Cyclicality and Firm-size in Private Firm Defaults

Thais Lærkholm Jensen, David Lando, Mamdouh Medhat*

This version: July 8, 2016

Abstract

The Basel II/III and CRD-IV Accords reduce capital charges on bank loans to smaller firms by assuming that the default probabilities of smaller firms are less sensitive to macroeconomic cycles. We test this assumption in a default intensity framework using a large sample of bank loans to private Danish firms. We find that controlling only for size, the default probabilities of small firms are, in fact, less cyclical than the default probabilities of large firms. However, accounting for firm characteristics other than size, we find that the default probabilities of small firms are equally cyclical or even more cyclical than the default probabilities of large firms. These results hold using a multiplicative Cox model as well as an additive Aalen model with time-varying coefficients.

Keywords: Capital charges · SME · Default risk · Macroeconomic cycles
JEL: G21, G28, G33, C41

*Lærkholm Jensen is at Department of Economics, University of Copenhagen, and the Danish Central Bank. Lando is at Department of Finance and FRIC Center for Financial Frictions, Copenhagen Business School. Medhat is at Faculty of Finance, Cass Business School, City University London. Send correspondence to David Lando, `dl.fi@cbs.dk`, +45 3815 3613, Solbjerg Plads 3, DK-2000, Frederiksberg, Denmark.

We are grateful to a Danish financial institution for providing data on loans to private firm. We thank Diana Bonfim, Laurence Deborgies-Sanches (discussant), Darrell Duffie, Johan Hombert, Harrison Hong (editor), Allan Mortensen, Lasse Heje Pedersen, Hanne Schultz, and attendees at FEBS/LabEx-ReFi 2013 for helpful discussions and suggestions. Support from the Center for Financial Frictions (FRIC), grant no. DNRF102, and Danmarks Nationalbank, is gratefully acknowledged.

1 Introduction

Small and medium-sized enterprises (SMEs) typically depend more heavily on funding from banks than do larger firms. It is therefore conceivable that SMEs are hit harder during a financial crisis in which banks' capital constraints are binding. As a way to facilitate bank funding of SMEs, the Basel II Accord prescribes lower capital charges for loans to the SME-segment. Technically, the reduction in capital charges is implemented by prescribing a lower asset correlation to be used when calculating capital charges. To the extent that asset correlation arises because of common dependence on macroeconomic shocks, the reduction corresponds to assuming that the default probabilities of SMEs are less sensitive to macroeconomic cycles than those of larger firms. These reductions in capital charges were recently reaffirmed and extended in the Basel III Accord and in the fourth Capital Requirements Directive (CRD-IV).[1]

This paper uses a large sample of loans to private Danish firms to test whether there is empirical support for the assumption that the default probabilities of smaller firms are less cyclical. Our results indicate that solely discriminating with respect to firm-size, the default probabilities of small firms are in fact less sensitive to macroeconomic cycles compared to the default probabilities of large firms. However, when we account for differences in firm-characteristics other than size, our results indicate that the default probabilities of small firms are as cyclical or even more cyclical than the default probabilities of large firms. This indicates that the size effect arises because of omitted variables. These results are robust to different regression models and different ways of dividing our sample into small and large firms. The evidence form our data suggests that a bank that properly accounts for firm characteristics in its default risk modeling will not experience a weaker effect of economic cycles in the SME segmemt, and hence the reduced capital charge will in fact make a bank with a high exposure to the SME-segment more risky.

According to the European Commission's (2016) annual report, SMEs comprised over 99% of all enterprises, accounted for 67% of total employment, and stood for over 70% of job growth in the non-financial business sector in the 28 EU countries. Hence their importance for the economy provides a political justification for the separate treatment. Furthermore, shifting banks' risk-exposure towards smaller firms could be in the interest of regulators—for instance because larger firms are more likely to have larger loans or loans across several banks, which could create contagion within and across banks in case of defaults. But such effects are already accounted for using so-called

[1]Article 273 of The Basel Committee on Banking Supervision's (2006) report states the following: *"Under the IRB approach for corporate credits, banks will be permitted to separately distinguish exposures to SME borrowers (defined as corporate exposures where the reported sales for the consolidated group of which the firm is a part is less than €50 million) from those to large firms. A firm-size adjustment [...] is made to the corporate risk weight formula for exposures to SME borrowers."*

‘granularity adjustments’. Our focus is on whether an additional adjustment of capital charges through lower correlation with the macroeconomic environment can be justified empirically.

