9%*	*2,0%*	*1,3%*	*2,6%*	*2,6%*	*1,1%*	*1,4%*	*1,6%*	*1,5%*	*1,9%*	*1,3%*	*1,4%*	*1,3%*	*1,4%*	*1,9%*	*1,9%*	*1,4%*	*1,5%*	*1,5%*	*1,7%*
RORWA	*3,9%*	*5,2%*	*6,5%*	*5,5%*	*14,7%*	*15,5%*	*3,4%*	*5,2%*	*5,2%*	*3,5%*	*11,4%*	*11,9%*	*4,1%*	*4,5%*	*5,8%*	*4,8%*	*10,9%*	*11,3%*	*3,7%*	*4,1%*	*5,8%*	*4,5%*	*12,2%*	*13,1%*
Revenues Per Credit	17,6	22,7	15,3	16,1	14,0	14,5	13,9	22,7	22,7	12,1	10,5	11,0	12,8	16,1	10,4	11,7	9,5	10,2	21,9	18,0	12,3	13,3	11,4	12,0
Loss Per Default	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9

*The table reports the comparative value of default (mis)classification by showing the amount of credits, amount of credit losses, the market share of said credits and losses, and the profit and loss associated with those credits and losses. Since the simulated approach employs a 5% cut-off, the avaliable credits are slightly higher than the credits granted. These have a high overlap between models. Loans are of equal size, and the hold-out, out-of-time market size based on total loans to credit institutions is SEK 77.7bn (before cut-off). The credit spread is assumed constant at 0.3%. LGD = 40%. default probability floor = 5 basis points. Market share = share of credits granted. Default share = share of defaults. Revenues = Spread charged * Loan value (i.e. where loan value equals the EAD, which also equals assets), if a default does not occur the upcoming year, in which no revenues are assumed. Losses = EAD * LGD. Profits = revenues - losses. For simplicity, the revenue are obtained on the same year as the ending balance sheet, and losses are taken up to one year in advance, meaning if a default occured in 2017, its losses are shown in 2016. ROA = return on assets, RORWA = return on risk-weighted assets. Risk-weighted assets are calculated according to the latest version of the International Convergence of Capital Measurement and Capital Standards document prepared for the Basel Committee of Banking Supervision, see Equation (11)-(15). Loan maturity, M, is one year, consitent with the default probability horizon. A = Altman (1968), Zmijewski (1984), S = Shumway (2001), AS = Altman and Sabato (2007), D = Dakovic et al. (2010), CM = Composite Model.*

Table 17: Economic Value of Default (Mis)classification using the Simulated Approach - New Clients

	Hold-Out (new clients): **2014**						Hold-Out (new clients): **2015**						Hold-Out (new clients): **2016**						Hold-Out (new clients): **2014-2016**					
Bank:	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**
Credits Granted	438	235	7 792	388	26 142	2 564	463	235	235	402	25 935	2 566	478	257	8 659	405	25 344	2 563	1 378	725	24 660	1 196	77 444	7 673
Defaults	9	2	84	1	101	3	11	2	2	3	105	6	6	2	54	1	81	1	26	6	222	5	287	10
Market Share Granted	*1,2%*	*0,6%*	*20,7%*	*1,0%*	*69,6%*	*6,8%*	*1,2%*	*0,6%*	*0,6%*	*1,1%*	*68,6%*	*6,8%*	*1,3%*	*0,7%*	*23,0%*	*1,1%*	*67,2%*	*6,8%*	*1,2%*	*0,6%*	*21,8%*	*1,1%*	*68,5%*	*6,8%*
Default Share	*4,5%*	*1,0%*	*41,8%*	*0,5%*	*50,2%*	*1,5%*	*5,2%*	*1,0%*	*1,0%*	*1,4%*	*49,5%*	*2,8%*	*4,1%*	*1,4%*	*37,2%*	*0,7%*	*55,9%*	*0,7%*	*4,7%*	*1,1%*	*39,8%*	*0,9%*	*51,4%*	*1,8%*
Defaults to Credits	2,1%	0,9%	1,1%	0,3%	0,4%	0,1%	2,4%	0,9%	0,9%	0,7%	0,4%	0,2%	1,3%	0,8%	0,6%	0,2%	0,3%	0,0%	1,9%	0,8%	0,9%	0,4%	0,4%	0,1%
Revenues (SEKk)	8 297	5 420	119 456	6 401	367 883	37 648	6 506	5 420	5 420	5 164	271 451	28 500	6 319	3 931	89 756	4 778	240 223	26 184	21 109	13 455	303 361	16 356	879 853	92 085
Losses (SEKk)	2 457	546	22 928	273	27 568	819	3 002	546	546	819	28 660	1 638	1 638	546	14 739	273	22 109	273	7 097	1 638	60 595	1 365	78 336	2 729
Profits (SEKk)	5 840	4 874	96 529	6 128	340 315	36 829	3 504	4 874	4 874	4 345	242 791	26 862	4 682	3 385	75 017	4 505	218 114	25 911	14 012	11 817	242 766	14 992	801 517	89 356
Avg. Rev to Avg. Loss	6,9%	8,0%	5,6%	6,0%	5,2%	5,4%	5,1%	8,0%	8,0%	4,7%	3,8%	4,1%	4,7%	5,5%	3,7%	4,3%	3,5%	3,7%	5,5%	6,6%	4,5%	5,0%	4,2%	4,4%
ROA	*1,9%*	*2,9%*	*1,8%*	*2,3%*	*1,9%*	*2,1%*	*1,1%*	*2,9%*	*2,9%*	*1,6%*	*1,4%*	*1,5%*	*1,4%*	*1,9%*	*1,2%*	*1,6%*	*1,3%*	*1,5%*	*1,4%*	*2,3%*	*1,4%*	*1,8%*	*1,5%*	*1,7%*
RORWA	*4,5%*	*6,2%*	*6,7%*	*6,6%*	*14,7%*	*15,9%*	*2,9%*	*6,2%*	*6,2%*	*4,7%*	*11,1%*	*11,3%*	*4,0%*	*4,3%*	*5,7%*	*5,6%*	*10,5%*	*11,8%*	*3,8%*	*5,0%*	*5,9%*	*5,6%*	*12,2%*	*13,0%*
Revenues Per Credit	18,9	23,1	15,3	16,5	14,1	14,7	14,1	23,1	23,1	12,8	10,5	11,1	13,2	15,3	10,4	11,8	9,5	10,2	15,3	18,6	12,3	13,7	11,4	12,0
Loss Per Default	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9	272,9

*The table reports the comparative value of default (mis)classification by showing the amount of credits, amount of credit losses, the market share of said credits and losses, and the profit and loss associated with those credits and losses. Since the simulated approach employs a 5% cut-off, the avaliable credits is slightly higher than the credits granted. These have a high overlap between models. Loans are of equal size, and the hold-out, out-of-time market size based on total loans to credit institutions is SEK 77.7bn (before cut-off). The credit spread is assumed constant at 0.3%. LGD = 40%. Default probability floor = 5 basis points. Market share = share of credits granted. Default share = share of defaults. Revenues = Spread charged * Loan value (where the loan value equals the EAD, which also equals assets), if a default does not occur the upcoming year, in which no revenues are assumed. Losses = EAD * LGD. Profits = revenues - losses. ROA = return on assets, RORWA = return on risk-weighted assets. Risk-weighted assets are calculated according to the latest version of the International Convergence of Capital Measurement and Capital Standards document prepared for the Basel Committee of Banking Supervision, see Equation (11)-(15). Loan maturity, M, is one year, consitent with the default probability horizon. For simplicity, the revenue are obtained on the same year as the ending balance sheet, and losses are taken up to one year in advance, meaning if a default occured in 2017, its losses are shown in 2016. A = Altman (1968), Zmijewski (1984), S = Shumway (2001), AS = Altman and Sabato (2007), D = Dakovic et al. (2010), CM = Composite Model.*

Table 18: Calibration and Power, Estimated on New and Existing Customers, 2014-2016

	One Year Prediction														
	Calibration						AUROC				AUPRC				
	Failures	Non Failures	Total	Default rate	Actual	Δ	1†††	2***	3†††	4***	1	2	3	4
Altman (1968)	2 355	282 920	285 275	0,8%	2,1%	-1,3%	-	76,9%	-	78,1%	-	1,7%	-	1,7%
Zmijewski (1984)	2 665	282 468	285 133	0,9%	2,1%	-1,2%	-	68,8%	-	72,5%	-	1,3%	-	1,4%
Shumway (2001)	322	240 028	240 350	0,1%	2,1%	-2,0%	-	67,9%	-	69,2%	-	1,3%	-	1,4%
Altman and Sabato (2007)	1 541	217 536	219 077	0,7%	2,1%	-1,4%	-	74,3%	-	75,6%	-	1,9%	-	2,4%
Dakovic et al. (2010)	1 501	283 732	285 233	0,5%	2,1%	-1,6%	-	76,4%	-	74,4%	-	1,7%	-	1,5%
Composite Model	**2 665**	**281 747**	**284 412**	**0,9%**	**2,1%**	**-1,2%**	-	**80,6%**	-	**83,0%**	-	**2,5%**	-	**3,2%**

This table shows a portion of the Summary Table of Results (see Table 13) with the the estimated number of failures and non-failures by model only for the out-of-time period (2014-2016). That is, estimated default rates multiplied by the nr of non-missing observations for sampled and hold-out firms, out-of-time. These can be compared to the actual default rate, to gauge the (mis)calibration of the Skogsvik adjusted estimated probabilites of default.

Table 19: Economic Value of Default (Mis)classification using Notional Credit Exposures - Existing Clients

	Sampled firms (existing clients) : **2014**						Sampled firms (existing clients) : **2015**						Sampled firms (existing clients) : **2016**						In sample (existing clients) : **2014-2016**						
Bank:	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**Aggregated financials**
Credits	160	38	4 373	128	16 593	1 131	179	40	4 372	125	15 733	1 216	184	26	4 535	137	15 322	1 300	523	104	13 280	390	47 648	3 647	65 592
Defaults	6	0	56	1	74	1	6	0	54	1	55	0	2	0	34	0	48	0	14	0	144	2	177	1	338
Market Share	*0,7%*	*0,2%*	*19,5%*	*0,6%*	*74,0%*	*5,0%*	*0,8%*	*0,2%*	*20,2%*	*0,6%*	*72,6%*	*5,6%*	*0,9%*	*0,1%*	*21,1%*	*0,6%*	*71,3%*	*6,0%*	*0,8%*	*0,2%*	*20,2%*	*0,6%*	*72,6%*	*5,6%*	
Default Share	*4,3%*	*0,0%*	*40,6%*	*0,7%*	*53,6%*	*0,7%*	*5,2%*	*0,0%*	*46,6%*	*0,9%*	*47,4%*	*0,0%*	*2,4%*	*0,0%*	*40,5%*	*0,0%*	*57,1%*	*0,0%*	*4,1%*	*0,0%*	*42,6%*	*0,6%*	*52,4%*	*0,3%*	
Defaults to Credits	3,8%	0,0%	1,3%	0,8%	0,4%	0,1%	3,4%	0,0%	1,2%	0,8%	0,3%	0,0%	1,1%	0,0%	0,7%	0,0%	0,3%	0,0%	2,7%	0,0%	1,1%	0,5%	0,4%	0,0%	0,5%
Revenues (SEKk)	2 772	386	247 095	607	497 953	100 248	2 490	152	153 304	743	375 365	77 060	2 017	129	132 741	767	304 156	96 094	7 279	667	533 140	2 118	1 177 474	273 401	**6 409 094**
Losses (SEKk)	1 033	0	24 632	10	23 815	34	1 142	0	28 499	40	17 735	0	119	0	16 240	0	47 533	0	2 294	0	69 370	50	89 083	34	160 832
Profits (SEKk)	1 739	386	222 463	597	474 138	100 214	1 348	152	124 805	703	357 630	77 060	1 898	129	116 502	767	256 623	96 094	4 985	667	463 770	2 067	1 088 391	273 368	6 248 262
Avg. Rev to Avg. Loss	10,5%		13,0%	47,8%	9,4%	264,0%	7,6%		6,7%	14,8%	7,4%		18,6%		6,2%		2,0%		8,7%		8,4%	21,7%	4,9%	223,2%	20,5%
ROA	*1,5%*	*2,8%*	*1,8%*	*2,4%*	*1,9%*	*2,0%*	*1,0%*	*2,9%*	*1,2%*	*1,6%*	*1,4%*	*1,5%*	*1,6%*	*2,8%*	*1,2%*	*1,6%*	*1,1%*	*1,4%*	*1,4%*	*2,8%*	*1,5%*	*1,8%*	*1,5%*	*1,6%*	*5,2%*
RORWA	*3,1%*	*4,9%*	*14,2%*	*5,7%*	*20,0%*	*22,3%*	*2,2%*	*5,0%*	*9,7%*	*4,4%*	*15,2%*	*17,9%*	*4,2%*	*6,2%*	*9,4%*	*4,8%*	*12,2%*	*15,8%*	*3,1%*	*5,1%*	*11,3%*	*4,9%*	*15,9%*	*18,4%*	
Revenues Per Credit	17,3	10,2	56,5	4,7	30,0	88,6	13,9	3,8	35,1	5,9	23,9	63,4	11,0	5,0	29,3	5,6	19,9	73,9	13,9	6,4	40,1	5,4	24,7	75,0	97,7
Loss Per Default	172,1	-	439,9	10,0	321,8	33,6	190,4	-	527,8	40,4	322,5	-	59,6	-	477,6	-	990,3	-	163,9	-	481,7	25,2	503,3	33,6	475,8

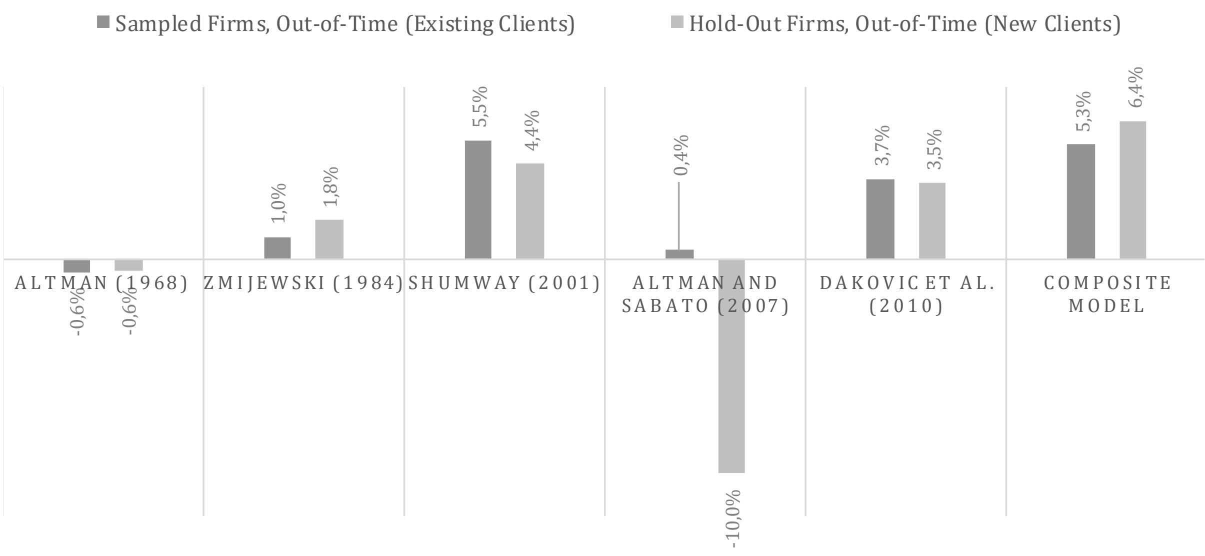
*The table reports the comparative value of default (mis)classification in a competitive environment for existing bank clients by showing the amount of credits, amount of credit losses, the market share of said credits and losses, and the profit and loss associated with those credits and losses. The aggregated financials column to the right includes the revenue as calculted using all external interest cost for the sampled firms, out-of-time. Note that it is approximately three to four times higher than the model-estimated interest, translating into a higher ROA that a compeitive environment would suggest. A = Altman (1968), Zmijewski (1984), S = Shumway (2001), AS = Altman and Sabato (2007), D = Dakovic et al. (2010), CM = Composite Model. The credit spread for the most creditworthy customer, k, varies with the 5th percentile in the charged bankrate in each year (2.0% 2014, 1.48% in 2015 and 1.33% in 2016). LGD = 40%. Default probability floor = 5 basis points. Banks do not reject any firms since all rejected firms are already excluded from this notional credit exposure approach. Market share = share of credits granted. Default share = share of defaults. Revenues = spread charged * loan value (where the loan value equals the EAD, which also equals assets), if default does not occur the coming 12 months. Losses = EAD * LGD. Profits = revenues - losses. For simplicity, the revenues are obtained on the same year as the ending balance sheet, and losses are taken up to one year in advance, meaning if a default occured in 2017, its losses are shown in 2016. ROA = return on assets, RORWA = return on risk-weighted assets. Risk-weighted assets are calculated according to the latest version of the International Convergence of Capital Measurement and Capital Standards document prepared for the Basel Committee of Banking Supervision, see Equation (11)-(15). Loan maturity, M, is one year, consitent with the default probability horizon.*

Table 20: Economic Value of Default (Mis)classification using Notional Credit Exposures - New Clients

	Hold-Out (new clients): **2014**						Hold-Out (new clients): **2015**						Hold-Out (new clients): **2016**						Hold-Out (new clients) : **2014-2016**						
Bank:	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**A**	**Z**	**S**	**AS**	**D**	**CM**	**Aggregated financials**
Credits	104	22	2 833	97	11 038	725	120	20	2 773	87	10 518	767	102	24	2 921	69	10 226	782	326	66	8 528	253	31 768	2 287	43 228
Defaults	4	0	34	0	47	0	3	0	28	2	52	0	0	0	23	0	47	0	7	0	85	2	146	0	240
Market Share	*0,7%*	*0,1%*	*19,1%*	*0,7%*	*74,5%*	*4,9%*	*0,8%*	*0,1%*	*19,4%*	*0,6%*	*73,6%*	*5,4%*	*0,7%*	*0,2%*	*20,7%*	*0,5%*	*72,4%*	*5,5%*	*0,8%*	*0,2%*	*19,7%*	*0,6%*	*73,5%*	*5,3%*	
Default Share	*4,7%*	*0,0%*	*40,0%*	*0,0%*	*55,3%*	*0,0%*	*3,5%*	*0,0%*	*32,9%*	*2,4%*	*61,2%*	*0,0%*	*0,0%*	*0,0%*	*32,9%*	*0,0%*	*67,1%*	*0,0%*	*2,9%*	*0,0%*	*35,4%*	*0,8%*	*60,8%*	*0,0%*	
Defaults to Credits	3,8%	0,0%	1,2%	0,0%	0,4%	0,0%	2,5%	0,0%	1,0%	2,3%	0,5%	0,0%	0,0%	0,0%	0,8%	0,0%	0,5%	0,0%	2,1%	0,0%	1,0%	0,8%	0,5%	0,0%	0,6%
Revenues (SEKk)	1 221	274	113 360	468	312 111	49 691	1 494	220	74 691	283	225 019	42 194	1 434	91	78 322	396	197 321	41 169	4 148	586	258 793	1 146	741 694	133 392	**3 943 807**
Losses (SEKk)	399	0	7 714	0	28 774	0	607	0	13 548	2 212	15 302	0	0	0	7 535	0	14 812	0	1 006	0	28 798	2 212	58 888	0	90 903
Profits (SEKk)	822	274	105 646	468	283 337	49 691	887	220	61 142	-1 929	209 718	42 194	1 434	91	70 787	396	182 509	41 169	3 142	586	229 996	-1 065	682 807	133 392	3 852 904
Avg. Rev to Avg. Loss	12,2%		17,9%	0,0%	4,6%		6,3%		5,6%	0,3%	7,3%				8,2%		6,2%		9,0%		9,0%	0,4%	5,8%		24,1%
ROA	*1,7%*	*6,2%*	*1,9%*	*2,2%*	*1,8%*	*2,0%*	*1,2%*	*6,6%*	*1,3%*	*-9,2%*	*1,4%*	*1,5%*	*1,8%*	*2,5%*	*1,2%*	*1,7%*	*1,3%*	*1,4%*	*1,6%*	*5,1%*	*1,5%*	*-1,6%*	*1,5%*	*1,6%*	*5,5%*
RORWA	*3,5%*	*7,2%*	*12,8%*	*6,8%*	*18,4%*	*24,0%*	*2,4%*	*7,5%*	*8,7%*	*-24,9%*	*14,7%*	*18,2%*	*4,0%*	*5,0%*	*9,6%*	*4,0%*	*13,5%*	*16,6%*	*3,2%*	*6,8%*	*10,3%*	*-4,3%*	*15,7%*	*19,4%*	
Revenues Per Credit	11,7	12,4	40,0	4,8	28,3	68,5	12,5	11,0	26,9	3,2	21,4	55,0	14,1	3,8	26,8	5,7	19,3	52,6	12,7	8,9	30,3	4,5	23,3	58,3	91,2
Loss Per Default	99,7	-	226,9	-	612,2	-	202,4	-	483,9	1105,8	294,3	-	-	-	327,6	-	315,2	-	143,7	-	338,8	1105,8	403,3	-	378,8

*The table reports the comparative value of default (mis)classification in a competitive environment for new bank clients by showing the amount of credits, amount of credit losses, the market share of said credits and losses, and the profit and loss associated with those credits and losses. The aggregated financials column to the right includes the revenue as calculted using all external interest cost for the companies in the hold-out, out-of-time population. Note that it is approximately three to four times higher than the model-estimated interest, translating into a higher ROA that a compeitive environment would suggest. A = Altman (1968), Zmijewski (1984), S = Shumway (2001), AS = Altman and Sabato (2007), D = Dakovic et al. (2010), CM = Composite Model. The credit spread for the most creditworthy customer, k, varies with the 5th percentile in the charged bankrate in each year (2.0% 2014, 1.48% in 2015 and 1.33% in 2016). LGD = 40%. Default probability floor = 5 basis points. Banks do not reject any firms since all rejected firms are already excluded from this notional credit exposure approach. Market share = share of credits granted. Default share = share of defaults. Revenues = spread charged * loan value (i.e. EAD, assets), if default does not occur the coming 12 months. Losses = EAD * LGD. Profits = revenues - losses. For simplicity, the revenues are obtained on the same year as the ending balance sheet, and losses are taken up to one year in advance, meaning if a default occured in 2017, its losses are shown in 2016. ROA = return on assets, RORWA = return on risk-weighted assets. Risk-weighted assets are calculated according to the latest version of the International Convergence of Capital Measurement and Capital Standards document prepared for the Basel Committee of Banking Supervision, see Equation (11)-(15). Loan maturity, M, is one year, consitent with the default probability horizon.*

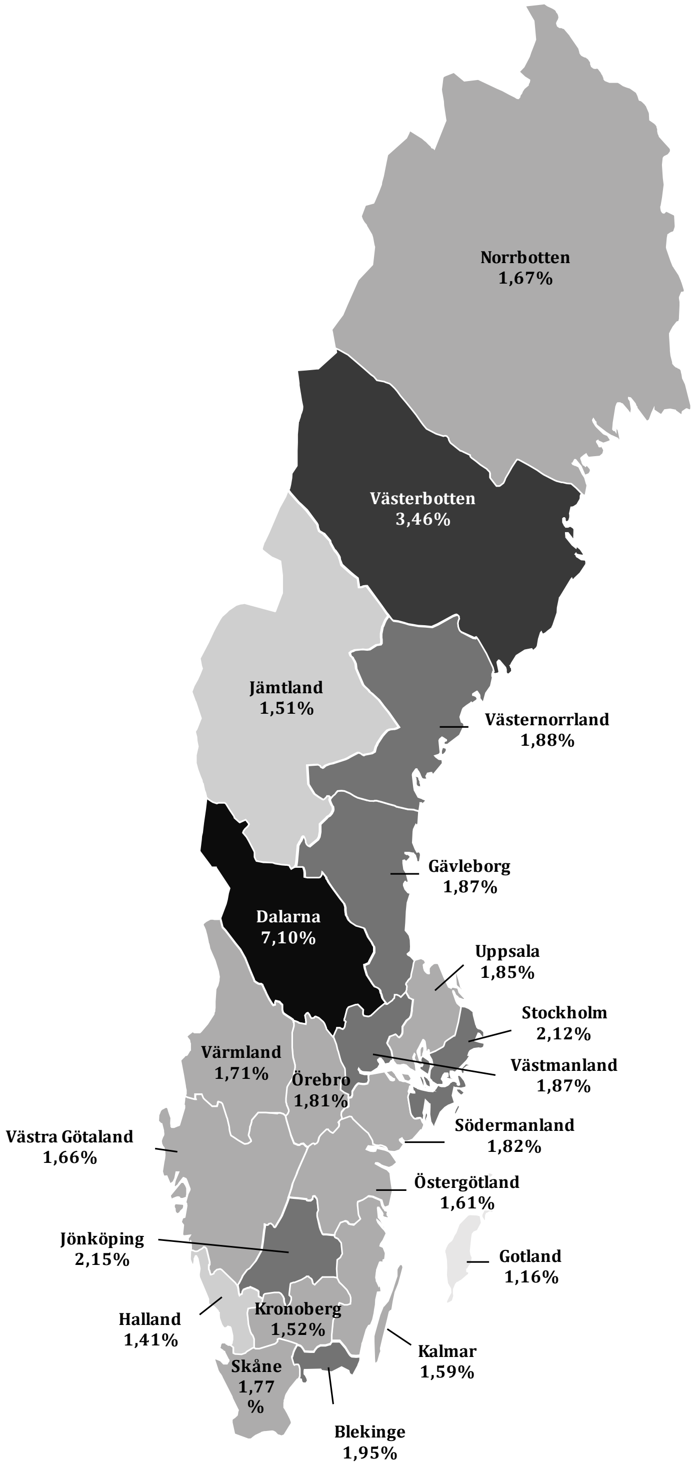
Figure 5: Difference in RORWA between the Notional Credit Exposure Method and the Simulated Approach



The table reports the benefit of using notional credit exposures (with a credit spread for the most creditworthy customer that varies by year), as compared to the simulated approach, when estimating the economic value of (mis)classification. The results are presented in percentage points, for existing and new clients respectively.

9. Appendix

Figure A1: Average Default Rates by County, 1998-2017



This figure shows the average default rate by Swedish county during the period 1998-2017. The number of observations range from 735,284 in Stockholm, to 32,072 in Blekinge.