Our data covers the period 2003-2012 and consists of obligor, loan, and default information for a large sample of private Danish firms. We devise two different ways of testing whether defaults of smaller firms are more or less cyclical. First, we use the standard multiplicative Cox regression model to estimate the effects of firm-specific and macroeconomic variables on default intensities. Our Cox models confirm previous findings that accounting ratios and macroeconomic variables play distinct roles in default prediction for private firms: (1) Accounting ratios are necessary for accurately ranking private firms according to default likelihood, but cannot by themselves capture the cyclicality of aggregate default rates, while (2) macroeconomic variables are indispensable for capturing the cyclicality of aggregate default rates, but do not aid in the ranking of firms with respect to default likelihood. Using our fitted Cox model, we find that when we split our sample with respect to firm size and keep all other firm characteristics fixed, the default probabilities of smaller firms do in fact exhibit less sensitivity to macroeconomic cycles. This is in the sense that the effects of macroeconomic variables are generally of smaller magnitude for smaller firms. However, when we account for the Cox model’s non-linear form and use averaging techniques adapted from other non-linear regression models, our results indicate that the default probability of the average small firm may be as cyclical or even more cyclical than the default probability of the average large firm.

Second, we investigate our data using the more flexible additive Aalen regression model. The additive model allows us to estimate time-varying effects of firm-specific variables, which is a potentially important component of the cyclicality of default probabilities that is missing from our Cox model. Furthermore, due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques. Our analysis based on the additive model reveals, in particular, that firm-size has a significantly time-varying and mostly negative effect. Hence, the effect of firm-size varies significantly the with business-cycle but is generally such that that larger firms have lower default probabilities. Moreover, the effect of firm-size reaches its largest (negative) magnitude during the financial crisis of ‘08-‘10, which indicates that larger firms were safer, not riskier, during the most recent recession. Lastly, using our fitted additive model and splitting the sample with respect to firm size, we again find no evidence that there is a difference in the signs, magnitudes, or significance of the effects of our macroeconomic variables for small firms compared to large firms. These results indicate that our main findings are robust to different regression models that capture cyclicality in different ways.

Our results indicate that a different treatment of capital charges solely based on firm-size is too simplistic, as it ignores other important characteristics that differ between small and large firms. This suggests that imposing solely size-based preferential treatment of capital charges may, in fact, increase the risks of banks with a high exposure to the SME-segment.

The outline of of the paper is as follows. Section 2 explains the piece of the Basel regulation that motivates our study. Section 3 reviews the literature. Section 4 details our data and variable selection. Section 5 provides our estimation methodology. Section 6 presents our regression results for identifying the firm-specific and macroeconomic variables that significantly predict defaults. Section 6 presents our results regarding the sensitivity of small and large firms' default probabilities to macroeconomic variables. Section 7 shows results related to robustness and model check. Section 8 concludes.

2 The Basel capital charge for loans

Basel II and III allow banks to estimate capital requirements for small and medium-sized corporations (SMEs) using a risk weight formula that includes a lower asset correlation with macroeconomic risk-drivers compared to that of larger corporations. The exact formulation in Basel II can be found in Articles 273 and 274 of The Basel Committee on Banking Supervision's (2006) report. The same provisions are carried forward and extended in the Basel III Accord and the recently adopted CRD-IV as Articles 153.4 and 501.1 of The European Parliament and the Council of the European Union's (2013) report.