Table A1: Proportional Hazard Test, Pseudo R-squared Rank and Average Marginal Effect Rank of Investigated Variables

Variable					Survival Analysis Statistics with Frailty						Average Marginal Effect, Simple Logit with RE			
Category	Sub-Category	Variable	Variable Description	Example Paper	Con.	Estimate (Coeff)	Pseudo R2	PH: Chi Sq stat.	Pseudo R2 rank	Pseudo R2 rank ex PH	Con.	Estimat e AME	P-Value AME	Rank AME
Financial Ratios	Acc. Qual	IAC_S	Items Affecting Comparab. / Sales	-	1	13,860	0,040%	2,81		133	1	0,50%	0,00	3
Financial Ratios	Acc. Qual	IAC_EBIT	Items Affecting Comparab. / EBIT	-	1	0,827	0,020%	0,03	46	156	1	0,03%	0,00	54
Financial Ratios	Acc. Qual	EXTRACS	Extraordinary Costs / Sales	-	1	0,017	0,000%	0,00	89	234	1	0,00%	0,98	
Financial Ratios	Acc. Qual	EXTRACNI	Extraordinary Costs / Net Income	-	1	0,000	0,000%	0,00	90	235	1	0,00%	.	
Financial Ratios	Acc. Qual	EXTRACTL	Extraordinary Costs / Total Liabilities	-	1	0,068	0,000%	0,00	80	219	1	0,00%	0,07	118
Financial Ratios	Acc. Qual	EXTRACTA	Extraordinary Costs / Total Assets	-	1	0,152	0,001%	0,02	77	215	1	0,01%	0,08	87
Financial Ratios	Acc. Qual	EXTRACEMP	Extraordinary Costs / # Employees	-	1	0,001	0,000%	0,03	83	224	1	0,00%	0,25	
Financial Ratios	Acc. Qual	EXTRACD	Extraordinary Costs / Dividend	-	1	-0,245	0,000%	0,00	79	218	1	0,00%	0,82	
Financial Ratios	Activity	CCC	DIO + DSO - DPO	-	1	-0,002	0,035%	18,05		136	1	0,00%	0,00	155
Financial Ratios	Activity	DIO	(Inventories/Sales)*365	Pederzoli & Torricelli (2010)	1	0,002	0,031%	33,36		142	1	0,00%	0,06	162
Financial Ratios	Activity	DSO	(Trade Receivables/Sales)*365	Karels & Prakash (1987)	1	-0,001	0,003%	15,64		203	1	0,00%	0,00	152
Financial Ratios	Activity	DPO	(Trade Payables/Sales)*365	Altman & Sabato (2007)	1	0,008	0,423%	11,89		45	1	0,00%	0,00	151
Financial Ratios	Activity	IC_turn	Sales / Invested Capital	-	1	-0,001	0,003%	0,81	74	201	1	0,00%	0,03	163
Financial Ratios	Activity	CE_turn	Sales / Capital Employed	Zavgren & Friedman (1988)	1	0,011	0,054%	1,28	35	126	1	0,00%	0,00	145
Financial Ratios	Activity	A_turn	Sales / Assets	Altman (1968)	1	0,073	0,102%	2,86		100	1	0,01%	0,00	109
Financial Ratios	Activity	FA_turn	Sales / Fixed Assets	-	1	0,000	0,004%	1,08	69	192	1	0,00%	0,01	166
Financial Ratios	Activity	A_WC_S	Avg. Working Capital / Sales	Karels & Prakash (1987)	1	0,172	0,034%	49,18		139	1	0,00%	0,01	115
Financial Ratios	Activity	A_OP_WC_S	Avg. Operating Working Capital / Sales	-	1	-0,727	0,115%	0,41	24	97	1	-0,02%	0,00	62
Financial Ratios	Activity	A_Tr_WC_S	Avg. Trade Working Capital / Sales	-	1	-0,638	0,035%	18,05		136	1	-0,05%	0,00	38
Financial Ratios	Activity	ln_CCC	Ln (DIO + DSO - DPO)	-	1	-0,047	0,014%	0,00	54	165	1	0,00%	0,00	123
Financial Ratios	Activity	ln_DIO	Ln ((Inventories/Sales)*365)	Altman & Sabato (2007)	1	0,050	0,026%	11,88		149	1	0,00%	0,11	
Financial Ratios	Activity	ln_DSO	Ln ((Trade Receivables/Sales)*365)	Karels & Prakash (1987)	1	-0,069	0,050%	3,05		128	1	0,00%	0,00	111
Financial Ratios	Activity	ln_DPO	Ln ((Trade Payables/Sales)*365)	Zavgren & Friedman (1988)	1	0,213	0,233%	40,85		63	1	0,01%	0,00	104
Financial Ratios	Activity	A_AR_INV	Avg. Accounts Receivables / Inv.	Zavgren & Friedman (1988)	1	-0,006	0,035%	4,26		138				
Financial Ratios	Activity	ln_A_AR_INV	ln (Avg. Accounts Receivables / Inv.)	Zavgren & Friedman (1988)	1	-0,067	0,113%	8,58		98	1	0,00%	0,00	121
Financial Ratios	Activity	Depr_S	Depreciation & Amortization / Sales	-	1	-1,434	0,023%	36,30		153	1	-0,05%	0,04	42
Financial Ratios	CFM	OP_CF_EBITDA	Operating Cash Flow / EBITDA	-	1	0,007	0,002%	45,06		206	1	0,00%	0,16	
Financial Ratios	CFM	OP_CF_EBIT	Operating Cash Flow / EBIT	-	1	0,004	0,001%	4,05		212	1	0,00%	0,22	
Financial Ratios	CFM	OP_CF_NI	Operating Cash Flow / net Income	-	1	-0,001	0,001%	1,59	78	216	1	0,00%	0,63	
Financial Ratios	CFM	FCF_EBITDA	Free Cash Flow / EBITDA	-	1	0,026	0,038%	28,68		134	1	0,00%	0,00	138
Financial Ratios	CFM	FCF_EBIT	Free Cash Flow / EBIT	-	1	0,008	0,009%	1,41	58	174	1	0,00%	0,02	148
Financial Ratios	CFM	FCF_NI	Free Cash Flow / Net Income	-	1	0,001	0,005%	6,49		191	1	0,00%	0,02	160
Financial Ratios	CFM	ACC_FCF_A	(Net Income + D&A) / Assets	Altman et al. (2016)	1	-2,997	1,580%	0,78	6	9	1	-0,15%	0,00	14
Financial Ratios	CFM	ACC_FCF_CE	(Net Income + D&A) / Capital Employed	-	1	-0,743	0,589%	12,63		38	1	-0,03%	0,00	52
Financial Ratios	CFM	FCF_A	Free Cash Flow / Assets	*Altman et al. (2016)*	1	0,193	0,008%	469,41		180	1	0,01%	0,16	
Financial Ratios	CFM	FCF_CE	Free Cash Flow / Capital Employed	-	1	0,039	0,002%	254,45		208	1	0,00%	0,16	
Financial Ratios	Changes	D_CCC	Change in CCC (DIO + DSO - DPO)	-	1	-0,002	0,008%	0,24	63	181	1	0,00%	0,30	
Financial Ratios	Changes	D_NWC_bloat	Change in DIO + DSO + DPO	-	1	0,006	0,166%	0,67	20	78				
Financial Ratios	Changes	D_OP_WC	Change in Operating Working Capital	-	1	0,000	0,199%	0,02	19	70	1	0,00%	0,00	165
Financial Ratios	Changes	D_FA	Change in Fixed Assets	-	1	0,000	0,009%	0,81	60	176				
Financial Ratios	Changes	OP_CF_A	Oper. Cash Flow / Assets	Bellovary et al (2007)*	1	-1,416	0,386%	163,09		50	1	-0,06%	0,00	32
Financial Ratios	Changes	OP_CF_CE	Oper. Cash Flow / Capital Employed	-	1	-0,364	0,196%	32,87		72	1	-0,01%	0,00	74
Financial Ratios	Changes	D_STDA	Change in (Short Term Debt / Assets)	Altman et al. (2016)	1	0,129	0,000%	0,78	84	225	1	0,00%	0,80	
Financial Ratios	Changes	D_IBDA	(Change in IBD) / Average Assets	Altman et al. (2016)	1	-3,756	1,170%	66,86		15	1	-0,10%	0,00	24
Financial Ratios	Changes	D_ValueAddPerE	Change in (EBIT + Salaries)/Employees	-	1	0,000	0,051%	53,19		127				
Financial Ratios	Changes	GR_ValueAddPerE	% GR in (EBIT + Salaries)/Employees	-	1	-0,824	0,518%	44,08		43				
Financial Ratios	Changes	GR_ROACE	% GR in ROACE	-	1	-0,022	0,046%	5,00		129				
Financial Ratios	Coverage	NDEBITDA	Net Debt / EBITDA	-	1	-0,003	0,005%	6,20		190	1	0,00%	0,23	
Financial Ratios	Coverage	NDEBIT	Net Debt / EBIT	-	1	-0,003	0,018%	0,05	48	158				
Financial Ratios	Coverage	ICR	EBIE / Interest Expense	-	1	-0,004	0,323%	118,65		57	1	0,00%	0,00	158
Financial Ratios	Coverage	ICR_EBITDA	EBITDA / Interest Expense	Altman & Sabato (2007)	1	-0,004	0,423%	100,32		46	1	0,00%	0,00	153
Financial Ratios	Coverage	DEBITDA	Debt / EBITDA	Altman et al. (2016)	1	-0,028	0,177%	6,03		75	1	0,00%	0,00	135
Financial Ratios	Coverage	IC_EBIE	Interest Cost / EBIE	-	1	-0,192	0,129%	0,96	21	91	1	-0,01%	0,00	105
Financial Ratios	Coverage	IC_EBITDA	Interest Cost / EBITDA	Altman & Sabato (2007)	1	-0,280	0,096%	2,67	25	101	1	-0,01%	0,00	89
Financial Ratios	Dividends	DIVTA	Dividend / Total Assets	-	1	-1,511	0,027%	157,16		145	1	0,01%	0,72	
Financial Ratios	Dividends	DIVTL	Dividend / Total Liabilities	-	1	0,190	0,003%	108,74		202	1	0,03%	0,00	48
Financial Ratios	Dividends	DIVLTL	Dividend / Long Term Liabilities	-	1	0,000	0,000%	35,77		217	1	0,00%	0,19	
Financial Ratios	Dividends	DIVCL	Dividend / Current Liabilities	-	1	-0,014	0,000%	107,25		228	1	0,02%	0,00	66
Financial Ratios	Dividends	DIVIBD	Dividend / Interest Bearing Debt	-	1	0,007	0,006%	41,58		185	1	0,00%	0,01	146
Financial Ratios	Dividends	DIVCASH	Dividend / Cash Balance	-	1	-0,400	0,150%	31,39		85	1	-0,01%	0,00	85
Financial Ratios	Dividends	DIVS	Dividend / Sales	-	1	-1,238	0,010%	184,92		171	1	0,02%	0,41	
Financial Ratios	Dividends	DIVNI	Payout Ratio	McKee (2003)	1	-0,304	0,155%	166,83		83	1	-0,01%	0,00	98
Financial Ratios	Dividends	DIVIE	Dividend / Interest Expense	-	1	-0,004	0,085%	25,62		114	1	0,00%	0,00	154
Financial Ratios	Dividends	DIVEMP	Dividend / # Employee	-	1	-0,008	0,202%	196,90		69	1	0,00%	0,00	150
Financial Ratios	Dividends	DIVD	Dividend Dummy	Dakovic et al. (2010)	1	-0,936	0,819%	11,76		29	1	0,04%	0,00	46
Financial Ratios	Efficiency	ValueAddPerE	(EBIT + Salary Cost)/Employees	-	1	-0,003	1,987%	0,03	4	5	1	0,00%	0,00	159
Financial Ratios	Growth	GR_A	% GR. in Assets	Altman et al. (2016)	1	-2,658	2,097%	215,03		4				
Financial Ratios	Growth	GR_FA	% GR. In Fixed Assets	-	1	-0,020	0,031%	6,97		140				
Financial Ratios	Growth	GR_E	% GR. in Equity	-	1	-0,470	0,630%	0,00	13	36				
Financial Ratios	Growth	GR_WC	% GR. in Working Capital	-	1	-0,039	0,026%	0,48	43	148	1	0,00%	0,00	140
Financial Ratios	Growth	GR_OP_WC	% GR. in Operating Working Capital	-	1	-0,011	0,007%	9,44		182	1	0,00%	0,14	
Financial Ratios	Growth	GR_Tr_WC	% GR. in Trade Working Capital	-	1	-0,074	0,144%	12,87		88	1	0,00%	0,00	126
Financial Ratios	Growth	GR_IC	% GR. in Invested Capital	-	1	-0,068	0,116%	11,11		96	1	0,00%	0,00	127
Financial Ratios	Growth	GR_CE	% GR. in Capital Employed	-	1	-1,044	1,011%	72,55		18	1	-0,03%	0,00	56
Financial Ratios	Growth	GR_AR	% GR. in Accounts Receivables	-	1	-0,315	0,346%	142,89		55	1	-0,01%	0,00	107
Financial Ratios	Growth	GR_AP	% GR. in Accounts Payables	-	1	-0,182	0,176%	136,68		76	1	0,00%	0,00	112
Financial Ratios	Growth	GR_S	% GR. in Sales	Altman et al. (2016)	1	-1,995	1,580%	1533,48						
Financial Ratios	Growth	GR_GP	% GR. in Gross Profit	-	1	-0,933	1,116%	416,82		17				
Financial Ratios	Growth	GR_EBITDA	% GR. in EBITDA	-	1	-0,067	0,223%	8,47		65	1	0,00%	0,00	131
Financial Ratios	Growth	GR_EBIT	% GR. in EBIT	-	1	-0,030	0,086%	22,50		110				

Financial Ratios	Growth	GR_EBIE	% GR. in EBIE	-	1	-0,031	0,087%	8,24		109				
Financial Ratios	Growth	GR_NI	% GR. in Net Income	-	1	-0,014	0,043%	1,34	37	131	1	0,00%	0,00	147
Financial Ratios	Growth	GR_OP_CF	% GR. in Operating Working Capital	-	1	-0,006	0,017%	1,68	49	159	1	0,00%	0,00	156
Financial Ratios	Growth	GR_FCF	% GR. Free Cash Flow	-	1	0,000	0,000%	1,53	85	226	1	0,00%	1,00	
Financial Ratios	Growth	D_Employees	Change in # of Employees	-	1	-0,019	0,060%	1,90	33	123				
Financial Ratios	Growth	GR_Employees	% GR in Employees	-	1	-1,716	0,519%	110,91		42				
Financial Ratios	Leverage	BEQ/TA	Equity / Total Assets	Dakovic et al. (2010)	1	-0,935	0,384%	108,32		52	1	-0,05%	0,00	39
Financial Ratios	Leverage	EQTL	Equity / Total Liabilities	Altman (1968)	1	0,178	0,300%	11,65		58	1	0,01%	0,00	93
Financial Ratios	Leverage	RETA	Retained Earnings / Total Assets	Altman (1968)	1	-1,093	0,574%	0,54	14	40	1	-0,06%	0,00	34
Financial Ratios	Leverage	DTA	Debt / Total Assets	Karels & Prakash (1987)	1	-0,501	0,065%	58,71		120	1	-0,03%	0,00	57
Financial Ratios	Leverage	LTLTA	Long Term Liabilities / Total Assets	Zavgren and Friedman (1988)	1	-0,793	0,132%	70,89		90	1	-0,04%	0,00	45
Financial Ratios	Leverage	NDA	Net Debt / Total Assets	-	1	-0,369	0,109%	277,48		99	1	-0,02%	0,00	64
Financial Ratios	Leverage	CETL	Capital Employed / Total Liabilities	Andrikopoulos et al. (2018)	1	0,149	0,182%	22,96		73	1	0,01%	0,00	96
Financial Ratios	Leverage	STDEQ	Short Term debt / Equity	Andrikopoulos et al. (2018)	1	-0,076	0,011%	3,19		168	1	0,00%	0,04	124
Financial Ratios	Leverage	IC_IBD	Interest Cost / Interest Bearing Debt	-	1	0,104	0,036%	1,31	39	135	1	0,00%	0,00	122
Financial Ratios	Leverage	IC_TL	Interest Cost / Total Liabilities	-	1	0,172	0,082%	1,52	30	116	1	0,01%	0,00	102
Financial Ratios	Leverage	STDA	Short Term debt / Assets	Altman et al. (2016)	1	1,303	0,031%	1,78	40	141	1	0,10%	0,00	25
Financial Ratios	Leverage	IC_adj_TL	Interest /(Liab. + Provisions + DTL)	-	1	5,513	0,553%	0,04	15	41	1	0,29%	0,00	7
Financial Ratios	Leverage	TLTA	Total Liabilities / Total Assets	Shumway (2001)	1	0,935	0,384%	108,16		51	1	0,05%	0,00	39
Financial Ratios	Leverage	TLEQ	Total Liabilities / Equity	Altman and Sabato (2007)	1	-0,024	0,128%	2,37	22	92	1	0,00%	0,00	136
Financial Ratios	Leverage	IC_S	Interest Cost / Sales	-	1	5,324	0,159%	4,30		80	1	0,32%	0,00	5
Financial Ratios	Leverage	IC_A	Interest Cost / Total Assets	Kalak & Hudson (2016)	1	11,618	0,968%	0,03	10	21	1	0,58%	0,00	2
Financial Ratios	Leverage	NDE	Net Debt / Equity	Andrikopoulos et al. (2018)	1	-0,029	0,065%	15,64		119	1	0,00%	0,00	133
Financial Ratios	Leverage	APTL	Acc. Payables / Total Liabilites	Terradez et al. (2015)	1	-1,124	0,397%	5,48		48	1	-0,06%	0,00	35
Financial Ratios	Leverage	ln_MM	Ln (Maturity Matching)	Duan et al. (2018)	1	0,680	1,010%	2,67	9	19	1	0,03%	0,00	55
Financial Ratios	Leverage	CLTA	Current Liabilities / Total Assets	Dakovic et al. (2010)	1	1,378	0,747%	10,05		30	1	0,07%	0,00	30
Financial Ratios	Liquidity	CASHA	Cash / Total Assets	Altman & Sabato (2007)	1	0,614	0,122%	355,00		94	1	0,03%	0,00	49
Financial Ratios	Liquidity	CASHE	Cash / Equity	-	1	-0,209	0,163%	145,60		79	1	-0,01%	0,00	83
Financial Ratios	Liquidity	CASHCA	Cash / Current Assets	Duan et al. (2018)	1	0,056	0,001%	271,80		210	1	0,00%	0,26	
Financial Ratios	Liquidity	CASHS	Cash / Sales	Andrikopoulos et al. (2018)	1	0,253	0,086%	39,24		111	1	0,01%	0,00	101
Financial Ratios	Liquidity	CASHCL	Cash / Current Liabilities	Pederzoli & Torricelli (2010)	1	0,107	0,232%	53,44		64	1	0,01%	0,00	108
Financial Ratios	Liquidity	WCTA	Working Capital / Total Assets	Altman (1968)	1	-0,427	0,086%	207,63		112	1	-0,02%	0,00	61
Financial Ratios	Liquidity	WCCL	Working Capital / Current Liabilities	Pederzoli & Torricelli (2010)	1	0,072	0,148%	20,58		86	1	0,00%	0,00	116
Financial Ratios	Liquidity	QUICK	(Curr. Assets - Inv.) / Curr. Liabilities	Altman et al. (2016)	1	0,093	0,222%	44,95		66	1	0,01%	0,00	110
Financial Ratios	Liquidity	MM	(Curr. Liab. - Cash)/ Total Liab.	Duan et al. (2018)	1	-0,250	0,279%	139,23		59	1	-0,01%	0,00	73
Financial Ratios	Liquidity	CACL	Current Assets / Current Liabilities	Shumway (2001)	1	0,072	0,148%	20,58		86	1	0,00%	0,00	116
Financial Ratios	Liquidity	ARTL	Acc. Receivables / Total Liabilites	Altman & Sabato (2007)	1	0,377	0,022%	66,63		154	1	0,02%	0,00	59
Financial Ratios	Other	saksu_S	Pledged Assets / Sales	-	1	0,000	0,001%	3,21		214	1	0,00%	0,73	
Financial Ratios	Other	saksu_E	Pledged Assets / Equity	-	1	0,000	0,045%	0,24	36	130	1	0,00%	0,00	167
Financial Ratios	Other	saksu_A	Pledged Assets / Assets	-	1	0,000	0,156%	8,74		82	1	0,00%	0,00	164
Financial Ratios	Other	EBIT_NI	EBIT / Net Income	-	1	-0,017	0,088%	0,00	28	108	1	0,00%	0,00	142
Financial Ratios	Other	Tax_A	Tax Cost / Total Assets	Kalak & Hudson (2016)	1	-4,401	0,063%	300,56		121	1	-0,11%	0,00	22
Financial Ratios	Other	FA_A	Fixed Assets / Total Assets	-	1	-0,358	0,175%	149,15		77				
Financial Ratios	Other	IC_TC	Interest Cost / Total Costs	-	1	4,638	0,088%	3,28		107	1	0,28%	0,00	8
Financial Ratios	Other	Oplev	(Δ EBIT)/(Δ Sales)	-	1	-0,019	0,008%	3,78		179				
Financial Ratios	Other	Oplev_Avg	Within Firm Avg. (Δ EBIT)/(Δ Sales)	-	1	-0,025	0,010%	1,56	57	172				
Financial Ratios	Other	IndOplev_Avg	Within Ind Avg. (Δ EBIT)/(Δ Sales)	-	1	-0,174	0,009%	0,20	61	177				
Financial Ratios	Other	S_CV	Coefficient of Variance of Sales	-	1	-0,003	0,021%	88,87		155				
Financial Ratios	Other	EBIT_CV	Coefficient of Variance of EBIT	-	1	-0,004	0,041%	2,49	38	132				
Financial Ratios	Other	NI_CV	Coefficient of Variance of Net Income	-	1	-0,002	0,019%	1,39	47	157				
Financial Ratios	Other	RETA_sd	Standard Deviation in RETA	-	1	0,003	0,004%	0,00	70	193				
Financial Ratios	Other	WCTA_sd	Standard Deviation in WCTA	-	1	0,003	0,003%	0,00	75	205				
Financial Ratios	Other	EQTA_sd	Standard Deviation in EQTA	-	1	0,003	0,004%	0,02	72	196				
Financial Ratios	Other	ROAA_sd	Standard Deviation in ROAA	-	1	0,019	0,014%	0,04	53	164				
Financial Ratios	Other	A_turn_sd	Standard Deviation in Asset Turnover	-	1	0,004	0,012%	0,44	55	166				
Financial Ratios	Other	NITA_sd	Standard Deviation in NITA	-	1	0,018	0,014%	0,00	52	163				
Financial Ratios	Other	TLTA_sd	Standard Deviation in TLTA	-	1	0,003	0,004%	0,02	72	196				
Financial Ratios	Other	CACL_sd	Standard Deviation in CACL	-	1	0,002	0,024%	2,11	44	151				
Financial Ratios	Other	IC_EBITDA_sd	Standard Deviation in IC_EBITDA	-	1	0,003	0,011%	9,71		170				
Financial Ratios	Other	ln_RelSize_sd	Standard Deviation in ln(RelSize)	-	1	0,280	0,076%	70,52		117				
Financial Ratios	Other	pledged	Recommend w/ Notation	*Altman et al. (2016)*	1	-0,279	0,091%	72,60		104	1	-0,01%	0,00	86
Financial Ratios	Other	cont_SH_contr	Conditional shareholder's contribution	-	1	0,461	0,121%	0,21	23	95	1	0,04%	0,00	47
Financial Ratios	Other	other_cont_contr	Other Cont. Liabilities. (Warranties etc)	-	1	-0,320	0,026%	2,89		146	1	-0,01%	0,00	77
Financial Ratios	Other	overdraft	Bank Overdraft Facility Granted	-	1	-0,053	0,003%	15,19		200	1	0,00%	0,97	
Financial Ratios	Other	UTRS	Untaxed Reserves / Sales	-	1	-6,256	0,822%	137,79		28	1	-0,26%	0,00	9
Financial Ratios	Other	UTRA	Untaxed Reserves / Total Assets	-	1	-10,589	2,187%	1,60	2	2	1	-0,31%	0,00	6
Financial Ratios	Other	UTREQ	Untaxed Reserves / Equity	-	1	-3,054	2,150%	2,24	3	3				
Financial Ratios	Other	UTRTL	Untaxed Reserves / Total Liabilities	-	1	-4,147	1,332%	27,56		13	1	-0,17%	0,00	12
Financial Ratios	Other	UTRRE	Untaxed Reserves / Retained Earnings	-	1	-0,149	0,412%	68,44		47				
Financial Ratios	Other	UTRNI	Untaxed Reserves / Net Income	-	1	-0,029	0,273%	0,30	16	60	1	0,00%	0,00	139
Financial Ratios	Other	UTR/EMP	Untaxed Reserves / # Employees	-	1	-0,013	1,875%	79,57		7				
Financial Ratios	Other	UTRDP	Untaxed Reserves Dummy	-	1	-1,559	2,838%	0,07	1	1	1	-0,09%	0,00	27
Financial Ratios	Profitab.	GP_M	Gross Profit / Sales	Karels & Prakash (1987)	1	-1,304	0,516%	3,69		44	1	-0,06%	0,00	31
Financial Ratios	Profitab.	EBITDA_M	EBITDA / Sales	-	1	-2,409	1,003%	11,03		20	1	-0,12%	0,00	20
Financial Ratios	Profitab.	EBIT_M	EBIT / Sales	Altman & Sabato (2007)	1	-2,139	0,893%	12,10		25	1	-0,11%	0,00	23
Financial Ratios	Profitab.	EBIE_M	EBIE / Sales	-	1	-2,544	0,854%	40,32		27	1	-0,12%	0,00	19
Financial Ratios	Profitab.	NI_M	Net Income / Sales	Andrikopoulos et al. (2018)	1	-2,602	0,615%	58,46		37	1	-0,13%	0,00	18
Financial Ratios	Rel. Perf.	RelROACE	ROACE-Avg. Population ROACE	-	1	0,000	0,000%	9,49		233	1	0,00%	0,92	
Financial Ratios	Rel. Perf.	RelVA	Capital Employed * RelROACE	-	1	0,000	0,007%	0,09	64	183	1	0,00%	0,05	
Financial Ratios	Rel. Perf.	RelROACEInd	ROACE-Avg. Industry ROACE	-	1	0,000	0,000%	4,76		230	1	0,00%	0,89	
Financial Ratios	Rel. Perf.	RelVAInd	Capital Employed * RelROACEInd	-	1	0,000	0,000%	0,00	87	231	1	0,00%	0,15	
Financial Ratios	Rel. Perf.	RelROACEIndKomm	ROACE - Avg. Industry-Muni ROACE	-	1	-0,003	0,005%	1,21	68	189	1	0,00%	0,10	
Financial Ratios	Rel. Perf.	RelVAIndKomm	Capital Employed * RelROACEIndKomm	-	1	0,000	0,000%	0,00	88	232	1	0,00%	0,54	

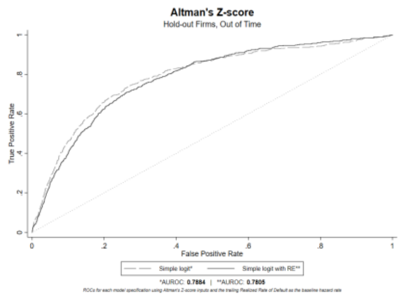
Financial Ratios	Rel. Perf.	GR_RelVA	% GR in RelVA	-	1	-0,006	0,006%	7,65		184	1	0,00%	0,11	
Financial Ratios	Rel. Perf.	GR_RelVAInd	% GR in RelVAInd	-	1	-0,003	0,002%	0,18	76	207	1	0,00%	0,29	
Financial Ratios	Rel. Perf.	GR_RelVAIndKomm	% GR in RelVAIndKomm	-	1	-0,004	0,005%	0,28	66	187	1	0,00%	0,09	157
Financial Ratios	Rel. Perf.	D_RelROACE	Δ in ROACE-Avg. Population ROACE	-	1	-0,241	0,083%	1,54	29	115	1	0,00%	0,00	114
Financial Ratios	Returns	ROAIC	EBIT / Average Invested Capital	-	1	-0,024	0,026%	1,16	42	147	1	0,00%	0,00	141
Financial Ratios	Returns	ROACE	EBIE / Average Capital Employed	-	1	-0,710	0,735%	13,25		32	1	-0,03%	0,00	51
Financial Ratios	Returns	EBIE/ATA	EBIE / Average Assets	Shumway (2001)	1	-2,745	1,692%	1,17	5	8	1	-0,13%	0,00	17
Financial Ratios	Returns	ROAE	Net Income / Average Equity	Altman et al. (2016)	1	-0,132	0,062%	1,60	32	122	1	0,00%	0,00	120
Financial Ratios	Returns	EBITDAROATA	EBITDA / Average Assets	Altman & Sabato (2007)	1	-2,846	1,878%	6,32		6	1	-0,14%	0,00	16
Financial Ratios	Returns	NITA	Net Income / Total Assets	Shumway (2001)	1	-2,903	1,487%	0,14	7	10	1	-0,15%	0,00	15
Financial Ratios	Returns	ValueAdd_A	(EBIT + Salary) / Total Assets	-	1	0,254	0,216%	15,13		67	1	0,02%	0,00	68
Financial Ratios	Returns	GPCA	Gross Profit / Current Assets	Duan et al. (2018)	1	-0,111	0,092%	12,59		103	1	0,00%	0,00	119
Financial Ratios	Size & Age	LNA	Ln(Total Assets)	Altman et al. (2016)	1	-0,492	1,340%	63,63		12	1	-0,02%	0,00	63
Financial Ratios	Size & Age	ln_Asq	Ln(Total Assets)^2	Altman et al. (2016)	1	-0,032	1,210%	65,77		14	1	0,00%	0,00	133
Financial Ratios	Size & Age	LNS	Ln(Sales)	Altman et al. (2016)	1	-0,328	0,663%	103,05		34	1	-0,01%	0,00	95
Financial Ratios	Size & Age	ln_Ssq	Ln(Sales)^2	Altman et al. (2016)	1	-0,020	0,588%	122,31		39	1	0,00%	0,00	143
Financial Ratios	Size & Age	bol_age_adj	Company Age	Altman et al. (2016)										
Financial Ratios	Size & Age	ln_Age	Ln (Company Age)	Altman et al. (2016)										
Financial Ratios	Size & Age	RelSize	Sales / Avg. Population Sales	Shumway (2001)	1	-0,074	0,153%	187,55		84	1	0,00%	0,00	130
Financial Ratios	Size & Age	RelSizeInd	Sales / Avg. Industry Sales	-	1	-0,082	0,158%	235,45		81	1	0,00%	0,00	129
Financial Ratios	Size & Age	RelSizeIndKomm	Sales / Avg. Industry-Muni Sales	-	1	-0,100	0,197%	149,06		71	1	0,00%	0,00	125
Financial Ratios	Size & Age	RelSizeSNI	Sales / Avg. SNI Sales	-	1	-0,085	0,139%	593,31		89	1	0,00%	0,01	132
Financial Ratios	Size & Age	RelSizeSNIKomm	Sales / Avg. SNI-Muni Sales	-	1	-0,110	0,124%	298,88		93	1	0,00%	0,02	128
Financial Ratios	Size & Age	ln_RelSize	Ln (Sales / Avg. Population Sales)	Shumway (2001)	1	-0,306	0,694%	68,90		33	1	-0,01%	0,00	94
Financial Ratios	Size & Age	ln_RelSizeInd	Ln (Sales / Avg. Industry Sales)	-	1	-0,351	0,899%	37,79		24	1	-0,01%	0,00	80
Financial Ratios	Size & Age	ln_RelSizeIndKomm	Ln (Sales / Avg. Industry-Muni Sales)	-	1	-0,347	0,881%	35,69		26	1	-0,01%	0,00	78
Financial Ratios	Size & Age	ln_RelSizeSNI	Ln (Sales / Avg. SNI Sales)	-	1	-0,368	0,922%	46,22		23	1	-0,01%	0,00	75
Financial Ratios	Size & Age	ln_RelSizeSNIKomm	Ln (Sales / Avg. SNI-Muni Sales)	-	1	-0,341	0,636%	36,93		35	1	-0,01%	0,00	81
Financial Ratios	Size & Age	ACC_FCF	Net Income + D&A	Altman et al. (2016)	1	-0,001	1,442%	194,42		11	1	0,00%	0,00	161
Financial Ratios	Size & Age	YoungC	Young Company (3-9 years)	Altman et al. (2016)	1	0,440	0,012%	0,27	56	167	1	0,01%	0,00	92
Non-Financial	Contagion	Current_Ind_Def	Default Rate - Industry	Altman et al. (2016)	1	26,700	0,241%	2,18	17	62	1	1,50%	0,00	1
Non-Financial	Contagion	Current_Komm_Def	Default Rate - Municipality	-	1	2,951	0,028%	0,13	41	144	1	0,16%	0,00	13
Non-Financial	Contagion	Current_SNI_Def	Default Rate - SNI	-	1	10,583	0,213%	1,98	18	68	1	0,47%	0,00	4
Non-Financial	Contagion	Current_Ind_Komm_Def	Default Rate - Muni-Industry	-	1	3,471	0,067%	0,02	31	118	1	0,17%	0,00	11
Non-Financial	Contagion	Current_SNI_Komm_Def	Default Rate - Muni-SNI	-	1	1,898	0,060%	0,01	34	124	1	0,05%	0,00	41
Non-Financial	Corp. Gov	Board_sal_S	(CEO & BOD Salary) / Sales	-	1	1,051	0,091%	107,28		105	1	0,04%	0,00	44
Non-Financial	Corp. Gov	Board_sal_OP	(CEO & BOD Salary) / EBIT	-	1	-0,012	0,090%	0,06	27	106	1	0,00%	0,00	144
Non-Financial	Corp. Gov	CEO_sal_S	CEO Salary / Sales	-	1	0,740	0,059%	5,51		125	1	0,05%	0,00	37
Non-Financial	Corp. Gov	ln_Board_members	Ln (# Board Members)	Wilson et al. (2011)	1	0,448	0,086%	26,95		113	1	0,01%	0,00	82
Non-Financial	Corp. Gov	D_Board_members	Change in # Board Members	*Wilson et al. (2011)*	1	-0,005	0,000%	2,86		221	1	0,00%	0,01	137
Non-Financial	Corp. Gov	GR_Board_members	% GR in Board Members	*Wilson et al. (2011)*	1	0,145	0,004%	3,22		195	1	0,01%	0,00	106
Non-Financial	Corp. Gov	CEO_BoardMove_Cum	Cumulative CEO on Board and Moved	*Wilson et al. (2011)*	1	0,600	0,005%	0,94	67	188	1	0,05%	0,00	36
Non-Financial	Corp. Gov	CEO_BoardChange_Cum	Cumulative CEO on Board Changed	*Wilson et al. (2011)*	1	0,251	0,011%	5,77		169	1	0,03%	0,00	50
Non-Financial	Corp. Gov	CEO_Appoint2Board_Cum	Cumulative CEO Appointed to Board	*Wilson et al. (2011)*	1	0,207	0,029%	13,87		143	1	0,03%	0,00	53
Non-Financial	Corp. Gov	AGM	General Meeting of Shareholders Held	-	1	-0,643	0,379%	7,24		53	1	-0,03%	0,00	58
Non-Financial	Corp. Gov	Rel_Loans	Loans to Related Parties	-	1	-0,158	0,000%	1,88	82	223	1	0,00%	0,85	
Non-Financial	Corp. Gov	SEV	Severance Pay Exists	-	1	0,089	0,000%	4,24		229	1	0,02%	0,65	
Non-Financial	Corp. Gov	CEO_BoardMove	CEO on Board and Moved	*Wilson et al. (2011)*	1	1,908	0,015%	1,06	51	162	1	0,39%	0,13	
Non-Financial	Corp. Gov	CEO_BoardChange	CEO on Board Changed	*Wilson et al. (2011)*	1	0,537	0,009%	0,27	58	174	1	0,06%	0,00	33
Non-Financial	Corp. Gov	CEO_Appoint2Board	CEO Appointed to Board	*Wilson et al. (2011)*	1	0,279	0,008%	1,92	62	178	1	0,02%	0,00	70
Non-Financial	Corp. Gov	CEO_LeaveBoard	CEO Leaves Board	*Altman et al. (2016)*	1	0,709	0,017%	1,36	50	160	1	0,08%	0,00	29
Non-Financial	Corp. Gov	CEO_OnBoard	CEO on Board	*Altman et al. (2016)*	1	0,067	0,004%	2,86		199	1	0,02%	0,00	65
Non-Financial	Corp. Gov	Chair_OnBoard	Chairman Exists	-	1	0,062	0,004%	6,22		198	1	0,01%	0,00	72
Non-Financial	Corp. Gov	Auditor_Change	Auditor Changed	Kluger & Shields (1989)	1	0,059	0,001%	4,85		211	1	0,01%	0,00	90
Non-Financial	Corp. Gov	CEO_Mult_Dir	CEO has Multiple Directorships	Wilson et al. (2011)	1	0,100	0,003%	5,43		204	1	0,01%	0,03	84
Non-Financial	Corp. Gov	Chair_Mult_Dir	Chairman has Multiple Directorships	Wilson et al. (2011)	1	0,153	0,010%	4,61		173	1	0,01%	0,02	99
Non-Financial	Corp. Gov	CEO_Chair	CEO is the Chairman	Chiampi (2013)	1	0,085	0,001%	3,53		213	1	0,02%	0,04	67
Non-Financial	Patents	Patent_appln_Cum	Cumulative nr of Patent Application	Buddelmeyer et al. (2009)	1	0,095	0,002%	2,97		209	1	0,01%	0,00	88
Non-Financial	Patents	Patent_appln	Patent Applied	Buddelmeyer et al. (2009)	1	0,124	0,000%	3,01		220	1	0,00%	0,78	
Non-Financial	Reporting	revber_Any	Any Auditor Remark - Dummy	Dakovic et al. (2010)	1	0,957	0,929%	0,30	11	22	1	0,08%	0,00	28
Non-Financial	Reporting	revber_Any_Cum	Cumulative Any Auditor Remark	*Dakovic et al. (2010)*	1	0,135	0,338%	20,50		56	1	0,01%	0,00	103
Non-Financial	Reporting	revber_ET_Cum	Cumulative Not Recommend	*Dakovic et al. (2010)*	1	0,271	0,178%	33,44		74	1	0,02%	0,00	60
Non-Financial	Reporting	revber_O_Cum	Cumulative Incomplete	*Dakovic et al. (2010)*	1	0,000	0,000%	0,00	90	235	1	0,00%	.	
Non-Financial	Reporting	revber_S_Cum	Cumulative Missing	*Dakovic et al. (2010)*	1	-0,072	0,004%	0,00	71	194	1	0,00%	0,01	112
Non-Financial	Reporting	revber_TK_Cum	Cumulative Recommend w/ Notation	*Dakovic et al. (2010)*	1	0,137	0,269%	15,65		61	1	0,01%	0,00	97
Non-Financial	Reporting	Latefiling_Cum	Cumulative Latefiling	*Altman et al. (2016)*	1	0,064	0,026%	19,00		150	1	0,01%	0,00	91
Non-Financial	Reporting	CEO_Latefiling_Cum	Cumulative Latefiling from CEO	*Altman et al. (2016)*	1	0,103	0,016%	4,94		161	1	0,01%	0,00	79
Non-Financial	Reporting	Chair_Latefiling_Cum	Cumulative Latefiling from Chairman	*Altman et al. (2016)*	1	0,001	0,000%	0,00	86	227	1	0,00%	0,00	149
Non-Financial	Reporting	CEO_LeaveBoard_Cum	Cumulative CEO Leaves Board	*Altman et al. (2016)*	1	0,419	0,023%	2,54	45	152	1	0,05%	0,00	43
Non-Financial	Reporting	Auditor_Change_Cum	Cumulative nr of Auditor Changes	*Altman et al. (2016)*	1	0,007	0,000%	1,74	81	222	1	0,01%	0,00	71
Non-Financial	Reporting	AUDNR	Not Recommend	*Altman et al. (2016)*	1	1,348	0,390%	3,24		49	1	0,21%	0,00	10
Non-Financial	Reporting	revber_O	Incomplete	*Altman et al. (2016)*	1	0,000	0,000%	0,00	90	235	1	.	.	
Non-Financial	Reporting	revber_S	Missing	*Altman et al. (2016)*	1	-0,164	0,006%	0,07	65	186	1	-0,01%	0,00	76
Non-Financial	Reporting	revber_TK	Recommend w/ Notation	*Altman et al. (2016)*	1	0,957	0,737%	0,09	12	31	1	0,09%	0,00	26
Non-Financial	Reporting	Latefiling	Late Filing of Financial Statements	Wilson et al. (2011)	1	0,421	0,096%	0,11	26	102	1	0,02%	0,00	69
Non-Financial	Reporting	revber_Any_Cum2	Cumulative ET/TK/S Remarks	*Dakovic et al. (2010)*	1	0,143	0,371%	25,94		54	1	0,01%	0,00	100
Non-Financial	Reporting	revber_Any2	Any ET/TK/S Remark	*Dakovic et al. (2010)*	1	1,122	1,162%	0,10	8	16	1	0,12%	0,00	21