The amount of equity capital required for funding a loan essentially depends on the unexpected loss (UL) per unit principal,

$$\mathrm{UL} = \mathrm{LGD}(\theta^* - \mathrm{PD}),$$

where PD is the loan's one-year probability of default, LGD is the loss rate given default, and

$$\theta^* = \Phi\left(\frac{1}{\sqrt{1-\rho}}\left(\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho}\Phi^{-1}(0.999)\right)\right). \tag{1}$$

Here, Φ is the standard normal distribution function and θ^* is the 99.9% worst case default rate. The correlation parameter, ρ, that goes into this calculation is PD-dependent and prescribed by the

Basel documentation (after removing the negligible terms exp(−50)) as

$$\rho(\mathrm{PD}) = 0.12\,(1 - \exp(-50\,\mathrm{PD})) + 0.24\exp(-50\,\mathrm{PD}).$$

For SMEs with annual sales, S, up to €50 million, this correlation is modified to

$$\rho(\mathrm{PD}) = 0.12\,(1 - \exp(-50\,\mathrm{PD})) + 0.24\exp(-50\,\mathrm{PD}) - 0.04\left(1 - \frac{\max(S-5;0)}{45}\right),$$

where the last term has the effect of lowering the capital requirement. The exact capital requirement depends on the contribution to Risk Weighted Assets (RWA), which depends on UL, the Exposure-at-Default (EAD), and a maturity adjustment whose exact form need not concern us here.

As an illustration of the economic magnitude of the reduction in capital requirement for SMEs, note that with an annual default probability of 1%, a loan to an SME with $S =$ €25 million achieves a deduction in asset correlation of 11.52% relative to an equally risky non-SME. Assuming an LGD of 15% and an effective maturity of 2.5 years (both provided as "representative averages" by the financial institution that provided us with the data), we calculate that this corresponds to deduction of 12.14% in capital requirement.

In the derivation of the capital charge formula, ρ represents correlation between asset values of different borrowers which arises because the asset values depend on a common economy-wide factor. Therefore, the reduction in correlation for SMEs corresponds to an assumption that these firms have default probabilities that are less sensitive to the economy-wide factor. It is this assumption that we will devise two ways of testing on a large Danish data set.

3 Related literature

There is conflicting evidence in the literature regarding the validity of the assumption in the Basel Accords that the default probabilities of smaller firms are less cyclical. Lopez (2004) employs the KMV approach on a sample of US, Japanese, and European firms and finds that average asset correlation is a decreasing function of probability of default and an increasing function of firm-size, which is line with the assumption in the Basel Accords. Dietsch and Petey (2004), however, use a one-factor credit risk model on a sample of French and German firms and find that SMEs have higher default risk than larger firms and that the asset correlations for SMEs are weak and decrease with firm-size. Chionsini, Marcucci, and Quagliariello (2010) use firm-size dependent linear regression of default probabilities on macroeconomic time-series for a sample of Italian firms and they find support for the size-dependent treatment in the Basel Accord, though not during severe financial crises like that of 2008-09. Jacobson, Linde, and Roszbach (2005) use a Monte-Carlo resampling method for a sample of Swedish firms and find little support for the hypothesis that SME loan portfolios are less risky, or require less economic capital, than corporate loans.

Our approach differs from these studies in that we use an intensity regression framework to directly estimate and study how default probabilities simultaneously depend on a large set of firm-characteristics, including size, as well as macroeconomic variables. This allows us to determine the firm-characteristics that are important for accurately ranking firms with respect to default likelihood as well as the macroeconomic variables important for capturing the cyclicality of default rates over time. Importantly, our approach has the advantage of allowing us to directly determine how the effects of macroeconomic variables on default probabilities differ for small and large firms while simultaneously controlling for firm-characteristics other than size. We apply averaging techniques to the Cox regression model resembling those commonly applied to generalized linear models—see, for instance, Wooldridge (2009, p. 582-83) for an overview. The additive regression model due to Aalen (1980, 1989), was applied in a default prediction setting by Lando, Medhat, Nielsen, and Nielsen (2013). They show that there is significant time-variation in how certain firm-specific variables influence the default probabilities of public US firms.

Statistical models using accounting ratios to estimate default probabilities date back to at least Beaver (1966) and Altman (1968), followed by Ohlsen (1980) and Zmijewski (1984)—we use many of the same accounting ratios as in these studies. Shumway (2001) was among the first to demonstrate the advantages of intensity models with time-varying covariates compared to traditional discriminant analysis, and was also among the first to include equity return as a market-based predictor of default probabilities—we use a similar estimation setup, although we do not have market-based variables for our private firms. Chava and Jarrow (2004) improved the setup of Shumway (2001) using covariates measured at the monthly level and showed the importance of industry effects—our data frequency is also at the monthly level and we correct for industry effects in all our regressions.