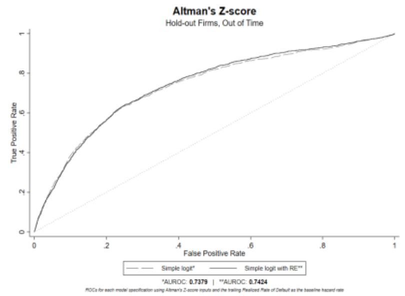
This table reports the examined variables, their definition, an example peer-reviewed paper that makes use of each variable as appropriate, as well as statistics pertaining to the univariate survival analysis tests and average marginal effects. PH = proportional hazard. If the proportional hazard chi-squared statistic is above 2.71, the variable is not propotional. The right-most column pseudo R-squared rank lists the order of which the variable ranks on pseudo R-squared, if the model converged and the proportional hazard assumption is not violated. The Pseudo R-squared rank ex. PH lists the same ranking, but with the relaxed proportional hazard assumption. The rank of the average marginal effect for the same variables are shown in the rightmost column. Variables that do not meet the proportional hazard assumptions can do so in a multivariate context. We refer to Table A3 for the global test of proportionality. *Bellovary et al. (2007) is a review of studies from the 1930's to 2007, not a specific bankruptcy study. Previous research in italics implies a very similar metric was used, either inversed or a metric aimed at capturing virtually the same risk. Acc. Qual = accounting quality. Corp. Gov = corporate governance. CFM = cash flow metrics. Rel. Perf = relative performance. Con. = Converged. RE = Random Effects.

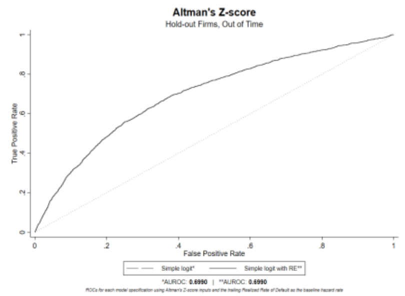
Table A2: Hold-Out Firms, Out-of-Time ROC and PR

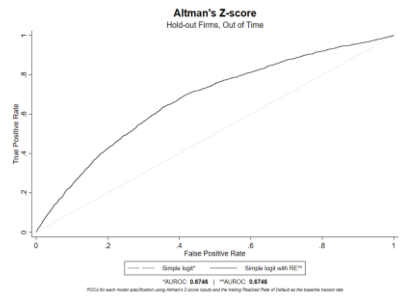
Altman (1968) - Hold-Out Firms, Out-of-Time

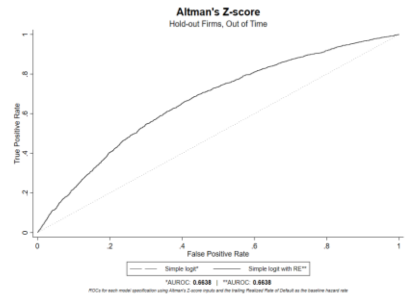
ROC

One-Year


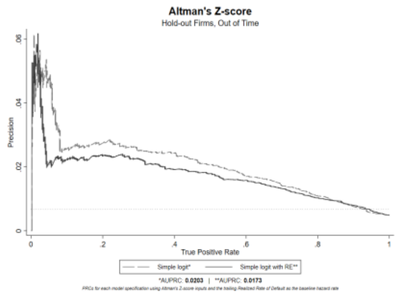
Two-Year


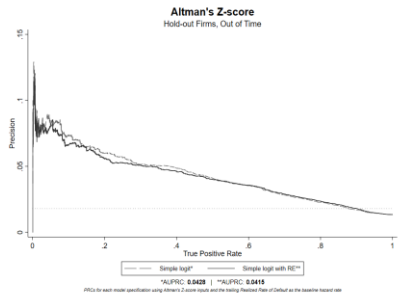
Three-Year


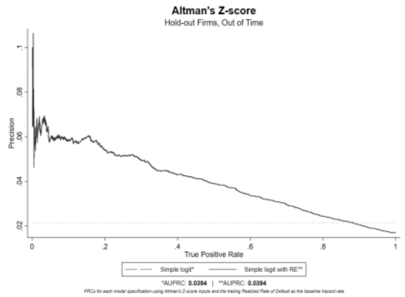
Four-Year


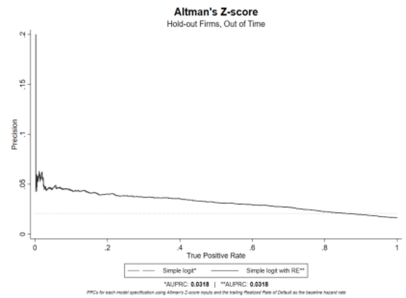
Five-Year


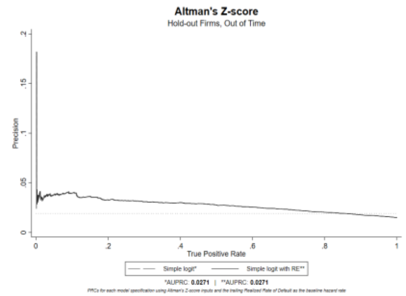
PR

One-Year


Two-Year


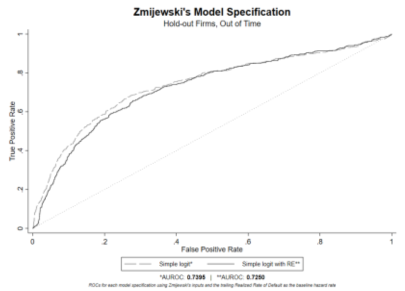
Three-Year


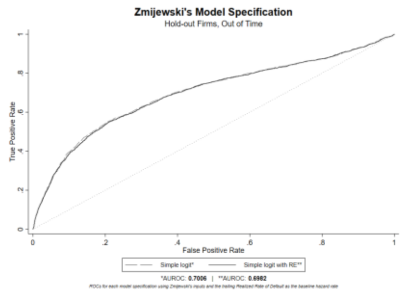
Four-Year


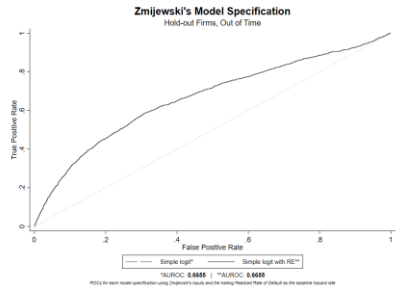
Five-Year


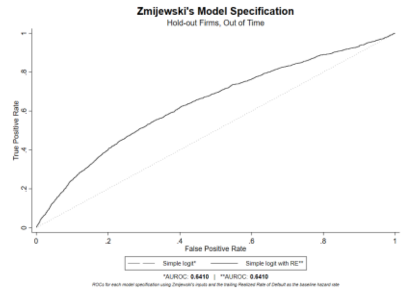
Zmijewski (1984) - Hold-Out Firms, Out-of-Time

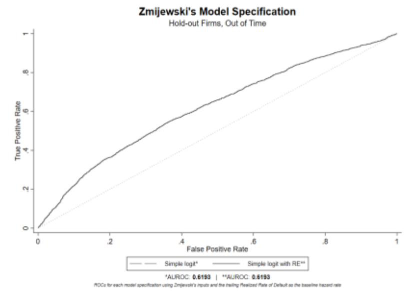
ROC

One-Year


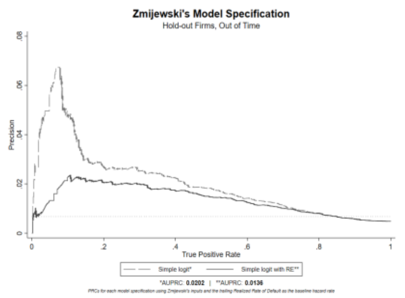
Two-Year


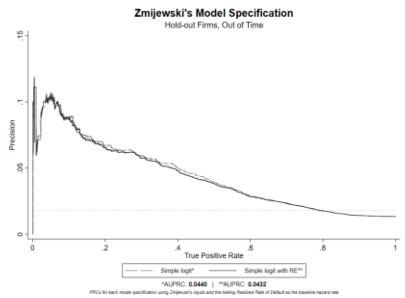
Three-Year


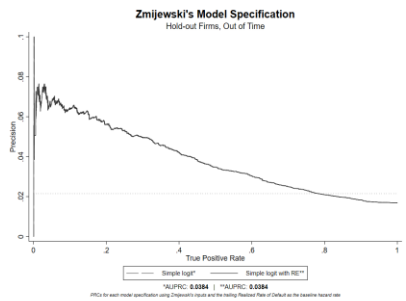
Four-Year


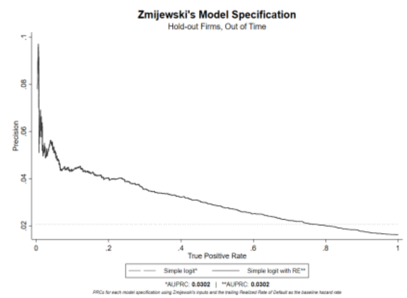
Five-Year


PR

One-Year


Two-Year


Three-Year


Four-Year


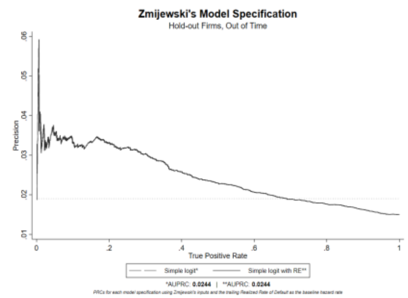
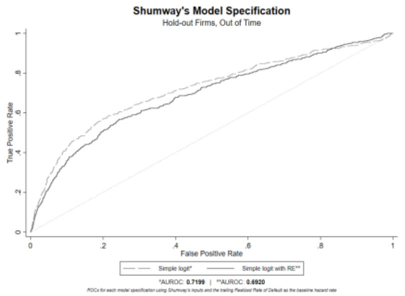
Five-Year


Table A2: Hold-Out Firms, Out-of-Time ROC and PR, Cont.

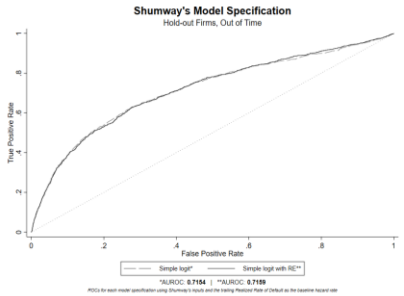
Shumway (2001) - Hold-Out Firms, Out-of-Time

ROC

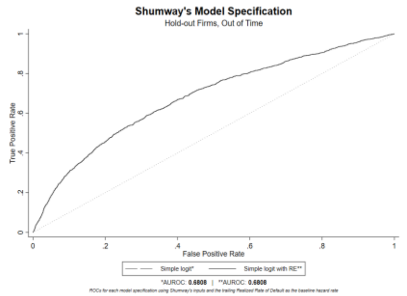
One-Year



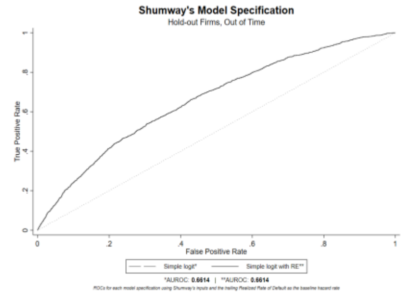
Two-Year



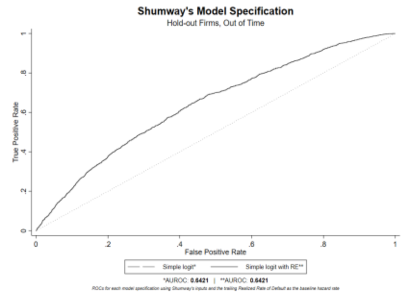
Three-Year



Four-Year

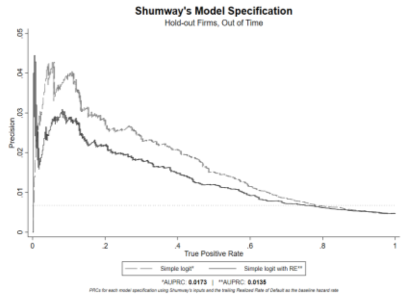


Five-Year

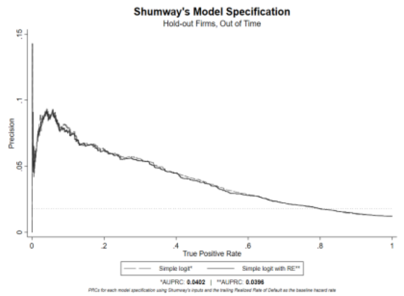


PR

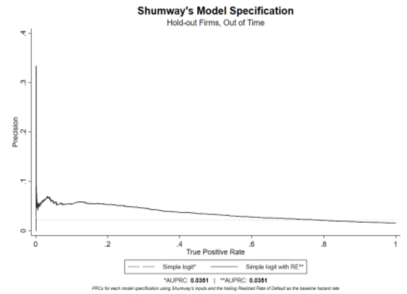
One-Year



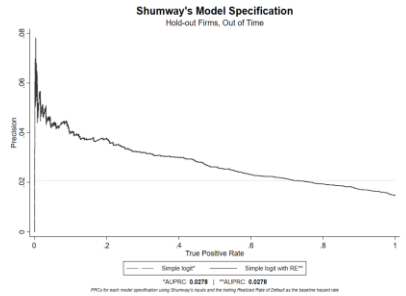
Two-Year



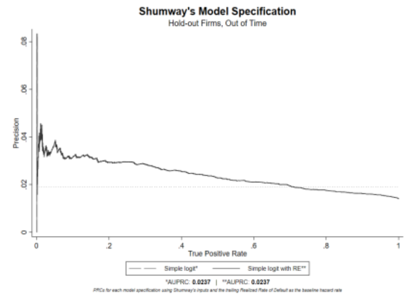
Three-Year



Four-Year



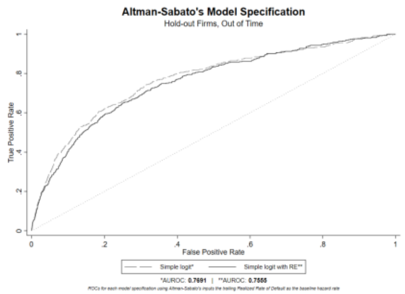
Five-Year



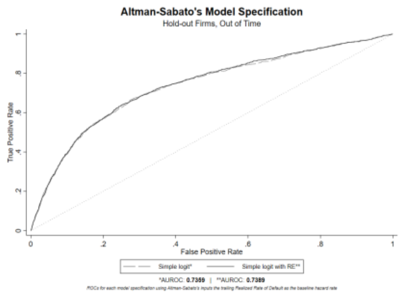
Altman and Sabato (2007) - Hold-Out Firms, Out-of-Time

ROC

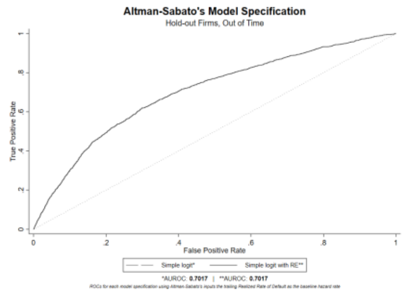
One-Year



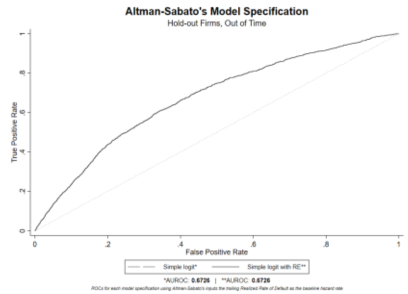
Two-Year



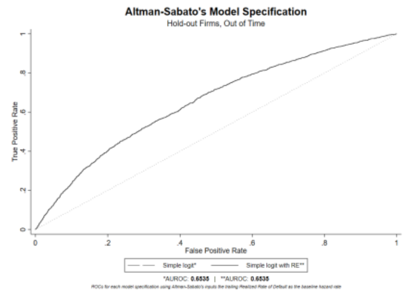
Three-Year



Four-Year

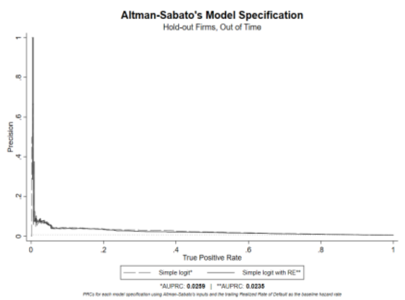


Five-Year

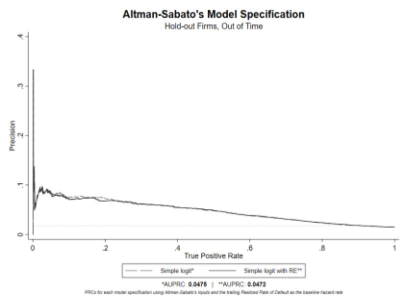


PR

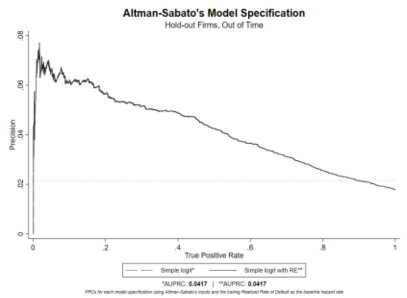
One-Year



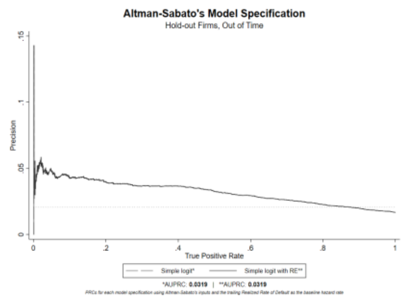
Two-Year



Three-Year



Four-Year



Five-Year

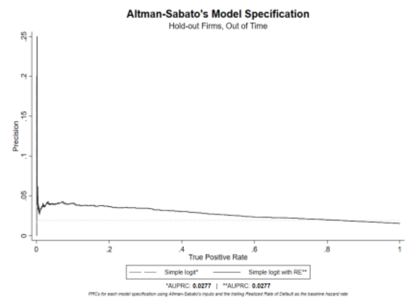
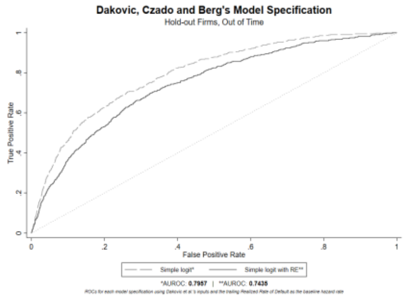


Table A2: Hold-Out Firms, Out-of-Time ROC and PR, Cont.

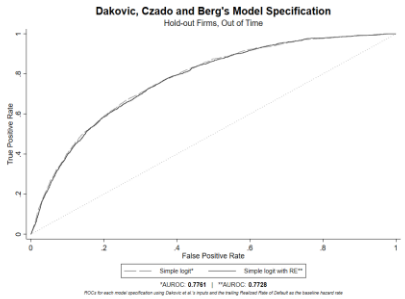
Dakovic et al. (2010) - Hold-Out Firms, Out-of-Time

ROC

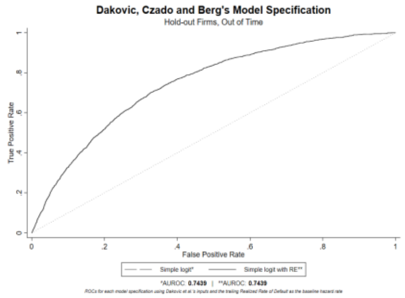
One-Year



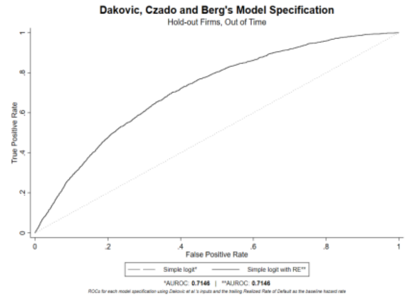
Two-Year



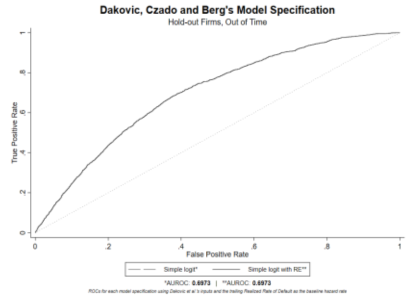
Three-Year



Four-Year

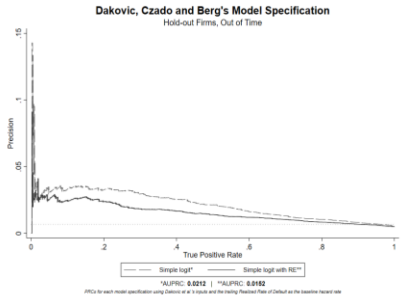


Five-Year

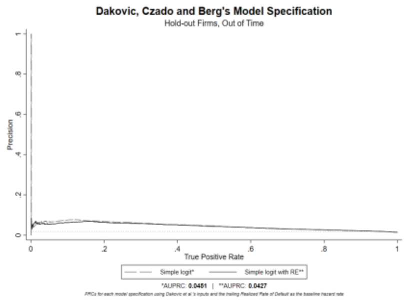


PR

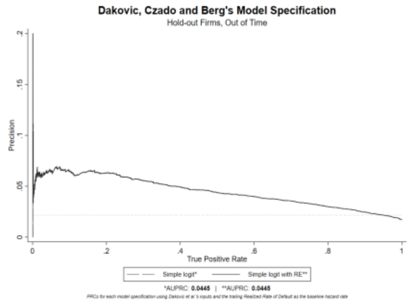
One-Year



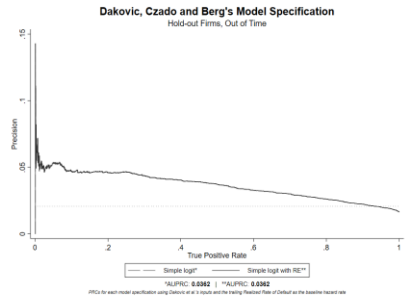
Two-Year



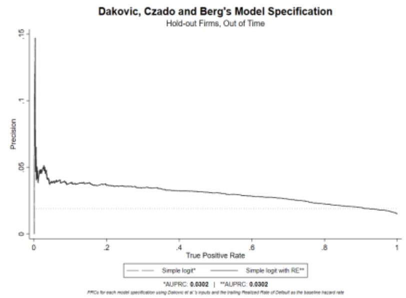
Three-Year



Four-Year



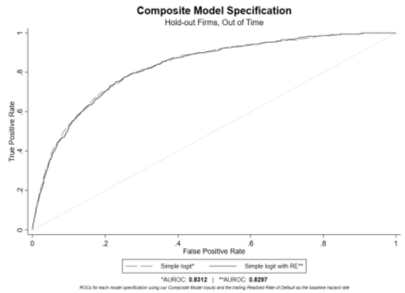
Five-Year



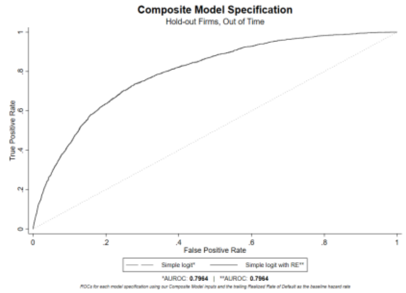
Composite Model - Hold-Out Firms, Out-of-Time

ROC

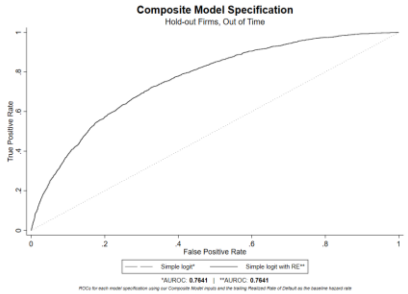
One-Year



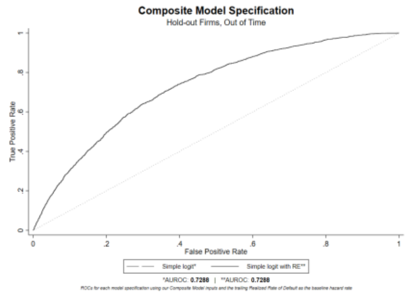
Two-Year



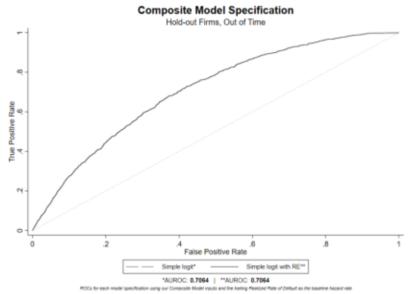
Three-Year



Four-Year

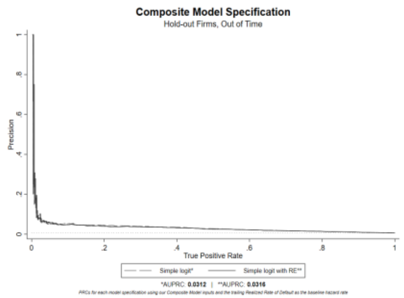


Five-Year

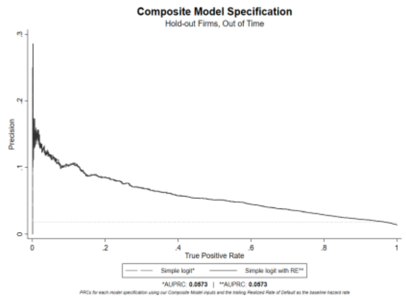


PR

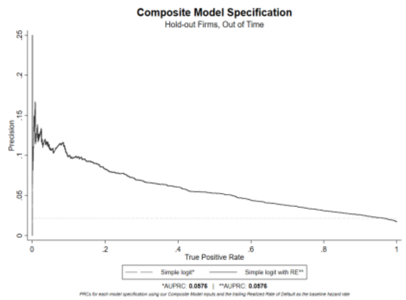
One-Year



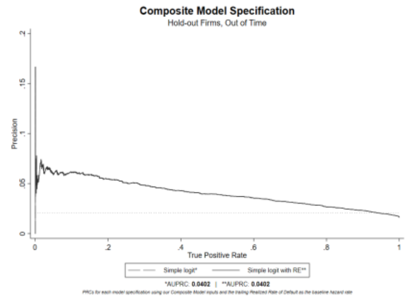
Two-Year



Three-Year



Four-Year



Five-Year

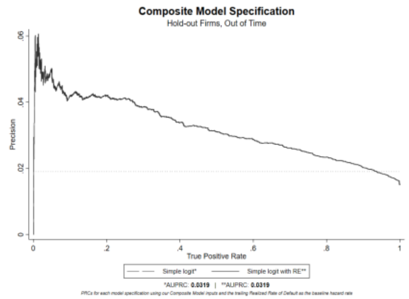


Table A3: One-Year Test of Proportionality for all Models

Altman (1968)	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	0.06	21.87	0.00***
WC/TA	0.08	67.35	0.00***
RE/TA	-0.08	115.88	0.00***
EBIT/TA	0.02	10.65	0.00***
EQ/TL	0.05	19.90	0.00***
A_turn	-0.01	0.32	0.57
Global test		165.73	0.00***

Zmijewski (1984)	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	0.04	12.04	0.00***
NI/TA	-0.02	2.51	0.11
TL/TA	-0.01	2.00	0.16
CA/CL	0.01	1.01	0.31
Global test		16.57	0.00**

Shumway (2001)	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	0.04	9.43	0.002***
NI/TA	0.03	12.42	0.00***
TL/TA	-0.02	2.90	0.09*
LNRELSIZE	-0.09	100.01	0.00***
NICV	-0.05	0.60	0.44
ΔRELROACE	0.01	1.42	0.23
Global test		122.22	0.00***

Altman and Sabato (2007)	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	0.057	19.36	0.00***
STD/EQ	0.0158	2.56	0.11
CASH/TA	0.1374	220.89	0.00***
EBITDA/ATA	0.0238	9.04	0.00***
RE/TA	-0.062	53.1	0.00***
EBITDA/IE	-0.006	0.23	0.63
Global test		257.7	0.00***

Dakovic et al. (2010)	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	-	-	-
REVANM	-	-	-
AGE	-	-	-
IND1	-	-	-
IND2	-	-	-
IND3	-	-	-
IND4	-	-	-
IND5	-	-	-
IND6	-	-	-
IND7	-	-	-
IND8	-	-	-
IND9	-	-	-
IND10	-	-	-
DIV	-	-	-
EQTAPOS	-	-	-
EQTAPOSSQ	-	-	-
EQTANEG	-	-	-
EQTANEGSQ	-	-	-
LNSIZE	-	-	-
LNSIZE^2	-	-	-
expCASH/L	-	-	-
EBIE/ATA POS	-	-	-
EBIE/ATA POSSQ	-	-	-
EBIE/ATA NEG	-	-	-
EBIE/ATA NEGS	-	-	-
EBIE/ATA NEGCU	-	-	-
CL/TA	-	-	-
Global test		-	-

Composite Model	Rho	Chi-Squared	Prob >Chi-Squared
Baseline Hazard	0.06	21.29	0.00***
EBIE/ATA	0.02	3.83	0.05*
RE/TA	-0.05	29.10	0.00***
BEQ/TA	0.06	44.74	0.00***
IE/TL	-0.03	4.41	0.04**
AUDNR	0.04	9.41	0.00**
LNA	0.00	0.00	0.990
LNS	-0.06	48.27	0.00***
DPO	-0.03	5.76	0.02**
UTR/EMP	0.03	35.83	0.00***
Global test		234.72	0.00***

Reports the individual and global test of survival curves' proportionality (the Cox semi-parametric PH assumption) for the one-year models using scaled Schoenfeld residuals. The Dakovic et al. (2010) model did not converge due to a flat region resulting in a missing likelihood.

Table A4: One-Year DTDDH Model Regression Output

Altman (1968)	Simple Logit	w. Random Effects
Baseline Hazard	19.886***	56.703***
	(2,82)	(4,24)
WC/TA	0.122*	0.304***
	(0,05)	(0,11)
RE/TA	-1.243***	-2.905***
	(0,05)	(0,12)
EBIT/TA	-2.586***	-4.256***
	(0,08)	(0,14)
EQ/TL	0.319***	0.824***
	(0,01)	(0,03)
A_turn	0.066***	0.154***
	(0,01)	(0,01)
constant	-5.856***	-13.868***
	(0,10)	(0,31)
Firm Random Effects		26.363***
		(1,32)
Observations	843694	843694
BIC	67594	64260
AIC	67675	64354

Zmijewski (1984)	Simple Logit	w. Random Effects
Baseline Hazard	20.009***	54.007***
	(2,76)	(3,95)
NI/TA	-2.880***	-3.644***
	(0,09)	(0,19)
TL/TA	0.582***	0.946***
	(0,06)	(0,12)
CA/CL	0.121***	0.285***
	(0,01)	(0,01)
constant	-6.341***	-13.901***
	(0,11)	(0,29)
Firm Random Effects		23.271***
		(1,13)
Observations	843495	843495
BIC	66041	68814
AIC	65971	68755

Shumway (2001)	Simple Logit	w. Random Effects
Baseline Hazard	15.480***	61.049***
	(2,90)	(4,46)
NI/TA	-2.383***	-3.275***
	(0,12)	(0,25)
TL/TA	0.331***	0,13
	(0,06)	(0,15)
LNRELSIZE	-0.286***	-0.689***
	(0,01)	(0,04)
NICV	(0,00)	-0.009*
	0,00	0,00
ΔRELROACE	0,00	0.104*
	(0,03)	(0,04)
constant	-5.738***	-14.623***
	(0,11)	(0,35)
Firm Random Effects		33.444***
		(1,83)
Observations	678522	678522
BIC	53928	57001
AIC	53837	56921

Altman and Sabato (2007)	Simple Logit	w. Random Effects
Baseline Hazard	13.488***	46.764***
	(2,95)	(4,39)
STD/EQ	0.049*	0,04
	(0,02)	(0,04)
CASH/TA	1.141***	2.598***
	(0,07)	(0,15)
EBITDA/ATA	-2.746***	-4.224***
	(0,09)	(0,15)
RE/TA	-1.072***	-2.396***
	(0,05)	(0,10)
EBITDA/IE	-0.001***	-0.001**
	0,00	0,00
constant	-5.348***	-12.551***
	(0,11)	(0,31)
Firm Random Effects		23.724***
		(1,27)
Observations	735142	735142
BIC	59179	56699
AIC	59098	56606

Dakovic et al. (2010)	Simple Logit	w. Random Effects
Baseline Hazard	22.828***	68.497***
	(2,84)	(5,08)
REVANM	0,00	0.422***
	(0,01)	(0,02)
AGE	0,00	0.050***
	0,00	0,00
IND1	(0,10)	(0,04)
	(0,28)	(0,77)
IND2	0.347*	1.223**
	(0,14)	(0,40)
IND3	0.569***	1.699***
	(0,10)	(0,27)
IND4	0.422***	1.377***
	(0,09)	(0,26)
IND5	0.491***	1.706***
	(0,09)	(0,25)
IND6	0.996***	3.120***
	(0,10)	(0,29)
IND7	0,14	0.625*
	(0,11)	(0,30)
IND8	0.333**	1.128***
	(0,11)	(0,30)
IND9	0.385**	1.315**
	(0,14)	(0,40)
IND10	0.398***	1.279***
	(0,09)	(0,25)
DIV	-0.462***	-0.631***
	(0,04)	(0,07)
EQTAPOS	-5.715***	-10.308***
	(0,22)	(0,51)
EQTAPOSSQ	8.513***	17.391***
	(0,24)	(0,61)
EQTANEG	(0,48)	-1.450*
	(0,25)	(0,57)
EQTANEGSQ	-1.139***	-2.320***
	(0,26)	(0,57)
LNSIZE	-0.717***	-2.467***
	(0,12)	(0,30)
LNSIZE^2	0.037***	0.138***
	(0,01)	(0,02)
expCASH/L	(0,00)	-0.316*
	(0,07)	(0,15)
EBIE/ATA POS	-1.061**	(0,80)
	(0,33)	(0,56)
EBIE/ATA POSSQ	4.796***	6.437***
	(0,51)	(0,93)
EBIE/ATA NEG	-9.335***	-13.607***
	(0,58)	(1,10)
EBIE/ATA NEGSQ	-22.709***	-33.486***
	(2,42)	(4,90)
EBIE/ATA NEGCU	-18.412***	-28.692***
	(2,59)	(5,47)
CL/TA	0.694***	1.887***
	(0,07)	(0,17)
constant	-3.068***	-8.202***
	(0,45)	(1,14)
Firm Random Effects		39.652***
		(2,22)
Observations	843979	843979
BIC	64978	61388
AIC	64652	61050

Composite Model	Simple Logit	w. Random Effects
Baseline Hazard	15.192***	49.532***
	(2,75)	(4,03)
EBIE/ATA	-2.143***	-3.301***
	(0,08)	(0,15)
RE/TA	-0.641***	-1.665***
	(0,06)	(0,13)
BEQ/TA	0.519***	1.424***
	(0,07)	(0,14)
IE/TL	6.262***	10.907***
	(0,18)	(0,40)
AUDNR	0.847***	1.502***
	(0,06)	(0,10)
LNA	-0.240***	-0.464***
	(0,02)	(0,05)
LNS	0.163***	0.395***
	(0,02)	(0,04)
DPO	0.004***	0.005***
	0,00	0,00
UTR/EMP	-0.007***	-0.012***
	0,00	0,00
constant	-5.140***	-12.014***
	(0,15)	(0,36)
Firm Random Effects		20.648***
		(0,97)
Observations	843 461	843 461
BIC	66 540	64 007
AIC	66 412	63 867

This table reports results from the binary logit regressions (simple logit and logit with random effects) for the Composite Model and all competing models using Swedish, private, independent, non financial/real estate limited liabilities. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion.

Table A5: AUROC statistical Significance by Sample

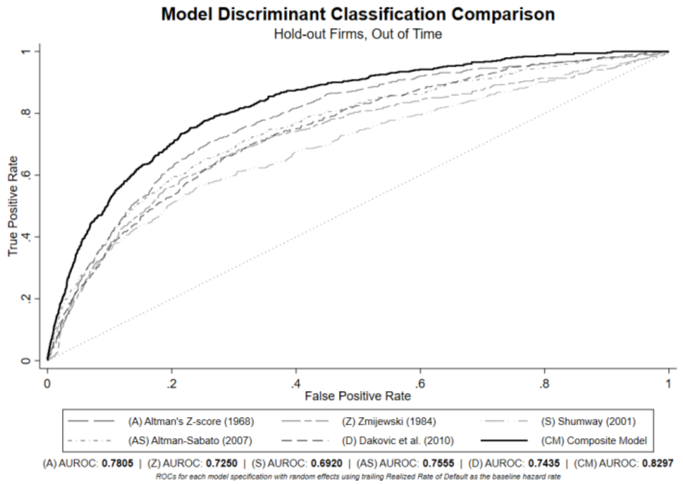
	Sampled Firms, In-Time										**Hold-Out Firms, In-Time**									
	Simple Logit					Simple Logit with Random Effects					Simple Logit					Simple Logit with Random Effects				
	One-Year	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**
Altman (1968)	73,37%	71,79%	67,74%	65,14%	62,96%	73,76%	72,21%	67,74%	65,14%	62,96%	73,85%	71,50%	67,94%	65,19%	62,72%	74,24%	72,06%	67,94%	65,19%	62,72%
Zmijewski (1984)	65,54%	68,34%	65,01%	61,89%	59,31%	65,62%	68,23%	65,01%	61,89%	59,31%	66,42%	68,40%	65,09%	62,29%	58,78%	66,84%	68,38%	65,09%	62,29%	58,78%
Shumway (2001)	65,88%	67,96%	66,26%	63,91%	62,12%	65,85%	67,83%	66,26%	63,91%	62,12%	66,13%	67,69%	66,21%	64,06%	61,94%	66,18%	67,90%	66,21%	64,06%	61,94%
Altman and Sabato (2007)	71,62%	72,17%	68,27%	65,42%	62,82%	71,72%	72,46%	68,27%	65,42%	62,84%	72,42%	71,81%	68,40%	65,24%	62,27%	72,46%	72,19%	68,40%	65,24%	62,27%
Dakovic et al. (2010)	78,51%	76,48%	72,21%	69,24%	66,88%	76,12%	76,69%	72,21%	69,24%	66,88%	78,14%	76,23%	72,17%	69,13%	66,39%	76,10%	76,48%	72,16%	69,13%	66,39%
Composite Model	74.91%†††	75.44%†††	72.74%***	70.2%***	67.48%***	74.87%†††	75.44%†††	72.74%***	70.2%***	67.48%***	75.65%†††	75.24%†††	72.46%***	69.95%***	67.16%***	75.7%†††	75.26%†††	72.45%***	69.95%***	67.16%***

	Sampled Firms, Out-of-Time										**Hold-Out Firms, Out-of-Time**									
	Simple Logit					Simple Logit with Random Effects					Simple Logit					Simple Logit with Random Effects				
	One-Year	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**	**One-Year**	**Two-Year**	**Three-Year**	**Four-Year**	**Five-Year**
Altman (1968)	77,01%	73,16%	71,26%	68,36%	65,78%	76,91%	73,77%	71,26%	68,36%	65,78%	78,84%	73,79%	69,90%	67,46%	66,38%	78,05%	74,24%	69,90%	67,46%	66,38%
Zmijewski (1984)	69,64%	69,43%	67,84%	63,87%	61,27%	68,75%	69,37%	67,84%	63,87%	61,27%	73,95%	70,06%	66,55%	64,10%	61,93%	72,50%	69,82%	66,55%	64,10%	61,93%
Shumway (2001)	69,67%	70,13%	69,06%	65,74%	64,35%	67,92%	70,25%	69,06%	65,74%	64,35%	71,99%	71,54%	68,08%	66,14%	64,21%	69,20%	71,59%	68,08%	66,14%	64,21%
Altman and Sabato (2007)	74,79%	72,59%	71,72%	68,26%	65,89%	74,30%	72,98%	71,72%	68,26%	65,89%	76,91%	73,59%	70,17%	67,26%	65,35%	75,55%	73,89%	70,17%	67,26%	65,35%
Dakovic et al. (2010)	81,02%	78,02%	74,89%	71,87%	68,92%	76,39%	77,83%	74,89%	71,87%	68,92%	79,57%	77,61%	74,39%	71,46%	69,73%	74,35%	77,28%	74,39%	71,46%	69,73%
Composite Model	80.54%†††	79.46%***	76.27%***	73.07%***	69.92%***	80.58%***	79.49%***	76.27%***	73.07%***	69.92%***	83.12%***	79.64%***	76.41%***	72.88%***	70.64%***	82.97%***	79.63%***	76.41%***	72.88%***	70.64%***

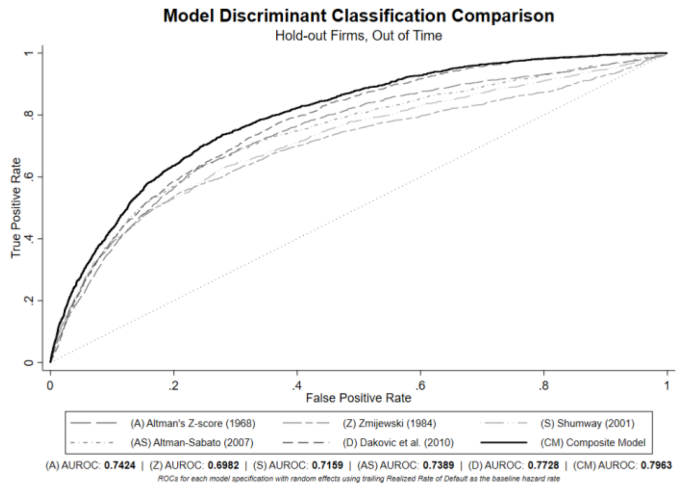
*This table shows the AUROC values. */ ** /*** Composite Model is significantly superior at the 10% / 5% / 1% level vis-à-vis all other models. † / †† / ††† Composite Model is significantly superior at the 10% / 5% / 1% level vis-à-vis all models except for the highest scoring AUROC/AUPRC model. All model specifications' AUROC and AURPC are statistically above their respective reference lines (i.e. better than random).*

Table A6: ROC Hold-Out Firm, Out-of-Time Model Comparison

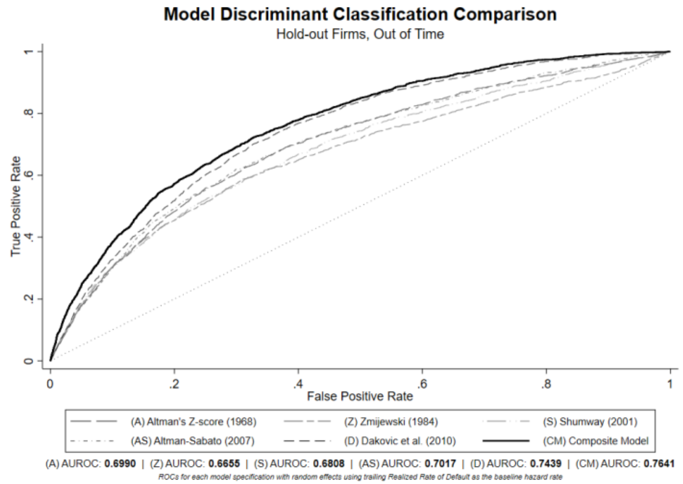
One-Year



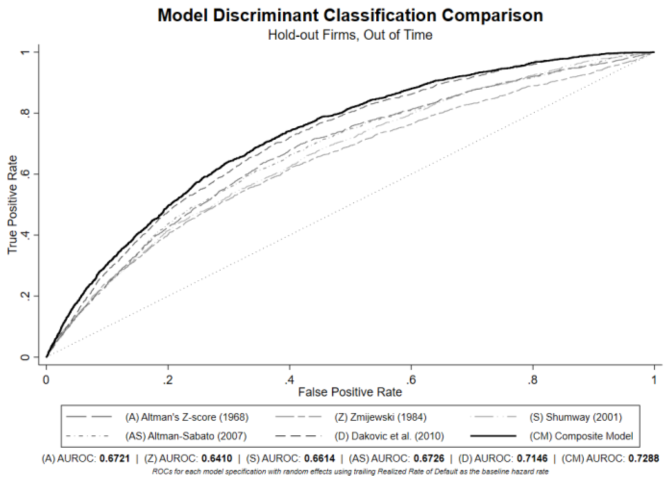
Two-Year



Three-Year



Four-Year



Five-Year

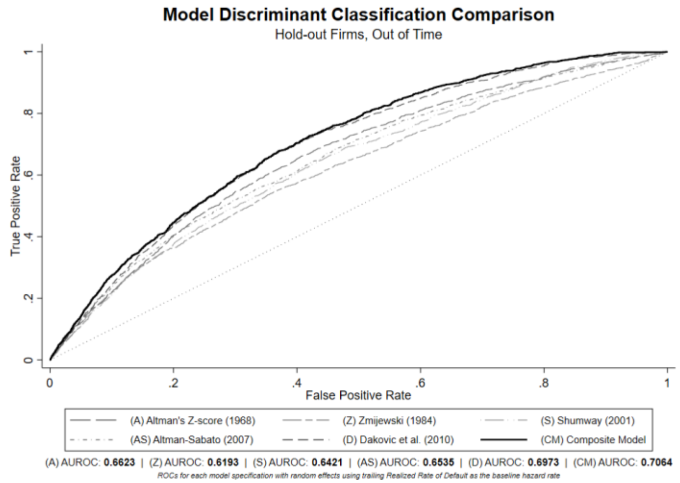
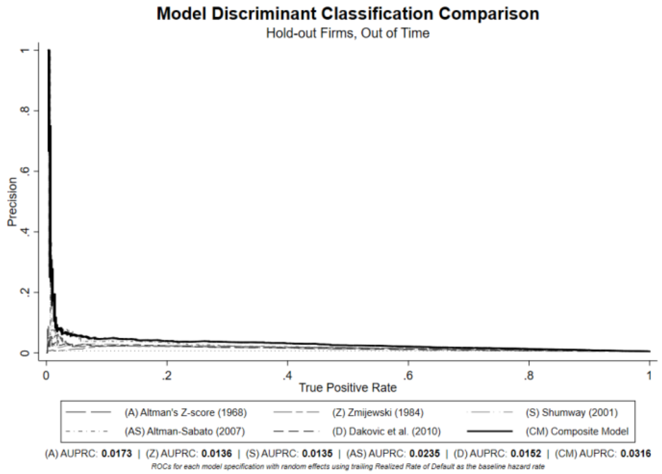
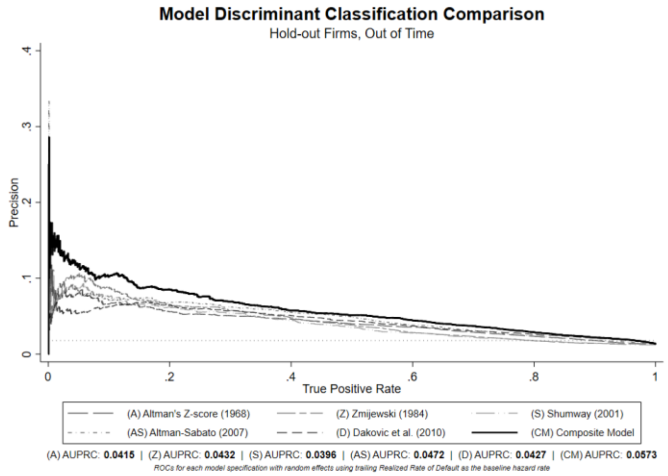


Table A7: PR Hold-Out Firm, Out-of-Time Model Comparison

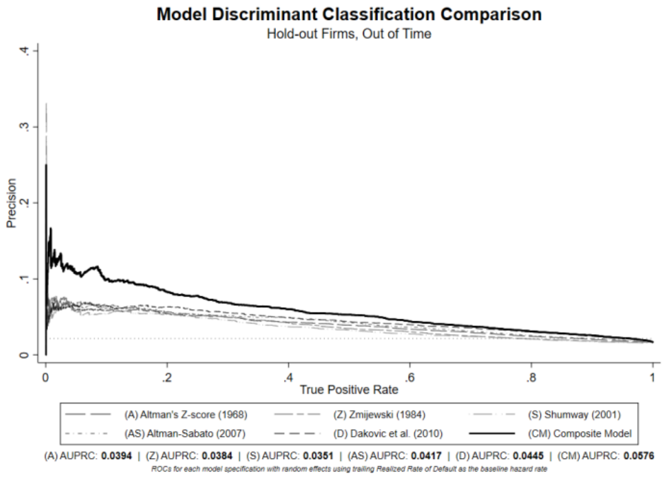
One-Year



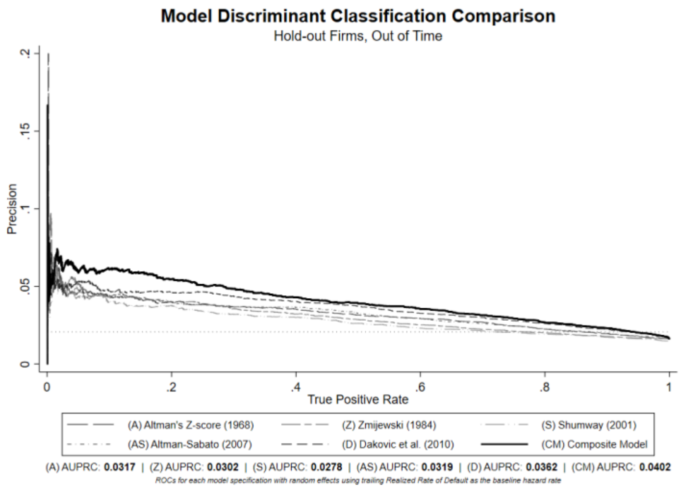
Two-Year



Three-Year



Four-Year



Five-Year

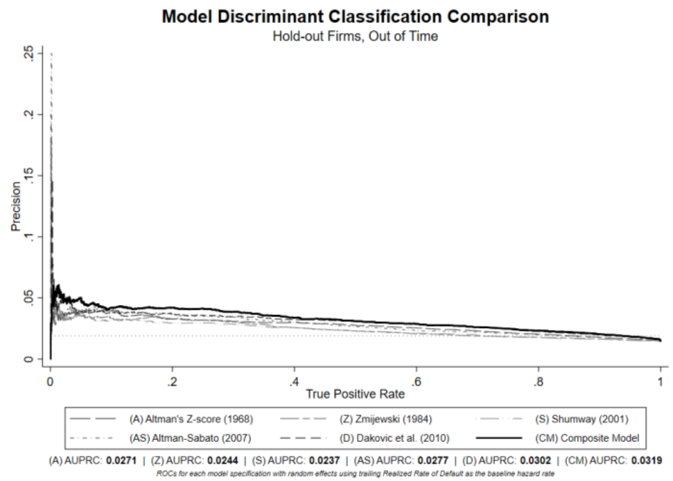


Table 10.1: Average Industry Margins by Year, 1998-2017

Gross Margins, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	41,1%	51,3%	41,1%	47,7%	33,7%	26,0%	60,1%	58,0%	48,0%	60,6%	36,0%	*45,3%*
1999	42,1%	51,4%	40,6%	47,9%	33,6%	26,2%	60,2%	57,5%	47,1%	60,7%	35,5%	*45,4%*
2000	43,2%	50,9%	40,5%	48,2%	33,4%	26,7%	60,5%	56,7%	46,6%	60,3%	35,5%	*45,4%*
2001	40,0%	49,6%	39,8%	46,1%	32,1%	25,5%	59,2%	55,7%	43,1%	57,5%	34,3%	*43,8%*
2002	38,9%	49,0%	38,9%	44,6%	31,5%	25,6%	58,8%	54,1%	41,1%	55,8%	33,3%	*42,8%*
2003	39,5%	48,2%	38,7%	43,9%	31,2%	26,1%	58,9%	52,5%	39,4%	55,2%	32,4%	*42,3%*
2004	39,1%	47,9%	38,2%	43,1%	31,0%	26,3%	57,7%	53,5%	39,2%	54,1%	31,7%	*41,9%*
2005	36,8%	47,6%	38,1%	42,7%	30,7%	26,2%	57,5%	53,6%	39,1%	53,5%	31,4%	*41,5%*
2006	36,9%	46,6%	38,1%	42,9%	31,1%	26,1%	58,0%	53,9%	38,7%	54,1%	31,1%	*41,9%*
2007	37,2%	46,2%	37,8%	43,7%	31,6%	27,1%	58,0%	55,1%	40,3%	54,3%	31,7%	*42,5%*
2008	36,1%	46,2%	37,6%	44,0%	31,7%	26,5%	58,4%	55,7%	39,7%	54,5%	31,4%	*42,8%*
2009	36,8%	46,7%	37,2%	44,0%	32,2%	26,9%	59,2%	56,8%	40,1%	54,6%	30,2%	*43,1%*
2010	37,9%	46,6%	37,9%	44,3%	33,1%	27,7%	59,7%	57,5%	41,4%	55,1%	34,0%	*43,8%*
2011	38,2%	46,7%	38,4%	44,7%	34,0%	27,9%	60,6%	59,3%	42,7%	55,4%	32,3%	*44,5%*
2012	37,3%	46,3%	38,6%	45,1%	34,9%	28,0%	61,4%	60,2%	43,2%	55,9%	33,1%	*45,1%*
2013	37,8%	47,2%	38,8%	45,4%	35,7%	28,0%	61,9%	60,4%	42,9%	56,3%	33,3%	*45,6%*
2014	36,5%	47,8%	39,1%	45,8%	36,3%	28,1%	63,0%	60,2%	43,6%	57,0%	33,5%	*46,1%*
2015	38,3%	48,6%	39,2%	46,4%	36,9%	27,1%	63,4%	60,8%	44,0%	57,7%	34,2%	*46,7%*
2016	39,8%	49,1%	39,4%	46,4%	37,7%	27,8%	63,7%	59,7%	43,3%	58,6%	34,8%	*47,2%*
2017	38,6%	49,0%	39,2%	46,7%	38,6%	29,0%	63,6%	60,8%	44,2%	58,9%	36,2%	*47,8%*
Total	*38,6%*	*48,0%*	*38,9%*	*45,2%*	*33,5%*	*26,9%*	*60,3%*	*57,2%*	*42,4%*	*56,5%*	*33,3%*	*44,3%*

EBITDA Margins, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	10,7%	16,8%	8,0%	7,3%	5,5%	3,6%	10,1%	10,8%	7,9%	11,0%	12,6%	*8,1%*
1999	11,2%	15,4%	7,9%	7,5%	5,6%	3,5%	10,2%	9,4%	6,5%	10,6%	12,2%	*8,0%*
2000	11,8%	15,8%	8,0%	7,6%	5,4%	3,4%	10,9%	7,2%	7,1%	10,3%	13,0%	*8,0%*
2001	13,2%	16,0%	7,3%	7,9%	5,1%	3,4%	9,7%	6,3%	6,4%	9,5%	13,0%	*7,5%*
2002	13,1%	16,6%	6,9%	7,4%	5,0%	3,4%	9,7%	5,1%	5,4%	8,7%	12,5%	*7,2%*
2003	12,9%	16,6%	6,6%	7,0%	4,6%	3,3%	10,1%	4,9%	5,1%	8,6%	11,8%	*6,9%*
2004	13,6%	16,9%	7,0%	7,0%	4,4%	3,0%	9,7%	6,2%	5,3%	8,9%	11,6%	*7,0%*
2005	12,2%	16,9%	7,3%	7,4%	4,3%	3,0%	8,8%	7,2%	4,8%	9,0%	12,1%	*7,1%*
2006	13,6%	16,4%	8,0%	8,3%	4,7%	3,3%	9,5%	7,4%	5,9%	10,1%	12,2%	*7,8%*
2007	11,6%	16,2%	8,3%	8,8%	4,3%	3,5%	9,8%	7,8%	4,9%	10,6%	13,4%	*8,0%*
2008	10,5%	15,2%	7,1%	8,4%	3,7%	3,0%	10,1%	6,8%	2,9%	10,1%	12,2%	*7,4%*
2009	8,9%	15,2%	5,0%	7,5%	4,0%	3,2%	10,8%	7,1%	3,1%	9,3%	10,6%	*7,0%*
2010	11,4%	15,2%	6,7%	8,1%	4,7%	3,2%	10,8%	8,2%	4,0%	10,4%	11,2%	*7,8%*
2011	12,0%	15,5%	7,4%	8,4%	4,7%	3,2%	11,0%	9,0%	4,8%	10,5%	11,0%	*8,0%*
2012	10,2%	13,9%	6,3%	7,5%	4,5%	3,0%	10,7%	8,6%	4,3%	9,9%	11,0%	*7,5%*
2013	9,8%	14,5%	6,1%	7,6%	4,7%	3,1%	11,3%	9,0%	3,9%	10,3%	11,0%	*7,7%*
2014	10,4%	15,0%	6,6%	8,0%	5,1%	3,3%	12,3%	8,8%	4,7%	10,7%	11,1%	*8,1%*
2015	10,3%	14,7%	7,0%	9,0%	5,5%	2,8%	12,9%	10,0%	3,8%	11,1%	11,5%	*8,6%*
2016	12,7%	14,9%	7,0%	8,6%	5,6%	3,2%	13,0%	9,1%	3,8%	11,9%	11,4%	*8,8%*
2017	12,0%	14,9%	6,9%	8,4%	5,5%	3,0%	12,1%	9,8%	2,3%	11,4%	12,2%	*8,6%*
Total	*11,6%*	*15,6%*	*7,1%*	*7,9%*	*4,8%*	*3,2%*	*10,8%*	*7,9%*	*42,4%*	*10,2%*	*11,8%*	*7,7%*

EBIT Margins, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	5,4%	6,3%	4,8%	4,1%	2,9%	1,8%	5,9%	7,7%	3,3%	5,8%	5,4%	*4,4%*
1999	5,5%	4,8%	4,5%	4,3%	2,9%	1,6%	5,9%	6,3%	1,8%	5,4%	4,7%	*4,1%*
2000	6,0%	5,1%	4,7%	4,4%	2,7%	1,6%	6,7%	4,0%	2,1%	5,2%	5,5%	*4,1%*
2001	7,5%	5,0%	3,9%	4,6%	2,4%	1,6%	5,7%	2,9%	1,1%	4,4%	5,3%	*3,7%*
2002	7,3%	6,0%	3,4%	4,0%	2,4%	1,6%	5,6%	1,4%	0,6%	3,6%	4,7%	*3,3%*
2003	6,4%	6,1%	3,2%	3,5%	1,9%	1,5%	6,1%	1,1%	0,2%	3,6%	3,8%	*3,0%*
2004	7,4%	6,7%	3,6%	3,7%	1,6%	1,2%	5,7%	2,6%	0,9%	4,1%	3,7%	*3,2%*
2005	6,0%	6,7%	4,2%	4,2%	1,5%	1,2%	4,9%	3,8%	0,7%	4,4%	4,2%	*3,4%*
2006	7,6%	5,4%	5,1%	5,3%	2,0%	1,6%	5,8%	4,4%	1,7%	5,8%	4,6%	*4,2%*
2007	5,2%	5,9%	5,4%	5,8%	1,6%	1,7%	6,3%	5,2%	1,1%	6,5%	6,3%	*4,6%*
2008	4,1%	4,3%	4,1%	5,4%	1,0%	1,3%	6,8%	4,1%	-0,8%	6,0%	5,1%	*4,0%*
2009	1,9%	4,0%	1,6%	4,3%	1,3%	1,5%	7,7%	4,5%	-0,6%	5,1%	3,0%	*3,4%*
2010	5,0%	4,9%	3,6%	5,2%	2,0%	1,6%	7,8%	5,7%	0,6%	6,5%	4,8%	*4,5%*
2011	5,8%	5,7%	4,5%	5,6%	2,2%	1,6%	8,3%	6,9%	1,9%	6,9%	3,9%	*4,9%*
2012	3,4%	3,6%	3,5%	4,8%	2,1%	1,4%	8,0%	6,6%	1,6%	6,5%	4,0%	*4,4%*
2013	3,0%	4,1%	3,3%	4,9%	2,3%	1,5%	8,8%	7,0%	1,2%	6,9%	4,1%	*4,7%*
2014	2,8%	4,6%	3,8%	5,3%	2,7%	1,6%	9,8%	6,7%	2,1%	7,4%	4,3%	*5,1%*
2015	2,5%	4,4%	4,4%	6,4%	3,2%	1,2%	10,5%	8,0%	1,6%	8,1%	4,9%	*5,8%*
2016	5,9%	4,6%	4,4%	6,1%	3,3%	1,5%	10,6%	6,8%	1,1%	9,0%	4,7%	*5,9%*
2017	6,3%	5,1%	4,4%	5,9%	3,3%	1,4%	9,7%	7,5%	-0,1%	8,4%	5,9%	*5,8%*
Total	*5,2%*	*5,1%*	*4,0%*	*5,0%*	*2,3%*	*1,5%*	*7,5%*	*5,2%*	*42,4%*	*6,0%*	*4,6%*	*4,3%*

Net Income Margins, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	2,4%	2,4%	2,3%	1,9%	1,2%	0,7%	3,3%	4,6%	1,9%	3,1%	2,6%	*2,1%*
1999	2,4%	2,3%	2,3%	2,1%	1,3%	0,9%	3,8%	4,2%	1,5%	3,4%	2,7%	*2,3%*
2000	2,7%	1,8%	2,2%	2,0%	0,9%	0,7%	3,6%	1,6%	0,9%	2,8%	2,5%	*1,9%*
2001	3,3%	1,1%	1,3%	1,7%	0,3%	0,2%	2,1%	-0,4%	-0,9%	1,3%	2,0%	*1,0%*
2002	2,7%	1,4%	0,9%	1,2%	0,2%	0,1%	1,6%	-1,4%	-1,3%	0,6%	1,0%	*0,6%*
2003	2,9%	2,2%	1,4%	1,5%	0,3%	0,5%	3,7%	0,7%	-0,4%	2,1%	1,8%	*1,3%*
2004	4,0%	3,0%	2,1%	2,0%	0,5%	0,5%	4,4%	2,6%	0,6%	3,1%	2,4%	*1,9%*
2005	4,4%	3,3%	2,8%	2,7%	0,7%	0,8%	4,5%	4,0%	1,7%	3,9%	3,0%	*2,5%*
2006	5,1%	2,9%	3,1%	3,1%	0,9%	0,8%	5,0%	4,0%	2,1%	4,5%	3,4%	*2,9%*
2007	2,4%	2,4%	3,1%	3,3%	0,4%	0,6%	4,6%	3,5%	0,7%	4,5%	3,3%	*2,7%*
2008	1,3%	1,4%	1,9%	2,6%	-0,5%	0,1%	3,6%	1,5%	-1,4%	3,3%	1,9%	*1,7%*
2009	1,2%	1,9%	0,9%	2,6%	0,1%	0,6%	5,5%	3,4%	-0,5%	3,9%	2,3%	*2,1%*
2010	2,7%	2,5%	2,1%	3,2%	0,8%	0,6%	5,6%	4,2%	0,5%	5,0%	3,0%	*2,9%*
2011	2,7%	2,0%	2,5%	3,2%	0,6%	0,4%	5,2%	4,4%	0,8%	4,6%	2,1%	*2,7%*
2012	1,2%	1,1%	1,8%	2,6%	0,5%	0,4%	5,1%	4,1%	0,7%	4,4%	2,2%	*2,5%*
2013	2,5%	2,1%	2,5%	3,3%	1,2%	0,8%	7,0%	6,0%	1,5%	5,9%	3,0%	*3,5%*
2014	1,7%	2,8%	3,0%	3,9%	1,6%	0,9%	8,2%	6,5%	2,4%	6,6%	3,7%	*4,1%*
2015	1,8%	3,3%	3,2%	4,6%	2,1%	0,8%	8,7%	7,1%	2,1%	7,3%	3,9%	*4,6%*
2016	4,4%	3,4%	3,6%	4,5%	2,2%	1,2%	8,8%	6,7%	1,9%	8,2%	4,4%	*4,9%*
2017	4,3%	3,5%	3,4%	4,3%	2,3%	1,0%	8,3%	7,2%	1,4%	7,7%	5,0%	*4,8%*
Total	*2,8%*	*2,3%*	*2,2%*	*2,9%*	*0,9%*	*0,6%*	*5,3%*	*3,8%*	*42,4%*	*4,4%*	*2,8%*	*2,7%*

The table reports average gross, EBITDA, EBIT and net income margins, for all active Swedish, non-financial, independent limited liability companies with at least two years of data. The results are presented by year and industry. Companies with missing industry classification are not identifiable in the table, but are included in the totals. Gross margins are not based on actual COGS due to the fact that 98% of the companies report their financials by nature, not function. Consequently, COGS are approximated as production costs, where production costs are defined as raw materials and consumables + goods for resale + other external costs.

Table 10.2: Average Industry Activity by Year, 1998-2017

Days of Inventory Outstanding

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	28,5	23,4	47,4	25,7	53,7	29,0	10,8	15,3	26,4	9,6	72,9	*33,9*
1999	28,1	24,1	46,0	24,4	51,9	29,5	11,5	14,9	23,4	9,5	69,6	*32,7*
2000	29,4	23,1	44,9	22,6	50,8	29,4	11,1	14,9	22,4	9,2	64,3	*31,3*
2001	24,8	22,0	45,5	21,6	50,9	29,1	10,8	15,0	23,9	9,1	61,8	*31,0*
2002	28,8	21,8	45,7	20,9	49,8	27,9	10,2	13,9	22,2	8,9	61,1	*30,3*
2003	28,3	20,0	46,6	20,3	49,7	28,2	9,5	14,1	24,0	8,8	61,8	*30,3*
2004	27,7	20,0	45,9	19,7	49,9	30,6	9,9	13,9	25,4	8,8	62,6	*30,2*
2005	29,1	19,8	44,9	18,9	49,5	29,8	9,5	13,3	24,3	8,6	64,8	*29,7*
2006	28,3	18,3	43,4	17,2	47,6	29,2	9,2	12,5	23,5	8,0	64,2	*28,2*
2007	27,7	17,3	43,3	16,3	47,6	29,3	9,6	11,7	22,3	8,1	62,1	*27,5*
2008	29,1	17,4	44,8	15,5	48,6	28,8	9,4	10,9	20,7	8,2	62,6	*27,4*
2009	31,0	19,7	50,3	15,2	48,9	29,5	9,1	11,4	21,3	8,3	68,9	*28,2*
2010	28,5	18,3	47,3	14,5	47,3	29,4	9,0	11,1	18,4	7,7	66,3	*27,0*
2011	32,6	17,7	45,2	12,8	46,1	29,9	8,4	9,9	17,2	7,5	62,9	*25,4*
2012	31,1	16,8	46,7	12,1	45,5	28,8	7,7	9,0	17,6	6,9	63,1	*24,9*
2013	32,7	16,5	48,2	11,7	44,2	29,4	7,5	8,7	15,8	6,8	62,1	*24,4*
2014	33,2	15,5	46,1	10,5	41,9	28,6	7,5	8,0	16,6	6,4	61,5	*23,1*
2015	31,3	16,2	44,9	9,6	40,2	28,5	6,7	8,0	13,7	6,3	61,7	*22,1*
2016	30,5	14,8	44,0	9,0	38,5	29,5	6,8	7,3	13,2	6,0	61,2	*21,3*
2017	32,0	16,0	41,9	8,3	36,0	27,0	6,4	6,8	13,0	5,5	57,5	*19,8*
Total	*29,6*	*18,9*	*45,7*	*15,6*	*47,0*	*29,1*	*8,9*	*11,3*	*19,9*	*7,9*	*63,7*	*27,4*

Days of Sales Outstanding

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	34,2	38,3	47,6	48,2	21,4	12,5	27,5	54,3	40,7	45,8	28,9	*36,5*
1999	37,2	39,3	47,6	48,1	21,2	12,5	27,8	56,3	45,2	46,5	29,7	*37,0*
2000	38,3	38,6	48,0	48,6	21,4	12,9	29,0	59,5	43,7	47,7	32,7	*37,6*
2001	37,3	37,0	47,7	48,1	21,0	13,0	28,5	53,1	43,4	46,4	30,4	*36,6*
2002	34,6	38,1	48,2	48,4	21,0	12,6	27,7	54,1	42,2	46,6	30,0	*36,6*
2003	34,9	38,1	47,7	48,1	20,4	13,0	27,0	53,8	43,6	45,6	29,8	*36,0*
2004	34,5	36,8	47,4	49,2	20,7	13,0	26,9	55,2	43,2	46,3	31,9	*36,3*
2005	38,4	37,7	48,5	50,3	21,0	13,7	28,4	55,9	44,8	48,3	33,7	*37,5*
2006	38,3	39,6	49,1	51,3	20,7	13,2	28,6	58,5	45,6	49,1	32,7	*38,0*
2007	36,6	40,0	47,8	50,2	21,1	14,2	29,4	60,0	46,8	49,5	30,0	*38,1*
2008	36,0	40,2	46,2	47,8	20,5	13,3	27,8	57,7	43,8	48,2	30,7	*37,0*
2009	37,9	39,2	48,2	48,0	20,5	13,1	26,0	56,8	47,3	49,1	31,9	*37,4*
2010	39,9	38,3	48,0	51,8	21,1	13,6	27,3	63,0	47,6	51,4	62,0	*40,4*
2011	36,2	38,3	46,4	48,4	20,2	13,8	26,5	59,4	46,4	50,2	32,5	*37,8*
2012	36,3	36,4	46,4	46,2	19,5	14,0	25,2	57,6	46,5	48,3	31,5	*36,5*
2013	38,1	36,3	46,9	46,3	19,2	13,4	24,0	59,4	43,4	48,8	30,8	*36,5*
2014	36,2	35,5	45,6	44,8	18,3	13,6	23,7	57,5	41,7	48,1	31,2	*35,6*
2015	36,2	36,3	45,0	43,9	17,6	13,2	25,2	57,1	42,0	48,7	32,2	*35,5*
2016	33,0	34,4	45,2	43,3	17,2	13,1	23,4	56,6	42,0	47,2	30,5	*34,7*
2017	31,6	33,7	42,6	39,4	15,7	13,1	20,9	52,4	40,6	44,3	28,5	*32,3*
Total	*36,4*	*37,7*	*47,2*	*47,3*	*20,0*	*13,2*	*26,4*	*57,1*	*44,1*	*47,9*	*32,9*	*36,7*

Days of Payables Outstanding

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	27,8	26,1	33,3	29,2	28,7	19,9	13,9	25,2	30,0	20,6	36,5	*27,0*
1999	28,3	27,2	33,2	29,2	28,7	20,7	15,4	27,3	33,1	21,2	37,0	*27,4*
2000	29,2	26,9	32,7	28,6	28,6	20,8	15,7	27,8	31,9	21,6	37,3	*27,0*
2001	29,9	26,4	32,0	27,8	28,5	20,9	14,9	24,3	31,0	20,6	36,4	*26,2*
2002	28,6	25,6	31,4	27,7	28,3	20,6	14,0	23,8	29,1	20,1	36,8	*25,8*
2003	28,0	24,6	30,8	27,4	27,9	20,6	13,3	24,4	29,2	19,1	36,5	*25,2*
2004	27,3	23,9	30,7	28,2	28,6	21,9	13,9	24,4	30,1	19,7	37,2	*25,7*
2005	30,5	24,9	30,8	28,9	29,0	21,5	14,6	23,5	28,5	20,2	36,9	*26,0*
2006	28,1	25,3	30,3	28,3	28,2	21,8	14,0	22,9	28,8	19,4	36,1	*25,3*
2007	27,1	25,6	29,7	27,2	28,1	23,7	14,7	22,0	28,2	19,0	36,1	*24,9*
2008	29,1	24,1	28,5	25,5	27,8	23,4	13,6	20,8	27,9	18,3	36,5	*24,1*
2009	28,9	23,8	29,6	25,5	27,0	23,3	13,4	18,5	29,0	17,9	36,1	*23,7*
2010	27,4	22,4	28,4	26,8	26,5	22,3	13,9	19,8	26,9	17,9	47,3	*24,2*
2011	25,4	22,4	27,1	24,5	25,1	22,5	13,1	17,0	25,1	17,1	32,8	*22,2*
2012	24,3	21,8	26,5	22,6	23,9	22,0	11,5	15,5	21,7	16,1	32,0	*20,9*
2013	26,1	20,7	26,3	22,2	23,7	21,6	11,1	16,0	22,7	15,8	31,6	*20,7*
2014	27,2	20,2	25,4	21,4	21,8	20,9	10,8	15,5	22,2	15,2	30,7	*19,7*
2015	26,1	19,9	24,6	20,7	21,1	20,4	10,3	15,8	21,5	14,7	30,6	*19,1*
2016	22,0	19,0	24,8	20,6	20,2	20,2	10,0	15,7	21,5	14,1	29,2	*18,6*
2017	22,5	18,1	22,7	17,9	17,1	17,8	8,8	13,2	18,6	12,4	24,6	*16,1*
Total	*27,3*	*23,4*	*29,5*	*25,1*	*26,0*	*21,4*	*12,9*	*20,3*	*26,6*	*18,0*	*35,1*	*23,4*

Cash Conversion Cycle, Days

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	35,2	36,4	62,1	45,4	47,3	21,9	25,0	46,9	39,0	36,2	66,8	*44,5*
1999	38,4	37,3	60,9	44,2	45,6	21,7	25,2	46,8	36,5	36,5	63,8	*43,6*
2000	39,4	35,2	60,7	43,7	44,5	21,7	25,6	49,1	35,1	37,1	61,2	*43,1*
2001	33,1	33,6	61,7	42,9	44,4	21,4	25,4	46,1	38,6	36,6	57,3	*42,6*
2002	36,9	35,6	63,1	42,7	43,7	20,6	25,4	47,0	37,1	37,0	56,3	*42,5*
2003	37,8	33,9	64,5	42,3	43,7	21,0	24,5	47,0	41,2	37,2	56,9	*42,6*
2004	37,2	33,4	63,7	42,1	43,5	22,5	24,2	47,8	41,4	37,2	59,0	*42,4*
2005	37,0	33,7	63,6	42,1	43,4	22,4	24,8	48,8	43,5	39,1	63,4	*43,0*
2006	42,3	33,8	63,0	41,8	41,8	21,0	25,3	51,0	42,8	39,8	62,8	*42,6*
2007	37,5	33,0	62,3	41,0	42,5	20,5	25,8	52,6	43,6	40,6	58,2	*42,4*
2008	36,3	34,5	63,3	39,1	43,2	19,1	24,8	50,9	39,3	40,2	59,1	*42,0*
2009	40,6	36,2	70,2	39,7	44,2	19,5	23,1	52,2	43,1	41,7	67,0	*43,7*
2010	42,1	35,7	68,1	41,0	44,0	21,7	23,8	57,8	43,0	43,0	89,5	*45,3*
2011	46,6	35,7	65,8	38,2	43,1	22,0	23,1	55,3	41,2	42,6	65,3	*42,8*
2012	44,5	32,8	67,6	36,9	42,5	21,4	22,9	52,9	44,5	41,1	65,1	*41,9*
2013	46,2	33,5	69,6	37,1	41,2	22,1	21,8	54,7	40,1	41,8	63,9	*41,8*
2014	45,8	31,9	68,0	35,1	40,1	21,9	22,0	53,1	39,2	41,9	65,3	*40,9*
2015	43,4	33,5	66,5	34,0	38,3	21,9	23,0	52,6	37,2	42,2	66,5	*40,2*
2016	42,8	31,2	65,3	33,0	37,0	23,4	21,6	51,5	37,0	41,0	66,5	*39,1*
2017	41,0	32,0	62,2	30,1	35,1	22,3	18,6	47,5	36,7	38,6	62,8	*36,7*
Total	*40,2*	*34,2*	*64,4*	*39,1*	*42,5*	*21,5*	*23,7*	*50,9*	*40,0*	*39,7*	*64,2*	*42,2*

The table reports the activity ratios for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. All activity ratios are calculated with sales as the denominator and assuming 365 days in a year. CCC = DIO + DSO - DPO. The results are presented by year and industry. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.3: Average Industry Return on Capital Ratios by Year, 1998-2017

Return on Average Capital Employed, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	13,2%	12,3%	15,8%	17,0%	13,6%	13,5%	17,7%	25,3%	13,8%	18,0%	11,7%	15,8%
1999	15,1%	12,0%	17,1%	20,2%	15,2%	15,9%	20,5%	26,7%	14,0%	19,7%	11,5%	17,6%
2000	15,5%	12,7%	17,8%	20,8%	13,8%	15,5%	23,3%	19,4%	17,2%	19,4%	12,6%	17,3%
2001	20,3%	11,4%	14,8%	20,2%	12,2%	14,4%	19,4%	16,3%	7,6%	15,5%	11,5%	15,0%
2002	17,6%	12,8%	12,6%	17,1%	11,9%	15,7%	19,7%	12,9%	11,7%	14,2%	10,5%	13,9%
2003	14,9%	12,5%	12,2%	14,3%	9,8%	13,5%	21,2%	13,2%	10,9%	14,1%	9,0%	12,8%
2004	13,7%	13,8%	13,3%	15,7%	9,2%	11,1%	19,6%	16,9%	12,1%	15,2%	7,9%	13,1%
2005	14,6%	13,6%	15,8%	18,7%	10,1%	11,6%	17,9%	22,7%	15,5%	18,0%	8,8%	15,0%
2006	18,7%	12,0%	20,1%	24,0%	13,5%	15,7%	21,8%	25,6%	17,6%	22,7%	12,1%	19,1%
2007	14,0%	14,1%	21,2%	25,6%	12,1%	17,0%	22,8%	26,7%	19,8%	24,2%	13,0%	19,8%
2008	10,9%	10,7%	15,6%	21,8%	8,2%	14,1%	25,5%	25,5%	13,3%	22,4%	11,1%	16,9%
2009	11,3%	9,6%	8,9%	18,3%	9,6%	15,9%	27,9%	23,2%	8,8%	19,8%	9,6%	15,5%
2010	16,4%	11,6%	15,0%	22,2%	12,9%	14,8%	27,1%	26,7%	15,9%	23,9%	10,7%	18,9%
2011	15,5%	13,2%	17,4%	24,0%	12,5%	14,9%	27,1%	28,9%	19,8%	25,0%	11,1%	19,8%
2012	10,8%	9,3%	13,0%	20,7%	12,8%	13,3%	27,2%	27,6%	15,2%	23,0%	11,1%	18,2%
2013	11,1%	9,3%	12,2%	21,1%	13,5%	15,2%	29,2%	27,6%	18,3%	23,9%	11,9%	18,9%
2014	11,7%	11,7%	15,8%	24,2%	14,7%	16,7%	31,8%	29,6%	17,7%	25,3%	11,3%	20,9%
2015	14,1%	10,0%	17,9%	29,0%	16,9%	15,4%	32,2%	32,6%	17,9%	28,1%	13,7%	23,4%
2016	13,6%	10,9%	17,6%	25,9%	16,4%	17,3%	32,4%	32,1%	17,6%	30,8%	13,1%	23,4%
2017	14,1%	12,9%	19,0%	26,7%	16,7%	16,9%	32,1%	36,5%	18,7%	31,5%	15,6%	24,3%
Total	*14,4%*	*11,8%*	*15,6%*	*21,8%*	*12,7%*	*14,9%*	*25,3%*	*25,0%*	*15,4%*	*21,9%*	*11,3%*	*18,0%*

Return on Average Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	8,0%	7,7%	8,8%	7,8%	6,8%	6,8%	10,3%	12,4%	6,8%	9,3%	6,8%	8,1%
1999	8,8%	7,2%	9,2%	9,3%	7,8%	7,8%	12,2%	12,6%	6,5%	10,0%	6,8%	8,9%
2000	9,2%	7,5%	9,6%	9,6%	7,2%	7,8%	13,5%	9,7%	7,0%	9,7%	7,5%	8,9%
2001	12,3%	6,8%	7,9%	9,3%	6,3%	6,9%	11,5%	8,0%	4,2%	8,0%	7,4%	7,7%
2002	10,7%	8,2%	7,1%	8,1%	6,4%	7,8%	11,3%	5,9%	4,6%	7,3%	6,5%	7,3%
2003	9,2%	8,1%	6,7%	7,1%	5,3%	6,9%	12,3%	6,1%	4,6%	7,2%	5,5%	6,8%
2004	9,4%	8,8%	7,5%	7,5%	4,7%	5,9%	11,6%	7,9%	4,8%	7,9%	5,4%	6,9%
2005	9,2%	9,1%	8,8%	8,9%	5,4%	5,5%	10,4%	10,0%	6,2%	9,2%	5,9%	7,7%
2006	11,0%	8,0%	10,7%	11,3%	6,8%	7,7%	12,4%	12,3%	7,7%	11,6%	7,0%	9,7%
2007	8,6%	8,4%	11,6%	12,4%	5,8%	7,8%	12,7%	12,8%	8,2%	12,4%	8,3%	10,0%
2008	6,5%	7,3%	9,2%	11,2%	4,0%	6,6%	14,4%	12,0%	5,4%	11,6%	7,2%	8,8%
2009	6,3%	6,3%	5,0%	9,3%	4,9%	7,4%	16,0%	11,6%	5,0%	10,3%	6,1%	8,1%
2010	9,5%	7,3%	8,3%	11,2%	6,6%	7,5%	15,7%	13,3%	7,3%	12,4%	6,3%	9,9%
2011	9,0%	8,2%	10,1%	12,3%	6,8%	7,2%	15,8%	15,0%	8,8%	13,2%	6,8%	10,6%
2012	6,5%	6,1%	7,7%	10,3%	6,7%	7,3%	15,7%	14,6%	7,6%	12,4%	6,7%	9,8%
2013	6,9%	6,6%	7,6%	10,9%	7,2%	7,9%	17,3%	15,6%	8,3%	13,2%	7,2%	10,4%
2014	6,8%	7,8%	9,1%	12,7%	8,4%	8,4%	19,3%	16,8%	9,5%	14,4%	7,5%	11,8%
2015	6,9%	6,8%	10,4%	15,3%	9,7%	8,3%	20,1%	18,3%	10,6%	16,0%	8,5%	13,3%
2016	9,7%	7,3%	10,5%	13,8%	9,5%	8,7%	20,8%	18,7%	10,1%	17,3%	8,9%	13,4%
2017	9,6%	8,5%	11,2%	14,6%	9,7%	8,4%	20,4%	20,1%	9,6%	17,8%	10,5%	13,9%
Total	*8,7%*	*7,6%*	*8,7%*	*10,9%*	*6,8%*	*7,4%*	*15,0%*	*12,9%*	*7,3%*	*11,7%*	*7,1%*	*9,6%*

Return on Average Equity, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	16,3%	13,6%	15,5%	14,8%	13,8%	15,0%	15,6%	20,3%	11,9%	15,9%	11,3%	14,8%
1999	14,2%	9,0%	15,4%	16,9%	12,2%	15,6%	16,9%	20,3%	9,6%	15,2%	8,3%	14,3%
2000	15,1%	11,7%	15,6%	16,1%	10,5%	14,9%	18,1%	9,9%	13,5%	14,1%	10,4%	13,4%
2001	20,7%	9,7%	9,0%	14,6%	6,3%	8,9%	14,3%	3,5%	2,3%	8,3%	6,5%	8,9%
2002	14,3%	12,8%	9,7%	11,4%	6,5%	12,9%	13,4%	3,8%	0,9%	8,3%	5,7%	8,7%
2003	8,4%	10,4%	8,4%	10,9%	4,9%	12,1%	18,9%	7,6%	10,0%	10,6%	6,3%	9,1%
2004	13,3%	17,6%	11,4%	14,6%	5,4%	9,9%	17,2%	13,0%	4,3%	12,4%	5,3%	10,6%
2005	15,1%	16,8%	15,4%	17,3%	7,8%	11,8%	14,7%	17,2%	9,2%	15,5%	9,0%	13,1%
2006	23,3%	15,0%	18,1%	21,1%	9,4%	12,6%	17,3%	17,1%	8,6%	19,2%	10,3%	15,6%
2007	13,9%	16,0%	17,9%	21,2%	5,9%	13,1%	18,4%	17,2%	12,6%	19,7%	13,9%	15,1%
2008	11,7%	9,0%	10,1%	16,5%	1,7%	8,1%	18,3%	14,1%	4,8%	16,4%	7,3%	10,8%
2009	12,4%	9,6%	5,5%	16,6%	6,0%	12,5%	24,6%	17,5%	0,6%	16,8%	8,3%	12,5%
2010	15,6%	15,3%	12,9%	20,3%	7,9%	11,5%	22,9%	19,5%	9,9%	20,1%	7,5%	15,2%
2011	11,4%	13,4%	14,2%	19,7%	5,3%	9,7%	20,9%	21,8%	14,3%	19,3%	3,9%	14,2%
2012	14,0%	7,9%	10,3%	17,1%	5,5%	9,3%	20,5%	18,3%	8,2%	17,7%	5,8%	12,8%
2013	11,3%	11,3%	10,9%	19,0%	7,7%	12,0%	28,0%	21,3%	10,4%	21,7%	5,2%	15,6%
2014	15,8%	16,0%	15,7%	23,6%	10,1%	13,8%	31,1%	26,3%	11,8%	23,8%	9,9%	18,8%
2015	13,6%	11,5%	15,1%	27,9%	11,7%	11,9%	29,9%	27,0%	12,4%	26,4%	12,4%	20,6%
2016	14,9%	15,0%	15,2%	25,4%	11,4%	12,9%	29,7%	26,7%	10,6%	28,9%	13,6%	20,8%
2017	13,2%	16,1%	17,9%	25,3%	10,2%	9,3%	28,9%	29,6%	10,1%	27,4%	15,5%	20,4%
Total	*14,4%*	*12,8%*	*13,1%*	*19,0%*	*8,0%*	*12,0%*	*21,5%*	*17,9%*	*8,9%*	*18,1%*	*8,7%*	*14,3%*

EBITDA Return on Average Total Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	14,3%	17,6%	13,2%	12,7%	11,5%	11,5%	16,9%	16,1%	12,8%	15,8%	13,3%	13,4%
1999	15,4%	17,1%	13,8%	14,1%	12,5%	11,8%	18,1%	15,4%	12,1%	16,0%	13,3%	14,1%
2000	15,8%	17,5%	14,1%	14,5%	11,9%	11,5%	19,1%	12,4%	12,6%	15,6%	13,8%	13,9%
2001	18,4%	17,3%	12,5%	14,3%	11,2%	11,6%	17,2%	11,6%	10,5%	14,1%	13,6%	13,0%
2002	16,9%	18,2%	11,6%	13,1%	11,2%	12,4%	17,2%	9,6%	10,3%	13,3%	12,8%	12,6%
2003	15,6%	18,1%	11,1%	12,0%	10,0%	11,4%	17,5%	9,6%	10,1%	12,8%	11,8%	11,8%
2004	15,9%	18,7%	11,8%	12,2%	9,5%	10,4%	16,4%	11,1%	10,0%	13,3%	11,6%	11,8%
2005	15,1%	19,0%	12,8%	13,4%	9,9%	9,9%	15,0%	12,7%	10,8%	14,1%	11,8%	12,4%
2006	16,9%	17,7%	14,5%	15,4%	11,1%	11,8%	16,6%	14,5%	11,8%	16,1%	12,5%	14,0%
2007	15,0%	17,9%	15,0%	16,4%	10,0%	11,8%	16,7%	14,6%	11,9%	16,4%	13,4%	14,1%
2008	12,8%	17,0%	12,4%	15,2%	8,3%	10,6%	18,1%	13,8%	8,9%	15,4%	11,9%	12,8%
2009	12,4%	16,3%	8,6%	13,4%	9,3%	11,7%	19,6%	13,3%	8,8%	14,2%	10,9%	12,3%
2010	15,6%	17,0%	12,0%	15,2%	11,0%	11,6%	19,1%	15,0%	10,8%	16,3%	10,5%	13,9%
2011	14,6%	17,8%	13,4%	16,2%	11,0%	11,3%	19,1%	16,4%	12,0%	16,8%	11,4%	14,5%
2012	12,4%	15,8%	10,9%	14,0%	10,9%	11,2%	19,0%	15,8%	10,6%	15,7%	11,1%	13,5%
2013	12,4%	16,4%	10,6%	14,6%	11,4%	11,7%	20,3%	16,4%	11,2%	16,3%	11,6%	14,1%
2014	12,6%	17,6%	12,1%	16,3%	12,6%	12,2%	22,2%	17,4%	12,2%	17,3%	11,8%	15,3%
2015	13,0%	16,6%	13,3%	18,8%	13,9%	12,0%	23,0%	19,0%	12,8%	18,7%	12,8%	16,7%
2016	15,8%	17,0%	13,3%	17,3%	13,7%	12,4%	23,7%	19,4%	12,6%	19,8%	12,9%	16,8%
2017	15,0%	17,8%	14,0%	18,1%	13,9%	12,1%	23,4%	20,7%	11,9%	20,3%	14,5%	17,3%
Total	*14,8%*	*17,4%*	*12,6%*	*15,0%*	*11,2%*	*11,5%*	*19,1%*	*14,9%*	*42,4%*	*16,0%*	*12,3%*	*13,9%*

The table reports the return on capital ratios for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. All return ratios use average balance sheets when possible, and ending balance sheet when there is no beginning balance sheet. Return on capital employed divide EBIE by the total assets less non interest bearing debt. That is, cash and its financial income is consistently included in the ratio. Return on average assets also use EBIE, but total assets in the denominator. Due to negative operating working capital and low fixed assets, invested capital was often too negative to get a reasonable mean values on return on invested capital. The results are presented by year and industry. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.4: Average Industry Growth Rate and Cash Conversion Ratios by Year, 1998-2017

Growth in Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998												
1999	13,2%	10,0%	11,4%	15,6%	9,1%	8,5%	13,3%	20,0%	10,5%	13,8%	8,5%	*11,9%*
2000	14,2%	9,5%	11,4%	14,2%	7,6%	6,8%	15,7%	16,4%	10,1%	12,8%	9,0%	*11,0%*
2001	13,8%	6,1%	6,6%	10,8%	5,5%	6,2%	10,1%	7,1%	5,1%	7,2%	7,9%	*7,3%*
2002	10,1%	7,7%	4,8%	7,4%	5,0%	4,8%	9,2%	2,6%	3,8%	4,6%	5,5%	*5,5%*
2003	9,5%	8,3%	5,5%	7,9%	4,7%	4,6%	13,5%	7,2%	4,6%	6,2%	6,4%	*6,4%*
2004	12,2%	12,6%	8,8%	10,6%	5,2%	4,0%	12,6%	11,3%	6,6%	9,6%	5,6%	*8,2%*
2005	10,3%	11,8%	10,4%	13,0%	6,6%	4,6%	9,9%	15,3%	9,7%	11,4%	5,5%	*9,6%*
2006	12,0%	9,2%	12,2%	15,5%	8,0%	7,1%	10,8%	17,6%	11,7%	12,9%	7,1%	*11,3%*
2007	10,2%	10,1%	11,8%	13,3%	6,9%	10,5%	10,1%	15,4%	13,4%	12,5%	9,6%	*10,7%*
2008	5,9%	4,5%	4,2%	7,3%	2,0%	4,1%	8,6%	10,7%	6,5%	6,7%	6,2%	*5,3%*
2009	7,7%	3,9%	-0,2%	6,9%	2,7%	4,2%	11,5%	8,3%	4,8%	4,5%	4,4%	*4,5%*
2010	13,0%	7,2%	8,1%	11,7%	6,1%	4,6%	10,1%	12,1%	7,7%	9,5%	6,5%	*8,6%*
2011	8,4%	9,0%	7,0%	10,7%	4,2%	3,6%	8,4%	12,6%	7,7%	8,6%	5,5%	*7,4%*
2012	6,3%	2,4%	2,9%	6,6%	3,7%	3,2%	9,2%	10,4%	6,5%	6,8%	4,7%	*5,6%*
2013	3,2%	3,0%	2,6%	8,2%	3,9%	3,7%	10,6%	10,7%	6,5%	7,1%	4,8%	*6,1%*
2014	3,1%	5,3%	5,6%	11,1%	5,6%	4,7%	11,9%	11,2%	11,2%	8,6%	4,8%	*8,0%*
2015	4,4%	5,0%	7,7%	16,5%	7,9%	5,7%	12,6%	14,0%	12,2%	11,3%	7,1%	*10,8%*
2016	9,2%	6,0%	7,1%	11,8%	7,1%	5,8%	12,2%	16,7%	12,9%	12,4%	7,3%	*10,0%*
2017	10,7%	6,3%	6,7%	13,0%	5,9%	3,7%	9,5%	15,5%	14,2%	11,2%	7,6%	*9,4%*
Total	*9,4%*	*7,2%*	*7,3%*	*11,1%*	*5,7%*	*5,3%*	*11,0%*	*12,3%*	*8,8%*	*9,3%*	*6,5%*	*8,3%*

Growth in Sales, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998												
1999	10,2%	11,5%	11,1%	16,7%	11,9%	9,9%	12,8%	23,8%	13,2%	14,9%	6,5%	*13,3%*
2000	16,8%	12,8%	13,1%	17,0%	11,1%	12,1%	13,5%	23,0%	15,7%	15,4%	8,7%	*13,8%*
2001	20,8%	10,9%	8,3%	13,3%	8,6%	11,8%	16,0%	20,7%	9,5%	11,3%	10,5%	*11,1%*
2002	10,1%	8,4%	5,6%	9,1%	8,9%	10,0%	13,7%	7,6%	10,8%	7,5%	5,9%	*8,4%*
2003	8,3%	10,3%	6,3%	8,5%	7,5%	6,9%	13,8%	8,1%	7,1%	7,6%	4,2%	*7,8%*
2004	12,0%	13,4%	9,9%	12,0%	7,6%	6,8%	11,2%	15,4%	10,2%	10,2%	6,0%	*9,7%*
2005	13,4%	16,1%	11,9%	15,7%	9,8%	10,8%	9,1%	20,0%	13,5%	13,6%	5,4%	*12,2%*
2006	14,9%	6,9%	13,8%	18,5%	11,6%	12,0%	13,2%	21,5%	13,9%	15,1%	8,1%	*14,0%*
2007	15,1%	15,3%	14,2%	18,4%	10,5%	11,0%	13,1%	19,7%	16,1%	15,5%	13,4%	*14,1%*
2008	9,2%	8,5%	4,5%	12,8%	6,4%	10,2%	15,3%	17,8%	15,8%	11,7%	9,1%	*10,0%*
2009	2,7%	4,1%	-6,3%	6,7%	4,9%	7,9%	14,1%	10,5%	9,9%	3,9%	1,4%	*4,7%*
2010	20,3%	11,9%	11,8%	15,3%	9,5%	5,9%	10,7%	14,9%	12,5%	12,3%	8,5%	*11,6%*
2011	13,0%	14,9%	12,0%	18,2%	9,7%	7,7%	13,5%	20,3%	19,3%	14,2%	7,7%	*13,2%*
2012	6,1%	5,3%	2,1%	10,1%	6,9%	8,4%	12,3%	13,2%	12,6%	9,1%	5,1%	*8,2%*
2013	4,8%	2,5%	1,7%	10,1%	7,4%	8,1%	11,8%	10,5%	8,3%	7,4%	7,6%	*7,8%*
2014	2,3%	7,0%	7,3%	14,1%	9,8%	8,8%	13,2%	13,3%	14,6%	10,6%	6,0%	*10,7%*
2015	8,4%	4,7%	9,3%	17,3%	11,0%	9,8%	12,6%	15,9%	14,1%	11,7%	6,3%	*12,2%*
2016	8,2%	5,2%	7,7%	14,4%	10,7%	8,7%	15,5%	18,3%	17,5%	12,9%	8,2%	*12,1%*
2017	12,6%	7,2%	9,0%	13,6%	7,6%	6,1%	8,2%	17,6%	14,1%	11,2%	8,8%	*10,3%*
Total	*11,1%*	*9,3%*	*8,3%*	*13,8%*	*9,1%*	*9,2%*	*12,8%*	*16,2%*	*13,2%*	*11,3%*	*7,2%*	*10,8%*

Operating Cash Flow To Net Income, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998												
1999	-93%	157%	47%	93%	102%	99%	134%	8%	-26%	68%	142%	*84%*
2000	170%	17%	19%	79%	59%	95%	101%	62%	162%	79%	71%	*68%*
2001	88%	-131%	-24%	45%	58%	67%	91%	81%	172%	65%	58%	*49%*
2002	154%	8%	52%	59%	76%	60%	88%	91%	81%	76%	72%	*70%*
2003	29%	19%	29%	41%	66%	69%	75%	61%	50%	88%	-42%	*58%*
2004	61%	-47%	35%	59%	75%	24%	116%	29%	14%	71%	35%	*60%*
2005	91%	-36%	58%	84%	63%	58%	86%	106%	93%	84%	37%	*71%*
2006	-60%	-9%	51%	71%	64%	98%	88%	82%	21%	70%	67%	*67%*
2007	63%	40%	42%	68%	67%	109%	95%	67%	46%	87%	8%	*70%*
2008	122%	79%	67%	80%	49%	39%	117%	91%	103%	95%	-13%	*72%*
2009	94%	-27%	28%	85%	80%	60%	113%	85%	94%	68%	34%	*71%*
2010	22%	36%	37%	48%	65%	62%	64%	75%	68%	62%	98%	*60%*
2011	-72%	86%	51%	54%	45%	71%	101%	76%	64%	70%	54%	*60%*
2012	-81%	15%	46%	61%	59%	90%	78%	62%	6%	91%	-10%	*64%*
2013	80%	78%	66%	79%	88%	56%	95%	88%	48%	81%	78%	*80%*
2014	144%	74%	65%	93%	75%	84%	98%	88%	72%	82%	42%	*81%*
2015	79%	55%	86%	98%	88%	98%	111%	72%	108%	94%	75%	*92%*
2016	23%	124%	61%	55%	67%	69%	84%	115%	119%	70%	56%	*69%*
2017	155%	82%	80%	82%	85%	98%	91%	111%	79%	96%	121%	*90%*
Total	*57%*	*36%*	*48%*	*72%*	*71%*	*75%*	*96%*	*78%*	*73%*	*80%*	*53%*	*72%*

Operating Cash Flow to EBITDA, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998												
1999	6,3%	16,8%	21,4%	31,5%	31,1%	37,1%	40,8%	36,7%	26,0%	28,6%	20,6%	*29,7%*
2000	11,5%	8,1%	23,4%	28,8%	29,0%	30,5%	34,6%	33,7%	42,8%	27,5%	20,8%	*27,9%*
2001	29,2%	7,2%	9,1%	18,7%	20,8%	39,5%	27,3%	31,1%	40,3%	19,6%	32,7%	*21,1%*
2002	28,9%	10,7%	13,9%	20,8%	20,4%	33,2%	26,6%	19,0%	22,8%	21,7%	8,6%	*20,5%*
2003	28,2%	14,3%	24,7%	25,2%	24,9%	29,5%	40,8%	41,4%	9,6%	26,8%	19,1%	*26,7%*
2004	27,1%	20,0%	27,2%	29,7%	32,1%	26,1%	44,2%	42,0%	19,3%	34,0%	21,0%	*31,6%*
2005	21,7%	23,5%	34,4%	37,4%	38,7%	43,0%	53,3%	36,6%	52,2%	38,5%	25,9%	*38,2%*
2006	21,2%	15,0%	28,2%	37,6%	35,1%	42,3%	47,9%	42,4%	38,5%	38,0%	37,4%	*36,4%*
2007	28,2%	12,9%	28,1%	37,1%	33,1%	36,9%	42,1%	47,3%	39,9%	41,4%	33,5%	*36,2%*
2008	24,0%	23,8%	33,3%	38,4%	33,8%	29,7%	44,0%	50,0%	39,8%	39,3%	28,5%	*36,5%*
2009	47,3%	17,7%	30,2%	38,2%	35,2%	33,2%	57,6%	53,2%	67,9%	41,3%	26,1%	*38,5%*
2010	35,9%	13,5%	30,2%	38,8%	38,4%	38,7%	54,6%	49,9%	58,9%	43,9%	33,9%	*40,1%*
2011	26,7%	20,7%	32,9%	37,8%	38,2%	42,1%	47,6%	46,5%	49,8%	42,7%	26,4%	*39,2%*
2012	15,3%	35,1%	37,2%	32,9%	34,3%	45,4%	44,0%	49,6%	30,2%	46,3%	26,1%	*38,6%*
2013	19,5%	31,1%	43,7%	45,7%	43,9%	36,6%	54,4%	63,4%	54,2%	47,0%	26,3%	*45,2%*
2014	26,0%	25,6%	40,0%	43,7%	41,6%	45,1%	63,5%	57,7%	33,4%	48,8%	31,2%	*45,0%*
2015	43,8%	23,9%	46,4%	58,8%	50,9%	51,8%	62,4%	57,3%	59,3%	57,0%	34,4%	*53,7%*
2016	46,5%	39,9%	45,7%	49,7%	49,5%	50,8%	57,9%	77,2%	61,7%	55,0%	39,8%	*52,0%*
2017	65,2%	24,4%	43,0%	50,2%	48,0%	39,8%	62,3%	75,5%	88,0%	54,2%	45,4%	*51,5%*
Total	*27,9%*	*20,4%*	*29,7%*	*37,9%*	*35,6%*	*38,2%*	*48,0%*	*48,4%*	*44,2%*	*39,6%*	*28,1%*	*37,2%*

The table reports average growth rate and cash conversion ratios, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. Operating cash flow is estimated using earnings and changes in operating working capital, which is defined as inventories, trade receivables and other operating current receivables (taxes, prepaid expenses, etc.), less trade payables and other operating current payables (accrued liabilities, salary, advance payment etc.). The first year is removed from the table since there are no beginning balance sheets, which implies a 100% conversion for all observations. Note that the survivorship bias in the data becomes apparent in the growth rates, as the average sales growth of 10.8% exceeds the growth rate of the economy. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.5: Average Industry Solidity and Liquidity Ratios by Year, 1998-2017

Equity to Total Assets (Solidity), %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	25,9%	21,4%	25,8%	25,6%	23,0%	22,0%	33,6%	32,4%	27,1%	28,0%	26,9%	*25,7%*
1999	26,8%	22,0%	26,9%	26,4%	23,9%	23,1%	34,6%	33,7%	27,9%	28,9%	27,7%	*26,7%*
2000	27,7%	21,9%	27,7%	26,9%	24,4%	23,5%	35,3%	33,9%	28,6%	29,3%	28,1%	*27,3%*
2001	29,1%	21,8%	28,5%	27,4%	24,4%	22,9%	35,1%	34,3%	29,4%	29,5%	28,4%	*27,5%*
2002	30,1%	22,3%	29,1%	27,6%	24,7%	23,5%	34,8%	34,5%	29,9%	29,7%	28,6%	*27,8%*
2003	28,8%	22,2%	29,6%	27,7%	24,6%	23,5%	35,4%	34,3%	30,9%	30,4%	28,3%	*28,1%*
2004	30,9%	22,8%	30,5%	28,1%	25,0%	23,9%	36,5%	36,2%	30,4%	31,1%	29,3%	*28,7%*
2005	30,9%	23,8%	31,8%	29,4%	25,7%	24,8%	38,1%	37,7%	31,5%	32,6%	30,4%	*29,9%*
2006	32,1%	24,5%	32,8%	30,2%	26,5%	25,5%	39,8%	38,5%	33,1%	33,7%	30,7%	*30,8%*
2007	30,9%	24,1%	33,2%	30,5%	25,8%	24,1%	39,0%	39,2%	33,2%	33,5%	30,8%	*30,7%*
2008	30,2%	24,3%	34,0%	30,8%	25,0%	23,4%	39,1%	38,7%	33,5%	33,3%	30,5%	*30,5%*
2009	30,7%	24,8%	34,3%	31,0%	25,3%	23,7%	39,2%	39,1%	34,1%	34,0%	31,2%	*30,9%*
2010	29,9%	25,1%	34,5%	31,2%	25,7%	24,3%	39,5%	39,4%	34,9%	34,5%	29,9%	*31,2%*
2011	30,7%	23,7%	35,2%	30,9%	26,2%	24,3%	39,6%	40,3%	33,4%	34,7%	31,0%	*31,5%*
2012	29,4%	24,3%	35,2%	30,5%	26,5%	24,2%	38,9%	39,6%	34,3%	34,5%	30,8%	*31,3%*
2013	29,9%	25,2%	36,1%	31,5%	27,3%	25,2%	40,2%	40,9%	35,4%	35,7%	31,4%	*32,3%*
2014	31,1%	26,5%	37,6%	32,9%	29,0%	25,9%	42,5%	42,6%	37,2%	37,4%	32,7%	*34,0%*
2015	31,1%	28,0%	38,1%	34,0%	30,5%	27,2%	43,5%	43,6%	38,4%	38,5%	33,1%	*35,1%*
2016	32,5%	28,6%	38,4%	33,9%	31,0%	27,7%	44,1%	44,9%	38,3%	39,0%	33,6%	*35,5%*
2017	33,0%	29,7%	38,7%	34,1%	32,3%	28,0%	44,3%	44,5%	37,2%	39,4%	34,2%	*36,1%*
Total	*30,0%*	*24,3%*	*32,2%*	*30,4%*	*26,3%*	*24,4%*	*39,0%*	*38,7%*	*33,2%*	*33,5%*	*30,4%*	*30,6%*

Retained Earnings to Total Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	12,9%	10,5%	11,4%	8,7%	7,2%	7,7%	10,4%	10,1%	8,0%	9,4%	13,6%	*9,0%*
1999	13,5%	11,1%	12,5%	9,4%	7,6%	7,5%	11,4%	11,4%	7,7%	10,4%	14,2%	*9,7%*
2000	14,6%	11,1%	13,4%	10,6%	8,6%	8,5%	12,6%	12,5%	9,3%	11,4%	14,8%	*10,8%*
2001	15,1%	11,9%	15,5%	12,1%	9,7%	9,3%	14,6%	13,7%	11,4%	12,8%	15,9%	*12,2%*
2002	17,0%	12,0%	16,5%	13,3%	9,9%	9,7%	15,3%	14,5%	10,4%	13,4%	16,5%	*12,8%*
2003	17,5%	12,0%	16,8%	13,5%	10,0%	9,2%	14,7%	13,1%	12,2%	13,2%	16,1%	*12,7%*
2004	18,6%	12,0%	16,7%	13,4%	9,9%	9,1%	15,3%	13,0%	11,3%	13,2%	17,1%	*12,7%*
2005	18,2%	12,9%	17,0%	13,7%	10,0%	9,5%	16,8%	14,2%	11,2%	13,7%	17,7%	*13,1%*
2006	18,6%	14,1%	17,8%	13,9%	10,5%	9,7%	17,9%	14,8%	11,8%	14,4%	17,5%	*13,7%*
2007	19,9%	14,0%	18,7%	14,3%	11,3%	10,0%	17,7%	16,1%	13,0%	14,7%	18,3%	*14,2%*
2008	20,0%	15,1%	21,1%	15,4%	11,3%	10,3%	17,6%	16,0%	12,9%	15,4%	18,8%	*14,9%*
2009	19,7%	15,0%	22,5%	16,1%	10,2%	9,5%	16,4%	16,9%	13,2%	15,6%	19,3%	*14,8%*
2010	17,3%	15,4%	20,9%	15,1%	9,9%	9,9%	16,7%	16,9%	14,0%	15,1%	15,9%	*14,2%*
2011	18,8%	14,5%	21,4%	15,2%	10,7%	10,4%	18,4%	18,4%	13,4%	16,0%	19,4%	*15,0%*
2012	20,0%	15,9%	23,2%	16,4%	11,2%	10,8%	18,4%	19,6%	15,9%	17,0%	19,3%	*15,8%*
2013	19,4%	16,1%	23,1%	16,1%	11,5%	10,5%	18,0%	19,3%	15,1%	16,8%	18,9%	*15,7%*
2014	20,5%	16,8%	23,8%	16,4%	12,5%	10,8%	18,5%	20,0%	16,7%	17,3%	19,8%	*16,3%*
2015	20,0%	18,3%	24,1%	16,4%	13,6%	11,9%	19,0%	21,0%	19,2%	18,0%	20,1%	*17,0%*
2016	20,3%	18,9%	24,3%	17,5%	14,6%	12,5%	19,7%	20,8%	19,6%	18,2%	20,8%	*17,6%*
2017	20,8%	19,6%	24,4%	18,6%	16,7%	13,7%	21,2%	21,2%	20,7%	19,5%	21,0%	*18,9%*
Total	*18,1%*	*14,4%*	*18,6%*	*14,6%*	*10,8%*	*9,9%*	*16,8%*	*16,5%*	*13,6%*	*14,9%*	*17,8%*	*14,1%*

Working Capital to Total Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	21,1%	13,5%	28,6%	23,9%	27,3%	21,8%	23,4%	29,7%	20,8%	20,3%	24,3%	*24,3%*
1999	20,3%	13,8%	28,7%	24,1%	27,0%	21,6%	24,1%	30,7%	20,7%	21,2%	23,3%	*24,5%*
2000	20,8%	12,8%	28,9%	23,8%	26,7%	21,1%	24,2%	29,8%	20,8%	21,1%	22,3%	*24,3%*
2001	22,1%	12,1%	29,8%	24,2%	26,5%	20,3%	24,3%	30,6%	22,3%	21,8%	22,0%	*24,6%*
2002	22,6%	13,1%	30,2%	24,7%	26,5%	21,4%	24,7%	31,5%	22,4%	22,4%	21,7%	*24,9%*
2003	20,1%	13,0%	30,4%	24,7%	25,9%	21,6%	25,9%	31,3%	23,6%	23,4%	21,9%	*25,1%*
2004	20,8%	12,2%	30,0%	24,5%	24,6%	20,7%	25,9%	32,3%	23,1%	23,4%	21,8%	*24,7%*
2005	20,8%	12,4%	29,8%	24,8%	24,0%	20,3%	26,0%	32,7%	24,7%	24,1%	22,6%	*24,7%*
2006	23,1%	12,4%	30,5%	25,6%	24,4%	20,6%	27,8%	33,9%	25,3%	25,3%	22,4%	*25,5%*
2007	21,1%	12,2%	31,3%	26,2%	23,8%	18,9%	27,9%	35,2%	25,7%	26,2%	22,9%	*25,7%*
2008	20,6%	12,2%	32,4%	27,1%	23,4%	18,7%	29,7%	35,2%	26,2%	26,8%	23,2%	*26,1%*
2009	21,1%	13,3%	32,8%	27,4%	23,7%	19,2%	30,6%	36,0%	27,9%	27,9%	24,3%	*26,7%*
2010	21,3%	14,1%	33,6%	28,0%	24,5%	19,9%	31,5%	37,3%	29,3%	28,9%	24,6%	*27,5%*
2011	23,0%	13,7%	34,7%	28,7%	25,2%	20,2%	32,5%	39,2%	28,9%	30,1%	24,6%	*28,4%*
2012	22,0%	13,9%	35,2%	29,0%	26,2%	20,8%	33,7%	40,1%	30,5%	31,2%	25,4%	*29,2%*
2013	22,3%	15,1%	36,0%	29,9%	26,6%	21,9%	35,3%	41,5%	31,3%	32,4%	26,5%	*30,1%*
2014	23,2%	16,1%	36,9%	31,0%	28,3%	22,5%	37,4%	42,9%	33,4%	34,0%	27,6%	*31,6%*
2015	23,5%	17,2%	37,2%	32,1%	29,2%	23,1%	38,2%	43,3%	35,1%	34,9%	27,9%	*32,4%*
2016	25,4%	17,3%	37,2%	31,8%	29,1%	23,2%	37,9%	44,2%	34,9%	35,4%	28,2%	*32,5%*
2017	26,1%	19,4%	37,5%	32,1%	30,6%	24,0%	38,7%	43,1%	34,3%	36,0%	29,6%	*33,3%*
Total	*22,0%*	*13,9%*	*32,0%*	*27,6%*	*26,1%*	*21,0%*	*30,5%*	*36,4%*	*27,5%*	*27,6%*	*24,4%*	*27,4%*

Cash to Current Assets, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	31,4%	33,1%	23,0%	27,4%	26,0%	29,0%	49,2%	42,3%	36,8%	38,0%	28,5%	*30,8%*
1999	30,1%	32,6%	23,4%	28,4%	27,5%	30,5%	50,2%	43,1%	37,7%	38,6%	29,2%	*31,8%*
2000	30,5%	32,6%	23,8%	28,9%	28,0%	31,6%	50,8%	42,3%	38,2%	38,5%	29,6%	*32,3%*
2001	34,2%	32,7%	23,9%	29,1%	28,2%	32,2%	51,1%	42,9%	38,1%	38,6%	30,4%	*32,6%*
2002	36,0%	33,9%	24,2%	29,5%	28,9%	33,3%	51,8%	42,9%	38,1%	39,0%	30,4%	*33,1%*
2003	33,8%	34,5%	24,1%	29,5%	29,1%	33,4%	52,7%	41,8%	39,2%	39,6%	29,8%	*33,3%*
2004	34,9%	35,2%	24,6%	29,7%	29,5%	32,7%	53,3%	42,5%	37,8%	39,9%	29,3%	*33,6%*
2005	33,4%	35,6%	25,4%	30,9%	30,4%	33,2%	53,6%	43,6%	39,9%	40,7%	30,6%	*34,5%*
2006	34,3%	34,6%	26,1%	31,7%	31,3%	33,5%	54,0%	43,2%	39,8%	41,6%	31,8%	*35,2%*
2007	34,5%	34,1%	27,2%	32,2%	31,3%	33,6%	54,2%	44,8%	40,3%	41,7%	33,0%	*35,8%*
2008	34,1%	35,1%	28,7%	33,6%	31,1%	33,4%	55,9%	46,6%	40,5%	43,0%	32,2%	*36,7%*
2009	32,2%	35,7%	28,4%	34,1%	32,5%	33,2%	58,3%	47,1%	41,8%	43,3%	32,7%	*37,4%*
2010	32,1%	36,6%	28,2%	34,0%	33,4%	33,4%	58,9%	47,0%	42,5%	43,5%	34,8%	*37,9%*
2011	33,1%	36,4%	29,4%	34,7%	34,1%	33,9%	58,6%	48,2%	42,4%	44,0%	33,4%	*38,5%*
2012	31,7%	37,9%	29,9%	34,7%	35,0%	34,8%	58,6%	49,3%	43,0%	44,3%	33,4%	*39,0%*
2013	32,1%	38,4%	30,0%	35,2%	36,5%	35,3%	60,0%	50,3%	44,7%	45,4%	34,5%	*40,0%*
2014	33,4%	40,3%	30,7%	36,9%	37,9%	35,9%	61,5%	51,9%	45,8%	46,6%	36,0%	*41,4%*
2015	33,9%	41,2%	31,9%	39,4%	39,6%	36,4%	62,1%	52,6%	46,6%	47,5%	36,8%	*42,8%*
2016	36,2%	41,4%	31,8%	37,6%	40,0%	36,0%	60,3%	52,7%	47,8%	47,1%	37,2%	*42,4%*
2017	37,0%	42,6%	31,8%	37,8%	39,7%	34,9%	59,4%	51,3%	47,2%	46,4%	38,1%	*42,2%*
Total	*33,4%*	*36,2%*	*26,8%*	*33,3%*	*32,5%*	*33,4%*	*56,2%*	*46,6%*	*41,7%*	*42,5%*	*32,6%*	*36,7%*

The table reports average solidity and liquidity ratios, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. Note that working capital is all current assets less all current liabilities, irrespective of whether operating or financial items are captured. Retained earnings do not include the equity portion of untaxed reserves. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.6: Average Industry Leverage & Liquidity Ratios by Industry and Year, 1998-2017

Debt to EBITDA, x Turns

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	2,9x	3,1x	2,8x	2,2x	2,8x	2,4x	1,7x	1,4x	1,3x	1,9x	3,0x	2,4x
1999	3,0x	3,1x	2,8x	2,2x	2,7x	2,5x	1,7x	1,3x	1,3x	1,9x	3,3x	2,4x
2000	2,7x	3,0x	2,7x	2,1x	2,7x	2,5x	1,7x	1,5x	1,5x	1,9x	3,0x	2,3x
2001	2,2x	2,9x	2,7x	2,1x	2,5x	2,2x	1,7x	1,4x	1,2x	1,8x	3,0x	2,2x
2002	3,0x	2,7x	2,7x	2,0x	2,5x	2,2x	1,7x	1,3x	1,0x	1,9x	3,3x	2,2x
2003	3,4x	3,0x	2,7x	2,2x	2,5x	2,3x	1,8x	1,2x	1,6x	1,9x	3,2x	2,3x
2004	3,1x	2,8x	2,5x	2,2x	2,5x	2,3x	1,5x	1,2x	1,9x	1,9x	3,5x	2,2x
2005	2,5x	3,0x	2,4x	1,9x	2,1x	2,0x	1,6x	1,2x	1,2x	1,8x	3,3x	2,0x
2006	2,7x	3,0x	2,2x	1,8x	2,1x	1,8x	1,3x	1,0x	1,4x	1,6x	3,2x	1,9x
2007	2,3x	3,1x	2,2x	1,7x	2,1x	2,0x	1,3x	1,0x	1,2x	1,6x	2,9x	1,9x
2008	2,9x	3,3x	2,2x	1,8x	2,0x	1,8x	1,4x	0,8x	1,1x	1,7x	3,0x	1,9x
2009	2,8x	3,0x	2,1x	1,8x	2,1x	1,7x	1,3x	0,8x	1,0x	1,5x	3,5x	1,8x
2010	2,9x	3,0x	2,4x	1,9x	2,2x	1,8x	1,3x	1,0x	0,9x	1,7x	4,2x	2,0x
2011	3,4x	3,1x	2,2x	1,7x	2,1x	1,7x	1,3x	0,8x	0,8x	1,5x	3,4x	1,8x
2012	3,1x	3,0x	2,3x	1,8x	2,1x	1,9x	1,3x	1,0x	1,2x	1,5x	3,6x	1,9x
2013	3,3x	2,7x	2,2x	1,7x	1,9x	1,8x	1,2x	1,0x	1,1x	1,5x	3,3x	1,8x
2014	3,4x	2,9x	2,2x	1,6x	1,9x	1,6x	1,1x	0,9x	1,2x	1,4x	3,2x	1,7x
2015	2,4x	2,6x	2,2x	1,5x	1,8x	1,6x	1,2x	0,8x	0,9x	1,4x	3,7x	1,7x
2016	3,1x	2,5x	1,9x	1,4x	1,7x	1,7x	0,9x	0,7x	0,5x	1,3x	3,2x	1,5x
2017	3,0x	2,5x	1,9x	1,3x	1,6x	1,4x	1,0x	0,7x	0,6x	1,2x	2,8x	1,4x
Total	2,9x	2,9x	2,4x	1,8x	2,2x	2,0x	1,4x	1,0x	1,1x	1,6x	3,3x	2,0x

Cash as a % of Sales

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	16,2%	18,0%	12,2%	12,2%	9,7%	7,3%	23,3%	29,5%	22,3%	22,3%	18,6%	15,3%
1999	17,0%	19,4%	13,3%	13,4%	11,1%	8,4%	26,9%	35,2%	27,0%	25,4%	20,8%	17,5%
2000	17,3%	19,0%	13,5%	13,4%	11,1%	9,0%	28,4%	36,4%	28,3%	25,9%	21,4%	17,9%
2001	19,0%	17,7%	13,9%	13,1%	11,0%	9,0%	27,1%	32,3%	25,3%	25,2%	20,7%	17,3%
2002	21,0%	19,0%	14,6%	13,9%	11,7%	9,5%	27,5%	33,7%	26,0%	26,1%	21,6%	18,2%
2003	19,7%	19,8%	15,4%	14,7%	12,2%	9,9%	29,4%	35,0%	29,6%	27,9%	23,3%	19,2%
2004	23,0%	19,8%	16,1%	15,4%	12,9%	10,3%	32,1%	36,9%	28,8%	29,4%	23,6%	20,1%
2005	24,4%	20,9%	16,6%	16,6%	13,4%	10,9%	33,4%	38,7%	33,6%	30,9%	25,8%	21,2%
2006	24,5%	20,6%	17,1%	16,8%	13,9%	10,5%	34,9%	39,0%	35,7%	31,8%	27,4%	21,8%
2007	21,8%	22,4%	18,1%	17,0%	14,2%	10,8%	37,7%	42,9%	37,7%	32,9%	28,2%	22,9%
2008	22,3%	23,2%	19,5%	17,9%	14,7%	10,5%	39,4%	43,4%	37,5%	34,4%	29,8%	24,0%
2009	24,1%	23,9%	21,0%	18,7%	15,4%	10,4%	39,3%	44,1%	38,0%	36,8%	31,8%	25,2%
2010	23,7%	25,9%	21,9%	20,6%	16,6%	10,9%	42,3%	47,8%	39,9%	38,6%	50,6%	27,7%
2011	25,4%	22,4%	22,2%	19,5%	16,4%	11,5%	41,7%	46,5%	35,4%	38,2%	32,4%	26,4%
2012	22,8%	23,0%	22,7%	18,9%	16,3%	11,5%	41,0%	47,4%	37,8%	38,8%	33,0%	26,5%
2013	25,5%	23,9%	24,2%	19,6%	17,4%	11,7%	43,3%	51,1%	41,4%	40,7%	32,9%	27,8%
2014	27,6%	26,1%	24,3%	20,4%	18,0%	12,4%	45,7%	54,5%	44,4%	42,9%	36,1%	29,3%
2015	25,3%	27,6%	24,5%	21,2%	19,1%	12,4%	47,2%	55,5%	44,9%	43,4%	36,2%	30,1%
2016	24,5%	26,4%	24,5%	19,7%	18,6%	13,1%	44,3%	58,8%	49,3%	41,8%	36,2%	29,3%
2017	26,0%	26,0%	23,6%	17,7%	16,3%	11,9%	40,1%	51,2%	41,7%	38,5%	34,2%	26,7%
Total	22,5%	22,3%	18,2%	17,5%	14,5%	10,5%	37,1%	43,7%	35,9%	34,0%	29,6%	23,4%

Net-Debt (Cash) to EBITDA, x Turns

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	1,8x	1,8x	1,4x	0,3x	1,2x	1,1x	(0,2x)	(1,4x)	0,1x	(0,1x)	1,6x	0,7x
1999	1,8x	1,7x	1,3x	0,2x	1,0x	0,6x	(0,4x)	(1,5x)	(0,3x)	(0,1x)	1,5x	0,5x
2000	1,5x	1,6x	1,3x	0,2x	1,0x	0,6x	(0,6x)	(1,2x)	(0,4x)	(0,1x)	1,6x	0,5x
2001	1,1x	1,8x	1,1x	0,1x	0,8x	0,2x	(0,7x)	(1,1x)	(0,5x)	(0,3x)	1,4x	0,3x
2002	1,2x	1,7x	1,3x	0,3x	0,9x	0,3x	(0,6x)	(0,9x)	(0,3x)	(0,1x)	1,8x	0,5x
2003	1,4x	1,7x	1,1x	0,2x	0,8x	0,4x	(0,4x)	(1,0x)	(0,8x)	(0,3x)	1,8x	0,4x
2004	1,3x	1,4x	0,9x	0,2x	0,7x	0,6x	(0,6x)	(1,2x)	(0,2x)	(0,4x)	2,0x	0,3x
2005	1,0x	1,8x	0,8x	(0,1x)	0,4x	0,4x	(0,6x)	(1,5x)	(1,4x)	(0,5x)	1,7x	0,1x
2006	1,2x	1,9x	0,8x	(0,0x)	0,4x	0,3x	(0,9x)	(1,1x)	(0,4x)	(0,6x)	1,6x	0,1x
2007	0,8x	2,0x	0,8x	(0,0x)	0,4x	0,4x	(0,9x)	(1,2x)	(0,6x)	(0,4x)	1,4x	0,1x
2008	0,7x	2,0x	0,6x	0,0x	0,4x	(0,0x)	(0,9x)	(1,3x)	(0,9x)	(0,4x)	1,6x	0,1x
2009	1,6x	1,4x	0,6x	(0,1x)	0,4x	0,2x	(1,0x)	(1,4x)	(0,9x)	(0,6x)	2,1x	(0,0x)
2010	1,2x	1,6x	0,7x	(0,2x)	0,3x	0,1x	(1,2x)	(1,3x)	(1,3x)	(0,6x)	1,2x	(0,1x)
2011	1,6x	1,6x	0,6x	(0,2x)	0,4x	0,1x	(1,0x)	(1,4x)	(0,9x)	(0,6x)	2,0x	(0,0x)
2012	1,1x	2,0x	0,6x	(0,1x)	0,3x	0,4x	(0,9x)	(1,1x)	(1,1x)	(0,6x)	2,1x	(0,0x)
2013	1,2x	1,4x	0,4x	(0,3x)	0,1x	0,1x	(1,1x)	(1,4x)	(1,3x)	(0,6x)	1,6x	(0,2x)
2014	2,1x	1,2x	0,5x	(0,4x)	0,1x	0,0x	(1,3x)	(1,2x)	(1,3x)	(0,7x)	1,4x	(0,3x)
2015	0,5x	0,6x	0,2x	(0,6x)	(0,1x)	(0,2x)	(1,3x)	(2,0x)	(1,1x)	(0,9x)	1,8x	(0,4x)
2016	1,2x	1,0x	0,2x	(0,6x)	(0,2x)	(0,1x)	(1,1x)	(1,1x)	(1,6x)	(0,7x)	1,4x	(0,4x)
2017	1,0x	1,1x	0,4x	(0,4x)	(0,1x)	0,0x	(0,9x)	(1,1x)	(1,2x)	(0,6x)	1,2x	(0,3x)
Total	1,3x	1,6x	0,8x	(0,1x)	0,5x	0,3x	(0,9x)	(1,3x)	(0,9x)	(0,5x)	1,6x	0,1x

Short Term Debt to Equity, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	Total
1998	13,2%	40,0%	9,5%	7,7%	9,4%	9,7%	4,6%	3,7%	5,2%	12,4%	9,1%	10,0%
1999	14,6%	42,2%	10,1%	8,5%	9,5%	10,5%	4,8%	2,9%	5,6%	13,5%	9,8%	10,6%
2000	17,9%	47,7%	10,7%	9,1%	10,3%	11,0%	5,3%	3,3%	6,0%	14,7%	11,2%	11,5%
2001	14,8%	52,8%	11,6%	10,1%	11,3%	11,5%	5,7%	3,3%	6,8%	15,4%	13,5%	12,4%
2002	17,1%	60,4%	12,1%	11,2%	12,3%	12,3%	5,7%	3,9%	6,7%	17,2%	15,5%	13,6%
2003	29,5%	67,2%	15,5%	12,9%	14,9%	14,8%	6,0%	5,6%	8,5%	19,3%	19,6%	16,1%
2004	35,0%	78,8%	21,4%	16,8%	19,4%	17,9%	7,9%	7,9%	11,4%	23,0%	28,9%	20,4%
2005	38,5%	82,0%	22,6%	18,0%	20,7%	18,8%	9,2%	7,8%	10,6%	24,8%	27,9%	21,8%
2006	34,7%	80,6%	21,8%	17,4%	20,4%	18,7%	8,1%	8,4%	12,1%	24,4%	26,5%	21,2%
2007	35,1%	81,5%	20,7%	19,1%	20,9%	20,5%	8,3%	6,3%	11,7%	24,1%	26,5%	21,5%
2008	40,1%	85,6%	21,2%	19,0%	21,5%	21,2%	8,5%	7,6%	12,3%	24,3%	26,9%	21,9%
2009	42,0%	81,8%	20,9%	17,3%	20,4%	19,4%	7,6%	6,6%	9,5%	22,6%	26,8%	20,5%
2010	41,0%	79,3%	20,1%	17,5%	19,4%	18,1%	8,2%	5,3%	11,2%	21,1%	24,9%	19,7%
2011	45,0%	81,8%	19,1%	17,4%	18,6%	19,0%	8,1%	5,4%	9,9%	20,8%	29,9%	19,5%
2012	39,5%	78,4%	18,7%	16,4%	18,2%	18,1%	7,4%	5,2%	8,8%	19,5%	29,6%	18,6%
2013	42,4%	77,6%	17,9%	16,0%	17,6%	17,1%	6,5%	5,2%	8,2%	19,3%	31,1%	18,2%
2014	38,8%	71,3%	15,8%	13,9%	15,5%	15,3%	5,9%	4,0%	8,5%	17,2%	27,7%	16,1%
2015	38,2%	64,7%	15,1%	12,6%	13,7%	14,2%	5,4%	3,2%	5,1%	15,7%	27,2%	14,6%
2016	30,3%	62,3%	14,2%	12,9%	13,0%	13,4%	5,6%	3,2%	5,6%	15,4%	23,8%	14,1%
2017	33,6%	57,4%	13,2%	13,3%	12,1%	12,9%	5,2%	3,3%	7,4%	15,3%	22,2%	13,6%
Total	32,1%	69,5%	16,5%	14,6%	16,0%	15,7%	6,8%	5,1%	8,6%	19,1%	23,1%	16,9%

The table reports the average of debt- and net debt to EBITDA ratios, as well as cash as % of sales, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. Cash includes all cash and bank balances as well as investments in securities (assets that may be converted into means of payment within a few days). Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.7: Average Industry Interest Cost Ratios by Year, 1998-2017

	Interest Cost to Adjusted Total Liabilities											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	3,4%	4,0%	3,4%	2,9%	3,7%	3,6%	3,2%	2,3%	3,0%	3,0%	3,7%	*3,3%*
1999	3,1%	3,6%	3,1%	2,6%	3,4%	3,3%	2,9%	2,0%	2,7%	2,8%	3,4%	*3,0%*
2000	3,3%	4,0%	3,4%	2,8%	3,7%	3,5%	3,5%	2,8%	3,1%	3,2%	3,7%	*3,3%*
2001	3,6%	4,4%	3,7%	3,1%	3,9%	3,7%	4,2%	3,8%	3,6%	3,8%	4,0%	*3,7%*
2002	4,1%	4,5%	4,0%	3,4%	4,0%	4,0%	4,9%	4,7%	4,1%	4,4%	4,3%	*4,1%*
2003	3,1%	3,5%	3,3%	2,6%	3,4%	3,1%	3,1%	2,8%	2,9%	3,0%	3,5%	*3,1%*
2004	2,8%	2,9%	2,8%	2,3%	3,0%	2,7%	2,6%	2,2%	2,6%	2,5%	2,9%	*2,7%*
2005	2,4%	2,6%	2,5%	2,0%	2,7%	2,4%	2,2%	1,8%	2,4%	2,2%	2,7%	*2,4%*
2006	2,3%	2,8%	2,4%	1,9%	2,6%	2,4%	2,1%	1,8%	2,1%	2,1%	2,7%	*2,3%*
2007	2,9%	3,4%	2,7%	2,2%	2,9%	2,7%	2,6%	2,2%	2,4%	2,6%	3,1%	*2,7%*
2008	3,7%	4,2%	3,6%	2,8%	3,6%	3,5%	4,4%	3,5%	3,7%	3,7%	4,0%	*3,6%*
2009	2,3%	2,7%	2,6%	1,9%	2,7%	2,3%	2,4%	2,2%	2,4%	2,3%	2,7%	*2,4%*
2010	2,2%	2,3%	2,3%	1,7%	2,3%	2,1%	2,1%	1,8%	2,1%	1,9%	2,0%	*2,0%*
2011	2,9%	3,0%	2,8%	2,0%	2,7%	2,6%	2,9%	2,3%	2,2%	2,5%	2,9%	*2,5%*
2012	2,6%	2,8%	2,4%	1,8%	2,5%	2,2%	2,1%	1,7%	2,1%	2,1%	2,6%	*2,2%*
2013	2,4%	2,5%	2,2%	1,7%	2,2%	2,0%	1,7%	1,4%	1,8%	1,8%	2,2%	*1,9%*
2014	2,4%	2,3%	2,0%	1,5%	2,1%	1,8%	1,6%	1,3%	1,7%	1,7%	2,0%	*1,8%*
2015	1,9%	1,9%	1,9%	1,4%	1,9%	1,7%	1,6%	1,4%	1,6%	1,6%	1,9%	*1,7%*
2016	1,8%	1,7%	1,6%	1,2%	1,7%	1,5%	1,4%	1,3%	1,5%	1,4%	1,6%	*1,5%*
2017	1,9%	1,7%	1,6%	1,3%	1,7%	1,5%	1,4%	1,3%	1,4%	1,4%	1,6%	*1,5%*
Total	*2,8%*	*3,0%*	*2,8%*	*2,1%*	*2,8%*	*2,7%*	*2,6%*	*2,2%*	*2,4%*	*2,5%*	*2,9%*	*2,6%*

	Interest Cost to Assets											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	2,6%	3,0%	2,6%	2,2%	2,9%	2,9%	2,2%	1,6%	2,2%	2,2%	2,6%	*2,5%*
1999	2,3%	2,7%	2,3%	2,0%	2,7%	2,6%	1,9%	1,4%	2,0%	2,0%	2,4%	*2,3%*
2000	2,3%	3,0%	2,5%	2,1%	2,9%	2,8%	2,3%	2,0%	2,3%	2,3%	2,6%	*2,5%*
2001	2,4%	3,3%	2,7%	2,3%	3,0%	3,0%	2,8%	2,7%	2,6%	2,7%	2,8%	*2,8%*
2002	2,6%	3,3%	2,8%	2,4%	3,0%	3,0%	3,0%	3,0%	2,8%	3,0%	2,9%	*2,9%*
2003	2,0%	2,6%	2,3%	1,9%	2,5%	2,3%	1,8%	1,8%	2,0%	2,0%	2,3%	*2,2%*
2004	1,8%	2,1%	2,0%	1,7%	2,2%	2,1%	1,6%	1,4%	1,8%	1,7%	2,0%	*1,9%*
2005	1,6%	1,9%	1,7%	1,4%	2,0%	1,9%	1,3%	1,1%	1,6%	1,5%	1,8%	*1,7%*
2006	1,5%	2,0%	1,7%	1,3%	1,9%	1,8%	1,3%	1,1%	1,5%	1,4%	1,8%	*1,6%*
2007	1,9%	2,4%	1,8%	1,5%	2,1%	2,0%	1,5%	1,3%	1,6%	1,6%	2,0%	*1,8%*
2008	2,3%	2,8%	2,2%	1,8%	2,5%	2,4%	2,1%	1,7%	2,0%	2,1%	2,3%	*2,2%*
2009	1,6%	1,8%	1,7%	1,3%	2,0%	1,7%	1,3%	1,2%	1,5%	1,4%	1,7%	*1,6%*
2010	1,5%	1,6%	1,5%	1,1%	1,7%	1,6%	1,1%	1,0%	1,3%	1,2%	1,3%	*1,4%*
2011	1,8%	2,1%	1,7%	1,3%	1,9%	1,8%	1,4%	1,1%	1,3%	1,4%	1,8%	*1,6%*
2012	1,8%	2,0%	1,6%	1,2%	1,8%	1,6%	1,2%	0,9%	1,2%	1,3%	1,7%	*1,5%*
2013	1,6%	1,8%	1,4%	1,2%	1,6%	1,5%	1,0%	0,8%	1,1%	1,1%	1,5%	*1,3%*
2014	1,6%	1,6%	1,3%	1,1%	1,5%	1,4%	0,9%	0,7%	1,0%	1,1%	1,3%	*1,2%*
2015	1,3%	1,3%	1,2%	0,9%	1,3%	1,2%	0,8%	0,7%	1,0%	0,9%	1,1%	*1,0%*
2016	1,2%	1,2%	1,0%	0,8%	1,2%	1,1%	0,7%	0,7%	0,9%	0,8%	1,0%	*0,9%*
2017	1,2%	1,1%	1,0%	0,9%	1,2%	1,1%	0,7%	0,7%	0,8%	0,8%	1,0%	*0,9%*
Total	*1,9%*	*7,2%*	*1,9%*	*1,5%*	*2,1%*	*2,0%*	*1,5%*	*1,3%*	*1,6%*	*1,6%*	*1,9%*	*1,8%*

The table reports the average of interest cost to asset and liabilites ratios, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. Interest cost to adjusted total liabilites is defined as (Financial costs - Financial expenses affecting comparability) / (Non-current liabilities + Current liabilities + Provisions + Deferred tax liability). Interest cost to assets are all financial expenses devided by total assets. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.8: Average Industry Efficiency and Fixed Cost Measures by Year, 1998-2017

Value Added Per Employee, SEKk

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	377	407	343	302	271	268	335	388	321	350	310	*313*
1999	397	418	358	320	285	280	351	401	326	360	316	*327*
2000	421	434	373	337	294	286	367	386	337	368	333	*338*
2001	448	441	379	354	298	296	382	401	336	375	354	*348*
2002	456	464	385	361	309	316	395	393	338	379	360	*356*
2003	489	480	394	368	313	320	419	392	345	389	366	*364*
2004	511	505	409	379	317	319	433	407	355	400	379	*374*
2005	515	529	428	395	324	328	431	427	364	414	395	*386*
2006	546	532	453	416	341	344	445	451	378	437	409	*406*
2007	552	555	487	441	352	364	448	470	396	459	453	*426*
2008	551	556	485	452	349	371	461	480	396	465	466	*431*
2009	556	553	451	441	354	383	482	478	404	455	448	*426*
2010	588	569	478	447	368	399	483	491	408	471	404	*438*
2011	603	582	505	465	375	401	486	503	414	483	481	*453*
2012	589	581	496	466	376	407	490	518	417	487	491	*454*
2013	589	588	497	474	381	413	501	525	415	492	494	*459*
2014	595	604	515	490	393	419	514	539	433	503	502	*472*
2015	624	624	538	518	412	422	524	559	449	524	527	*493*
2016	687	662	570	545	436	432	572	599	480	568	566	*525*
2017	717	706	607	596	473	451	610	668	536	617	619	*571*
Total	*536*	*540*	*445*	*438*	*350*	*358*	*463*	*479*	*396*	*452*	*434*	*419*

Change in Value Added Per Employee

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998												
1999	-15	14	17	23	20	17	21	32	-20	18	12	*19*
2000	11	15	22	20	13	6	23	7	8	15	20	*16*
2001	29	8	7	19	7	15	21	5	6	7	18	*11*
2002	18	22	6	12	14	20	21	-7	29	3	5	*10*
2003	19	29	11	9	9	15	31	-4	-21	10	0	*10*
2004	72	28	21	16	10	8	17	19	18	13	12	*14*
2005	29	36	25	23	15	21	9	30	0	21	12	*19*
2006	51	2	31	29	23	20	18	37	29	26	20	*25*
2007	-32	39	39	35	19	22	19	25	10	34	44	*29*
2008	12	6	-2	15	0	8	19	8	20	7	14	*6*
2009	-11	0	-34	-7	10	18	33	7	-2	-10	-22	*-2*
2010	65	22	26	24	25	22	9	23	18	21	26	*23*
2011	14	30	30	29	15	16	14	17	19	17	13	*20*
2012	-27	-5	-12	1	7	13	10	17	9	1	2	*3*
2013	-7	5	-1	14	16	12	20	15	18	6	5	*11*
2014	-16	24	20	24	19	18	21	24	44	15	4	*19*
2015	73	14	28	36	27	11	21	36	12	21	22	*26*
2016	8	30	28	30	27	12	37	49	18	39	43	*32*
2017	58	32	26	26	22	7	18	53	26	23	45	*25*
Total	*17*	*0*	*15*	*20*	*16*	*15*	*20*	*21*	*13*	*15*	*15*	*17*

Depreciation to Sales, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	5,5%	10,3%	3,2%	3,2%	2,5%	1,8%	4,2%	3,1%	4,6%	5,1%	7,2%	*3,7%*
1999	5,9%	10,4%	3,4%	3,2%	2,6%	1,9%	4,3%	3,2%	4,8%	5,2%	7,6%	*3,8%*
2000	6,0%	10,4%	3,3%	3,2%	2,6%	1,8%	4,2%	3,3%	4,7%	5,1%	7,5%	*3,8%*
2001	5,8%	10,7%	3,4%	3,3%	2,7%	1,8%	4,1%	3,5%	4,9%	5,1%	7,6%	*3,8%*
2002	5,8%	10,4%	3,5%	3,4%	2,7%	1,8%	4,1%	3,5%	4,6%	5,0%	7,6%	*3,9%*
2003	6,6%	10,3%	3,5%	3,4%	2,7%	1,8%	3,9%	3,8%	4,7%	4,9%	8,0%	*3,9%*
2004	6,5%	10,1%	3,4%	3,3%	2,7%	1,8%	3,9%	3,4%	4,2%	4,8%	8,0%	*3,8%*
2005	6,3%	10,1%	3,1%	3,2%	2,7%	1,8%	3,9%	3,0%	3,9%	4,5%	7,9%	*3,7%*
2006	6,0%	10,8%	2,9%	3,0%	2,7%	1,7%	3,8%	2,9%	4,0%	4,3%	7,6%	*3,5%*
2007	6,5%	10,2%	2,8%	3,0%	2,6%	1,7%	3,5%	2,6%	3,6%	4,0%	7,2%	*3,4%*
2008	6,4%	10,7%	3,0%	3,0%	2,7%	1,7%	3,3%	2,7%	3,7%	3,9%	7,0%	*3,4%*
2009	6,9%	11,1%	3,4%	3,1%	2,7%	1,7%	3,1%	2,5%	3,7%	4,1%	7,5%	*3,5%*
2010	6,3%	10,3%	3,1%	2,9%	2,6%	1,6%	2,9%	2,3%	3,3%	3,8%	6,3%	*3,3%*
2011	6,2%	9,8%	2,8%	2,8%	2,5%	1,7%	2,7%	2,0%	2,8%	3,4%	7,0%	*3,1%*
2012	6,3%	10,2%	2,8%	2,7%	2,4%	1,6%	2,7%	1,9%	2,6%	3,3%	6,9%	*3,0%*
2013	6,2%	10,1%	2,8%	2,7%	2,4%	1,6%	2,5%	1,8%	2,5%	3,2%	6,7%	*2,9%*
2014	6,6%	10,1%	2,7%	2,6%	2,3%	1,6%	2,4%	1,7%	2,4%	3,1%	6,6%	*2,9%*
2015	6,6%	10,1%	2,6%	2,5%	2,3%	1,6%	2,4%	1,8%	2,3%	3,0%	6,6%	*2,8%*
2016	6,2%	10,0%	2,6%	2,4%	2,2%	1,6%	2,3%	1,7%	2,3%	2,8%	6,5%	*2,7%*
2017	5,5%	9,6%	2,5%	2,4%	2,2%	1,6%	2,4%	1,8%	2,3%	2,8%	6,2%	*2,7%*
Total	*6,2%*	*7,2%*	*3,1%*	*2,9%*	*2,5%*	*1,7%*	*3,2%*	*2,6%*	*3,5%*	*4,0%*	*7,2%*	*3,4%*

The table reports the average of a few selected fixed cost and efficiency measures for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. Value added per employee is defined as the operating income, but before costs for personnel, divided by the number of employees in the business. It can be seen as both an efficiency measure and a scalability measure. Moreover, it is often preferred over EBIT since it is difficult to separate salary from dividends in small companies. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.9: Average Industry Auditor Remarks by Year, 1998-2017

Auditor Recommends, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	90,7%	91,5%	91,3%	90,4%	87,8%	86,5%	92,3%	92,5%	87,5%	91,9%	92,7%	*90,0%*
1999	93,6%	91,3%	91,4%	90,1%	87,3%	85,4%	92,1%	91,8%	86,9%	91,1%	92,2%	*89,5%*
2000	93,4%	91,2%	91,6%	90,5%	87,1%	85,9%	92,8%	92,0%	86,5%	90,7%	93,1%	*89,6%*
2001	92,4%	91,6%	90,7%	90,2%	86,4%	84,5%	91,7%	90,5%	86,0%	90,0%	92,0%	*88,9%*
2002	93,0%	90,4%	90,4%	88,5%	85,4%	84,3%	91,0%	89,9%	84,5%	89,5%	91,1%	*88,1%*
2003	92,5%	91,0%	89,6%	87,9%	83,6%	83,4%	90,3%	88,2%	86,5%	89,2%	90,4%	*87,2%*
2004	91,7%	90,8%	89,0%	86,7%	82,0%	81,2%	89,1%	88,8%	82,6%	88,5%	89,8%	*86,1%*
2005	89,2%	89,9%	88,1%	86,0%	80,9%	79,9%	88,1%	88,4%	81,7%	87,9%	89,0%	*85,3%*
2006	89,6%	89,5%	88,1%	85,7%	79,9%	80,1%	87,8%	87,4%	81,8%	87,5%	88,1%	*84,8%*
2007	86,4%	87,9%	88,5%	84,9%	79,2%	79,9%	87,2%	87,7%	81,1%	86,9%	89,3%	*84,4%*
2008	87,1%	88,7%	88,4%	85,1%	78,3%	79,6%	87,1%	87,7%	83,1%	87,0%	88,9%	*84,2%*
2009	86,3%	87,6%	87,0%	84,8%	78,4%	80,6%	86,8%	88,0%	83,3%	86,9%	89,1%	*84,1%*
2010	85,5%	86,6%	86,6%	82,2%	77,0%	80,2%	85,4%	87,2%	81,3%	85,4%	79,4%	*82,3%*
2011	76,5%	78,1%	78,9%	69,7%	65,4%	72,2%	70,1%	69,3%	63,9%	70,5%	73,8%	*70,0%*
2012	71,4%	73,0%	74,8%	62,9%	58,9%	67,6%	63,5%	62,0%	58,9%	63,5%	68,3%	*63,6%*
2013	71,4%	68,0%	71,2%	57,0%	53,9%	65,3%	58,0%	55,2%	51,8%	57,9%	63,8%	*58,4%*
2014	68,1%	66,4%	69,0%	52,7%	50,4%	61,7%	53,5%	51,1%	47,6%	53,3%	60,9%	*54,5%*
2015	64,0%	65,4%	66,8%	49,7%	47,9%	59,9%	49,5%	47,1%	43,6%	49,5%	59,0%	*51,4%*
2016	65,8%	64,1%	65,5%	47,6%	45,7%	58,2%	48,1%	45,1%	41,1%	47,4%	54,9%	*49,3%*
2017	63,6%	65,0%	65,9%	49,9%	46,8%	59,8%	47,6%	46,7%	45,7%	47,9%	55,0%	*50,3%*
Total	*83,2%*	*7,2%*	*84,3%*	*74,0%*	*72,1%*	*76,4%*	*76,0%*	*75,6%*	*70,9%*	*76,5%*	*79,9%*	*75,8%*

Auditor Recommends with Remarks, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	7,5%	7,0%	7,6%	8,3%	10,2%	11,0%	6,7%	6,1%	10,8%	6,8%	6,3%	*8,4%*
1999	4,9%	7,6%	7,2%	8,6%	10,5%	12,4%	6,9%	7,0%	10,5%	7,6%	6,7%	*8,8%*
2000	5,3%	7,5%	6,9%	8,2%	10,6%	11,5%	6,2%	6,9%	10,1%	7,9%	5,9%	*8,6%*
2001	6,4%	7,3%	8,0%	8,6%	11,4%	13,1%	7,2%	8,2%	11,0%	8,6%	6,7%	*9,4%*
2002	6,0%	8,5%	8,3%	10,1%	12,1%	13,0%	7,8%	8,5%	13,4%	9,1%	8,0%	*10,1%*
2003	5,9%	8,3%	9,0%	10,5%	13,6%	13,8%	8,4%	10,0%	11,3%	9,5%	8,7%	*10,9%*
2004	6,9%	8,5%	9,5%	11,8%	14,8%	15,5%	9,1%	9,9%	14,7%	9,9%	9,1%	*11,8%*
2005	8,1%	9,2%	10,3%	12,2%	15,8%	16,3%	10,4%	10,2%	15,2%	10,5%	10,0%	*12,5%*
2006	8,6%	9,7%	10,4%	12,6%	16,6%	16,2%	10,6%	11,0%	15,7%	10,9%	10,5%	*13,0%*
2007	10,1%	11,2%	10,1%	13,5%	17,4%	16,4%	11,4%	11,1%	15,9%	11,5%	9,8%	*13,5%*
2008	11,4%	10,3%	10,1%	13,4%	18,3%	16,8%	11,6%	10,9%	14,7%	11,5%	10,2%	*13,7%*
2009	11,9%	11,3%	11,5%	13,8%	18,2%	17,1%	11,9%	10,3%	14,9%	11,7%	9,8%	*13,9%*
2010	12,6%	11,9%	11,5%	14,5%	18,8%	16,6%	12,1%	10,0%	15,3%	12,0%	10,4%	*14,3%*
2011	12,5%	9,8%	9,7%	12,3%	15,3%	15,6%	10,3%	7,9%	14,3%	9,9%	9,4%	*12,0%*
2012	12,4%	10,7%	9,5%	11,2%	13,6%	14,7%	8,8%	6,9%	11,1%	8,7%	8,1%	*10,8%*
2013	10,1%	12,5%	9,7%	10,9%	12,3%	12,1%	7,7%	7,0%	10,2%	8,3%	8,4%	*10,2%*
2014	10,0%	11,1%	9,3%	10,0%	11,2%	11,5%	7,0%	6,2%	9,1%	7,4%	7,5%	*9,2%*
2015	10,4%	10,5%	8,5%	8,8%	9,9%	10,6%	6,4%	5,5%	7,5%	6,6%	6,9%	*8,3%*
2016	7,0%	10,7%	8,5%	9,0%	9,3%	10,6%	6,1%	5,1%	7,4%	6,4%	8,4%	*8,1%*
2017	8,1%	8,0%	7,8%	8,4%	9,0%	10,0%	5,5%	4,2%	6,5%	5,9%	7,0%	*7,5%*
Total	*8,9%*	*7,2%*	*9,1%*	*10,9%*	*13,5%*	*13,8%*	*8,7%*	*8,1%*	*11,9%*	*9,1%*	*8,4%*	*10,8%*

Auditor Does Not Recommend, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	1,6%	1,5%	1,1%	1,3%	2,0%	2,5%	1,0%	1,4%	1,6%	1,3%	1,0%	*1,6%*
1999	1,5%	1,1%	1,4%	1,3%	2,1%	2,2%	1,0%	1,2%	2,6%	1,3%	1,1%	*1,6%*
2000	1,3%	1,3%	1,4%	1,2%	2,3%	2,5%	1,0%	1,0%	3,3%	1,4%	1,0%	*1,7%*
2001	1,3%	1,0%	1,3%	1,3%	2,2%	2,3%	1,0%	1,3%	2,9%	1,4%	1,2%	*1,7%*
2002	1,0%	1,1%	1,3%	1,3%	2,5%	2,7%	1,2%	1,5%	2,1%	1,4%	0,9%	*1,7%*
2003	1,6%	0,8%	1,4%	1,6%	2,8%	2,9%	1,3%	1,8%	2,2%	1,3%	0,9%	*1,9%*
2004	1,4%	0,7%	1,5%	1,6%	3,2%	3,3%	1,8%	1,3%	2,7%	1,6%	1,1%	*2,1%*
2005	2,7%	0,9%	1,6%	1,7%	3,3%	3,8%	1,5%	1,3%	3,0%	1,5%	1,0%	*2,2%*
2006	1,8%	0,8%	1,5%	1,7%	3,5%	3,6%	1,6%	1,5%	2,5%	1,5%	1,4%	*2,2%*
2007	3,4%	0,9%	1,4%	1,6%	3,4%	3,6%	1,3%	1,2%	3,0%	1,5%	0,9%	*2,1%*
2008	1,4%	1,0%	1,5%	1,5%	3,4%	3,6%	1,2%	1,3%	2,2%	1,5%	0,9%	*2,1%*
2009	1,8%	1,1%	1,5%	1,4%	3,4%	2,4%	1,3%	1,8%	1,7%	1,4%	1,1%	*2,0%*
2010	1,2%	0,9%	1,4%	1,4%	2,9%	2,7%	1,3%	0,9%	2,3%	1,2%	1,0%	*1,8%*
2011	1,5%	0,6%	1,1%	1,3%	2,1%	1,8%	0,9%	0,9%	1,5%	0,8%	0,7%	*1,3%*
2012	1,3%	0,6%	1,0%	1,1%	2,0%	1,8%	0,7%	0,7%	1,6%	0,8%	0,6%	*1,2%*
2013	1,5%	0,6%	0,9%	0,8%	1,7%	1,7%	0,7%	0,6%	1,0%	0,6%	0,8%	*1,0%*
2014	1,8%	0,4%	0,9%	0,8%	1,3%	1,7%	0,5%	0,4%	0,7%	0,5%	0,8%	*0,9%*
2015	0,9%	0,6%	1,0%	0,7%	1,2%	1,7%	0,4%	0,4%	1,1%	0,5%	0,8%	*0,8%*
2016	1,2%	0,3%	0,9%	0,6%	1,1%	1,5%	0,3%	0,3%	0,8%	0,3%	0,5%	*0,7%*
2017	2,4%	0,8%	0,7%	0,5%	0,8%	0,9%	0,2%	0,3%	0,3%	0,2%	0,3%	*0,5%*
Total	*1,6%*	*7,2%*	*1,3%*	*1,2%*	*2,4%*	*2,5%*	*1,0%*	*1,0%*	*1,9%*	*1,1%*	*0,9%*	*1,5%*

Auditor Remarks Missing, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
1999	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2000	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	*0,0%*
2001	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2002	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2003	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2004	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2005	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2006	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2007	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2008	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	*0,0%*
2009	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	*0,0%*
2010	0,6%	0,6%	0,6%	1,9%	1,3%	0,5%	1,2%	1,9%	1,2%	1,3%	9,2%	*1,7%*
2011	9,6%	11,5%	10,3%	16,7%	17,1%	10,4%	18,7%	21,8%	20,3%	18,8%	16,2%	*16,7%*
2012	14,9%	15,7%	14,7%	24,9%	25,5%	15,9%	27,0%	30,5%	28,4%	27,1%	22,9%	*24,4%*
2013	16,9%	18,9%	18,1%	31,3%	32,1%	20,9%	33,6%	37,2%	36,9%	33,2%	27,0%	*30,4%*
2014	20,0%	22,1%	20,9%	36,5%	37,0%	25,0%	38,9%	42,3%	42,7%	38,8%	30,8%	*35,4%*
2015	24,7%	23,5%	23,7%	40,8%	41,0%	27,8%	43,7%	47,1%	47,8%	43,4%	33,3%	*39,5%*
2016	25,9%	24,8%	25,1%	42,8%	44,0%	29,7%	45,5%	49,6%	50,6%	46,0%	36,2%	*42,0%*
2017	25,9%	26,2%	25,6%	41,2%	43,4%	29,3%	46,6%	48,8%	47,5%	46,0%	37,7%	*41,7%*
Total	*6,3%*	*7,2%*	*5,3%*	*13,9%*	*12,0%*	*7,3%*	*14,3%*	*15,2%*	*15,3%*	*13,3%*	*10,8%*	*11,8%*

The table reports the average of auditor remark dummies, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. The audit obligation for companies below a certain size was partly repealed in 2010 (see grey shaded area), which is why there is a drastic uptick in the number of missing auditor remarks after that year. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.10: Average Industry Corporate Governance Indicators by Year, 1998-2017

	Late Filing Estimate, %											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	21,3%	17,7%	20,9%	21,6%	23,2%	25,7%	20,5%	23,1%	28,3%	20,8%	22,3%	*22,1%*
1999	12,8%	14,0%	13,9%	15,7%	16,9%	21,3%	14,0%	16,4%	21,9%	15,0%	15,7%	*16,1%*
2000	11,9%	12,2%	12,8%	14,6%	15,7%	18,0%	14,0%	15,2%	18,7%	14,1%	15,6%	*14,8%*
2001	10,2%	7,2%	9,5%	10,9%	12,0%	14,5%	10,3%	12,1%	13,4%	10,1%	10,3%	*11,0%*
2002	8,4%	6,2%	7,9%	9,5%	10,8%	12,4%	8,5%	10,8%	14,5%	8,9%	8,2%	*9,7%*
2003	4,1%	5,7%	7,1%	7,4%	8,9%	9,3%	7,1%	8,8%	9,7%	7,1%	6,0%	*7,8%*
2004	10,4%	8,6%	10,4%	11,7%	13,2%	15,2%	12,0%	12,4%	15,4%	11,1%	10,4%	*12,0%*
2005	11,9%	9,3%	11,6%	12,3%	14,9%	18,3%	13,1%	14,2%	18,7%	12,3%	11,3%	*13,4%*
2006	10,0%	9,2%	10,9%	11,9%	13,8%	18,0%	11,8%	12,7%	15,1%	11,3%	10,3%	*12,5%*
2007	8,1%	8,5%	8,9%	10,4%	11,4%	13,6%	10,2%	10,6%	15,0%	9,8%	8,4%	*10,5%*
2008	6,8%	4,8%	6,3%	7,4%	9,5%	9,8%	8,0%	7,5%	10,7%	7,2%	6,2%	*7,9%*
2009	6,2%	3,7%	5,2%	6,6%	8,4%	8,0%	5,7%	6,6%	8,5%	6,2%	6,8%	*6,8%*
2010	7,9%	6,9%	6,7%	8,7%	10,9%	10,1%	8,7%	7,8%	13,4%	8,0%	7,9%	*8,9%*
2011	3,5%	2,9%	3,6%	4,6%	5,9%	4,9%	4,7%	3,9%	7,2%	4,1%	4,2%	*4,7%*
2012	4,6%	1,8%	2,6%	3,8%	4,5%	4,3%	3,2%	3,0%	5,6%	3,2%	2,8%	*3,6%*
2013	6,8%	4,8%	5,6%	7,1%	9,5%	8,0%	7,3%	6,9%	10,6%	6,7%	6,3%	*7,5%*
2014	3,2%	1,9%	2,8%	3,8%	4,7%	4,5%	3,9%	3,5%	5,7%	3,4%	2,9%	*3,8%*
2015	3,8%	2,8%	3,1%	4,7%	5,8%	5,0%	4,3%	5,0%	6,5%	4,6%	4,7%	*4,8%*
2016	3,0%	2,0%	2,9%	4,3%	5,4%	5,4%	4,1%	4,0%	4,1%	3,8%	3,1%	*4,3%*
2017	16,8%	21,7%	18,1%	21,1%	21,5%	21,1%	18,9%	17,0%	18,7%	18,1%	25,4%	*20,0%*
Total	*8,5%*	*7,2%*	*8,8%*	*9,3%*	*11,1%*	*12,4%*	*9,0%*	*9,5%*	*12,5%*	*8,9%*	*9,0%*	*9,8%*

	Companies with Record From General Meetings of Shareholders, %											
Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	32,9%	30,3%	34,0%	30,7%	25,5%	28,0%	33,2%	33,3%	23,9%	32,8%	31,0%	*30,0%*
1999	31,3%	31,0%	34,7%	31,5%	26,9%	27,5%	34,4%	30,8%	21,9%	32,2%	31,4%	*30,4%*
2000	33,8%	28,3%	34,0%	31,7%	25,6%	27,3%	32,7%	27,4%	22,3%	30,8%	30,1%	*29,4%*
2001	32,7%	27,3%	31,7%	29,9%	24,3%	28,6%	31,0%	25,3%	20,6%	29,1%	27,3%	*27,9%*
2002	30,7%	26,7%	30,3%	28,0%	23,9%	28,9%	30,8%	22,9%	18,9%	27,1%	27,0%	*26,8%*
2003	29,4%	26,5%	29,6%	26,8%	22,8%	29,8%	32,9%	22,8%	20,3%	27,3%	26,1%	*26,4%*
2004	33,1%	30,1%	31,9%	29,8%	23,8%	29,0%	34,0%	27,3%	21,1%	30,0%	29,0%	*28,4%*
2005	38,1%	37,0%	39,9%	38,6%	29,4%	31,9%	45,2%	40,3%	30,7%	40,0%	33,4%	*36,3%*
2006	42,0%	37,5%	45,7%	43,4%	32,6%	36,2%	50,3%	44,2%	35,6%	44,9%	35,7%	*40,6%*
2007	39,1%	37,5%	45,7%	44,5%	31,6%	39,7%	48,6%	46,4%	37,6%	45,6%	36,5%	*40,9%*
2008	35,9%	36,3%	39,1%	40,8%	28,9%	39,3%	48,6%	43,8%	32,3%	42,6%	34,9%	*38,0%*
2009	40,6%	38,9%	39,3%	41,6%	31,8%	42,5%	49,6%	45,0%	33,8%	43,9%	37,5%	*39,7%*
2010	36,0%	34,3%	38,3%	40,2%	29,4%	39,1%	47,5%	45,4%	32,2%	42,5%	30,0%	*37,7%*
2011	35,2%	32,0%	38,0%	39,8%	29,4%	39,4%	46,5%	46,1%	32,8%	43,0%	28,7%	*37,7%*
2012	32,6%	28,3%	35,9%	37,9%	29,5%	38,9%	47,4%	45,7%	32,1%	42,7%	28,6%	*37,1%*
2013	18,7%	14,9%	22,5%	21,0%	17,2%	22,7%	27,3%	28,7%	18,3%	24,9%	14,6%	*21,5%*
2014	0,5%	1,0%	1,0%	0,8%	0,9%	0,8%	1,9%	1,4%	0,8%	1,0%	0,8%	*1,0%*
2015	0,5%	0,7%	0,7%	0,8%	0,7%	0,6%	0,8%	0,9%	0,9%	0,8%	0,7%	*0,8%*
2016	0,2%	0,4%	0,4%	0,6%	0,6%	0,5%	0,7%	0,9%	0,6%	0,7%	0,5%	*0,6%*
2017	0,7%	0,5%	0,5%	0,4%	0,4%	0,6%	0,6%	0,8%	0,7%	0,6%	0,5%	*0,5%*
Total	*28,2%*	*7,2%*	*30,6%*	*26,9%*	*21,9%*	*27,2%*	*31,6%*	*28,7%*	*21,6%*	*29,1%*	*24,2%*	*26,6%*

The table reports variables related to corporate governance and reporting, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. The late filing dummy is computed by comparing the date of the underlying financial statement (plus the allowed time before the statement is due) to the time-stamp of the most recent change recorded for said financial statement. As explained by correspondence with Bisnode, financial accounts' time of last change are not necessarily the same dates and times that the statements were received by the Swedish Company Registration Office: some later, manual corrections to the data have been made by Bisnode themselves for a portion of the companies. Although the data has been re-adjusted to correct for this on dates with a strikingly high number of coincident last-change-time-stamps (i.e. Bisnode modifications), the late filing dummy remains an imperfect estimate. AGM is dummy for a record from the general meeting of shareholders, and decrease drastically in 2014. No explanation has been found for why this is the case. Companies with missing industry classification are not identifiable in the table, but are included in the totals.

Table 10.11: Average Industry Miscellaneous Statistics by Year, 1998-2017

Logarithm of Assets (In SEKk)

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	7,9	7,9	7,7	7,2	7,2	7,4	6,9	7,2	7,2	7,2	7,6	*7,3*
1999	8,0	7,9	7,7	7,2	7,3	7,4	7,0	7,2	7,2	7,2	7,6	*7,3*
2000	8,0	8,0	7,8	7,3	7,3	7,5	7,1	7,2	7,3	7,3	7,7	*7,4*
2001	8,1	8,0	7,8	7,4	7,3	7,5	7,1	7,2	7,2	7,3	7,7	*7,4*
2002	8,1	8,0	7,8	7,4	7,3	7,5	7,1	7,2	7,2	7,3	7,8	*7,4*
2003	8,3	8,1	7,8	7,4	7,4	7,6	7,2	7,2	7,2	7,3	7,9	*7,4*
2004	8,3	8,1	7,9	7,5	7,4	7,5	7,3	7,3	7,2	7,3	7,9	*7,5*
2005	8,4	8,2	7,9	7,5	7,4	7,6	7,3	7,3	7,3	7,4	7,9	*7,5*
2006	8,4	8,3	8,0	7,6	7,4	7,6	7,3	7,4	7,3	7,4	7,9	*7,5*
2007	8,4	8,3	8,0	7,6	7,4	7,7	7,3	7,4	7,4	7,5	8,0	*7,6*
2008	8,3	8,3	8,0	7,6	7,4	7,7	7,3	7,4	7,3	7,5	8,1	*7,6*
2009	8,4	8,3	8,0	7,6	7,4	7,7	7,4	7,4	7,3	7,5	8,1	*7,6*
2010	8,4	8,3	8,0	7,5	7,4	7,7	7,4	7,4	7,3	7,5	7,9	*7,6*
2011	8,5	8,3	8,0	7,5	7,4	7,7	7,4	7,3	7,2	7,5	8,1	*7,5*
2012	8,4	8,2	8,0	7,5	7,4	7,7	7,3	7,3	7,2	7,4	8,1	*7,5*
2013	8,4	8,2	8,0	7,5	7,3	7,7	7,3	7,3	7,2	7,4	8,1	*7,5*
2014	8,5	8,2	8,0	7,5	7,3	7,7	7,3	7,4	7,2	7,4	8,1	*7,5*
2015	8,4	8,2	8,0	7,5	7,3	7,7	7,4	7,4	7,2	7,4	8,1	*7,5*
2016	8,5	8,3	8,1	7,5	7,4	7,7	7,4	7,5	7,3	7,5	8,1	*7,6*
2017	8,6	8,3	8,1	7,6	7,4	7,7	7,5	7,6	7,4	7,6	8,2	*7,6*
Total	*8,3*	*8,2*	*7,9*	*7,5*	*7,4*	*7,6*	*7,3*	*7,3*	*7,2*	*7,4*	*7,9*	*7,5*

Logarithm of Sales (In SEKk)

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	8,4	8,1	8,3	7,9	8,1	8,7	7,5	7,7	7,8	7,7	7,8	*8,0*
1999	8,4	8,1	8,3	8,0	8,2	8,7	7,5	7,7	7,7	7,7	7,8	*8,0*
2000	8,5	8,2	8,4	8,1	8,2	8,8	7,6	7,6	7,7	7,7	7,8	*8,0*
2001	8,6	8,2	8,4	8,1	8,2	8,8	7,6	7,7	7,7	7,7	7,9	*8,1*
2002	8,6	8,3	8,4	8,1	8,2	8,8	7,7	7,6	7,7	7,7	7,9	*8,1*
2003	8,6	8,3	8,4	8,1	8,2	8,8	7,8	7,6	7,7	7,7	7,9	*8,1*
2004	8,6	8,4	8,4	8,2	8,2	8,8	7,8	7,7	7,7	7,8	8,0	*8,1*
2005	8,7	8,4	8,5	8,2	8,2	8,8	7,8	7,7	7,7	7,8	8,0	*8,1*
2006	8,7	8,4	8,5	8,3	8,3	8,8	7,8	7,8	7,7	7,8	8,0	*8,2*
2007	8,7	8,4	8,6	8,3	8,3	8,9	7,8	7,8	7,8	7,9	8,1	*8,2*
2008	8,7	8,4	8,6	8,3	8,3	9,0	7,8	7,8	7,8	7,9	8,1	*8,2*
2009	8,6	8,4	8,5	8,3	8,3	9,0	7,9	7,7	7,8	7,9	8,1	*8,2*
2010	8,7	8,4	8,5	8,2	8,2	9,0	7,9	7,7	7,7	7,9	7,7	*8,1*
2011	8,7	8,5	8,6	8,3	8,2	8,9	7,9	7,7	7,7	7,9	8,1	*8,2*
2012	8,7	8,4	8,5	8,3	8,2	8,9	7,9	7,7	7,7	7,9	8,1	*8,1*
2013	8,7	8,4	8,5	8,2	8,2	8,9	7,9	7,7	7,7	7,8	8,1	*8,1*
2014	8,7	8,4	8,5	8,3	8,2	8,9	7,9	7,7	7,7	7,8	8,1	*8,1*
2015	8,7	8,5	8,6	8,3	8,2	8,9	7,9	7,7	7,7	7,9	8,1	*8,2*
2016	8,8	8,5	8,6	8,4	8,2	8,9	8,0	7,7	7,7	7,9	8,2	*8,2*
2017	9,0	8,6	8,7	8,5	8,4	9,0	8,1	7,9	7,9	8,1	8,3	*8,3*
Total	*8,6*	*8,4*	*8,5*	*8,2*	*8,2*	*8,9*	*7,8*	*7,7*	*7,7*	*7,8*	*8,0*	*8,1*

Dividend Paid, %

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	35,0%	37,0%	39,0%	35,0%	28,0%	32,0%	39,0%	38,0%	26,0%	38,0%	35,0%	*34,0%*
1999	35,0%	34,0%	38,0%	35,0%	28,0%	31,0%	38,0%	36,0%	25,0%	37,0%	32,0%	*33,0%*
2000	36,0%	32,0%	38,0%	35,0%	27,0%	31,0%	39,0%	32,0%	24,0%	35,0%	32,0%	*33,0%*
2001	35,0%	30,0%	36,0%	34,0%	26,0%	31,0%	36,0%	30,0%	22,0%	33,0%	30,0%	*31,0%*
2002	33,0%	30,0%	34,0%	33,0%	25,0%	33,0%	36,0%	27,0%	22,0%	31,0%	28,0%	*30,0%*
2003	32,0%	31,0%	33,0%	31,0%	24,0%	31,0%	38,0%	25,0%	22,0%	31,0%	27,0%	*29,0%*
2004	36,0%	34,0%	35,0%	34,0%	25,0%	32,0%	40,0%	31,0%	23,0%	35,0%	30,0%	*32,0%*
2005	42,0%	44,0%	45,0%	45,0%	32,0%	36,0%	52,0%	46,0%	36,0%	47,0%	39,0%	*41,0%*
2006	49,0%	45,0%	52,0%	51,0%	36,0%	41,0%	58,0%	52,0%	41,0%	53,0%	40,0%	*47,0%*
2007	44,0%	45,0%	53,0%	52,0%	36,0%	42,0%	58,0%	55,0%	43,0%	54,0%	42,0%	*47,0%*
2008	44,0%	44,0%	48,0%	50,0%	32,0%	41,0%	57,0%	53,0%	39,0%	52,0%	40,0%	*45,0%*
2009	43,0%	45,0%	45,0%	50,0%	34,0%	43,0%	58,0%	53,0%	40,0%	51,0%	43,0%	*45,0%*
2010	40,0%	42,0%	44,0%	47,0%	32,0%	40,0%	55,0%	54,0%	37,0%	50,0%	33,0%	*43,0%*
2011	37,0%	35,0%	42,0%	45,0%	29,0%	36,0%	52,0%	52,0%	36,0%	48,0%	29,0%	*41,0%*
2012	37,0%	32,0%	40,0%	42,0%	27,0%	35,0%	52,0%	51,0%	35,0%	47,0%	27,0%	*39,0%*
2013	33,0%	34,0%	40,0%	43,0%	28,0%	36,0%	53,0%	53,0%	34,0%	47,0%	28,0%	*40,0%*
2014	35,0%	37,0%	42,0%	45,0%	30,0%	37,0%	55,0%	53,0%	35,0%	50,0%	29,0%	*42,0%*
2015	37,0%	41,0%	45,0%	49,0%	33,0%	38,0%	56,0%	57,0%	37,0%	52,0%	34,0%	*45,0%*
2016	42,0%	41,0%	48,0%	50,0%	35,0%	38,0%	57,0%	58,0%	39,0%	55,0%	36,0%	*46,0%*
2017	36,0%	35,0%	42,0%	44,0%	31,0%	33,0%	53,0%	54,0%	37,0%	50,0%	30,0%	*42,0%*
Total	*38,0%*	*38,0%*	*42,0%*	*43,0%*	*30,0%*	*36,0%*	*50,0%*	*46,0%*	*33,0%*	*45,0%*	*33,0%*	*39,0%*

Untaxed Reserves Per Employee, SEKk

Year	Energy & Envir.	Materl's	Ind. Goods	Constr. Ind.	Shop. Goods	Conv. Goods	Health & Educ.	IT & Electr.	Telec. & Media.	Corp. Serv.	Other	*Total*
1998	108,8	126,6	66,7	38,0	37,9	37,9	46,1	56,9	47,6	60,6	111,9	*52,6*
1999	117,1	126,3	72,9	42,2	41,3	41,9	51,9	63,9	50,6	65,7	114,0	*57,0*
2000	127,3	129,5	75,7	46,1	42,1	42,7	56,8	64,4	56,0	67,8	119,1	*59,5*
2001	142,7	134,9	80,1	52,0	44,6	44,8	61,5	67,0	57,3	71,5	130,5	*63,6*
2002	147,9	147,5	83,1	57,1	48,1	50,4	66,2	68,1	56,1	74,2	138,9	*67,5*
2003	162,5	157,6	86,7	61,8	50,8	53,7	74,1	68,1	60,8	77,8	149,4	*71,6*
2004	162,3	165,4	83,8	61,3	48,3	49,4	74,3	63,5	54,0	75,9	148,1	*69,6*
2005	157,7	171,5	77,1	58,0	42,9	41,6	62,4	54,4	51,2	70,3	145,8	*64,2*
2006	161,5	176,3	79,2	60,8	42,7	41,9	61,8	53,7	46,8	71,9	145,5	*65,1*
2007	159,5	177,3	86,6	65,2	42,8	43,0	62,1	55,8	47,4	74,5	158,8	*67,9*
2008	159,9	182,1	94,9	70,9	42,7	45,8	66,6	59,0	48,9	78,8	174,8	*72,0*
2009	169,0	188,1	93,9	71,2	42,9	48,8	71,7	61,5	52,7	80,8	179,1	*73,4*
2010	169,2	189,9	93,1	69,9	43,3	49,4	74,2	64,9	52,0	82,5	159,1	*73,8*
2011	186,5	191,9	98,8	73,8	45,2	49,8	78,2	67,1	50,4	86,2	188,2	*77,3*
2012	194,4	201,0	106,4	78,6	48,3	55,7	86,7	80,0	57,9	94,6	202,3	*83,8*
2013	192,0	202,8	106,8	78,2	47,5	56,6	88,0	85,0	56,4	96,4	203,6	*84,0*
2014	189,5	202,3	105,5	76,8	47,9	55,4	88,8	86,7	56,4	96,8	205,9	*83,9*
2015	186,1	208,8	106,4	77,4	47,9	51,7	87,7	89,8	56,0	97,0	214,7	*84,3*
2016	185,7	218,3	106,5	77,7	48,8	51,0	88,1	88,5	57,1	98,0	226,8	*85,3*
2017	187,5	223,9	110,3	83,6	52,5	53,9	91,4	95,7	58,5	103,9	227,9	*89,9*
Total	*162,7*	*176,7*	*88,5*	*66,7*	*45,4*	*48,1*	*73,3*	*70,4*	*53,7*	*81,8*	*167,7*	*72,5*

The table reports miscellaneous metrics, for all active Swedish, non-financial, independent limited liability companies with atleast two years of data. The results are presented by year and industry. The dividend variable is a dummy for companies paying a dividend. Companies with missing industry classification are not identifiable in the table, but are included in the totals.