Structural models of credit risk, like the models of Black and Scholes (1973), Merton (1974), and Leland (1994), usually assume that a firm defaults when its assets drop to a sufficiently low level relative to its liabilities. The connection between structural models and intensity models was formally established by Duffie and Lando (2001), who showed that when the firm's asset value process is not perfectly observable, a firm's default time has a default intensity that depends on the firm's observable characteristics as well as other covariates. Studies demonstrating the importance of covariates implied from structural models, like distance-to-default or asset volatility, include

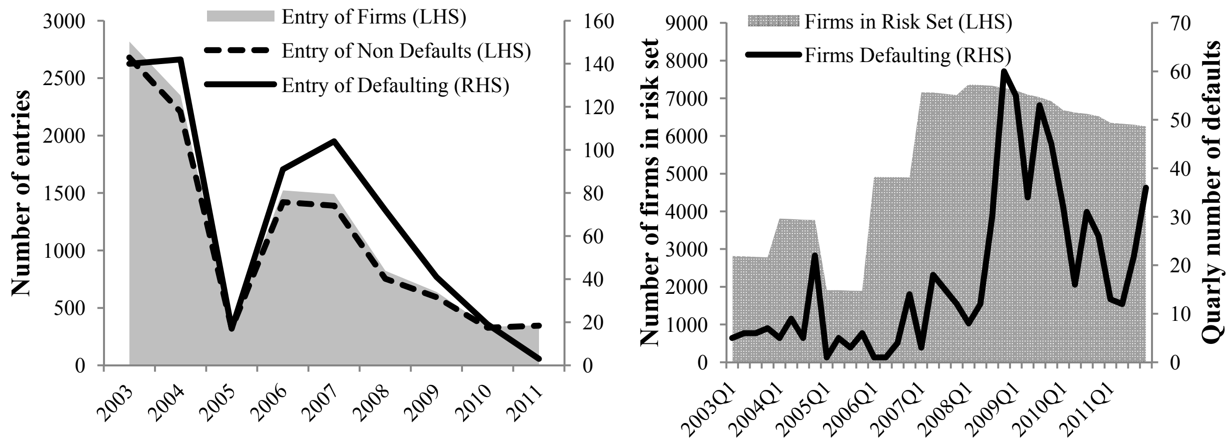

Figure 1. Entry and at-risk pattern in the sample. The left panel shows the yearly number of firms entering the sample (grey mass) along with the yearly number of entries that do not default (black, solid line) or eventually do default (black, dashed line). The right panel shows the quarterly number of firms at risk of defaulting (i.e. in the "risk set"; grey mass) along with the actual number of defaulting firms in each quarter (black, solid line).

Duffie, Saita, and Wang (2007), Bharath and Shumway (2008), Lando and Nielsen (2010), and Chava, Stefanescu, and Turnbull (2011) among many others.

Default studies using data on public firms and demonstrating the importance of employing macroeconomic variables include McDonald and de Gucht (1999), Peseran, Schuermann, Treutler, and Weiner (2006), Duffie et al. (2007), Lando and Nielsen (2010), Figlewski, Frydman, and Liang (2012), among many others. Recent default studies of private firms that also employ macroeconomic variables include Carling, Jacobson, Lindé, and Roszbach (2007), who use Swedish data, and Bonfim (2009), who uses Portuguese data. We employ many of the same macroeconomic variables as in these studies.

4 Data and variables

This section presents our data and the variables that we employ as firm-specific and macroeconomic drivers of default probabilities.

Our raw data comprises 28,395 firms and 114,409 firm-year observations of obligor and loan histories, accounting statements, and default indicators over the period 2003 to 2012. The data is obtained from a large Danish A-IRB (advanced internal ratings-based approach) financial institution. A firm is included in this dataset if it has an engagement over DKK 2 million in at least one of the years underlying the period of analysis. An engagement is defined in terms of loans or granted credit lines. After removing sole proprietorships, government institutions, holding companies without consolidated financial statements, firms that do not have Denmark as their residency,

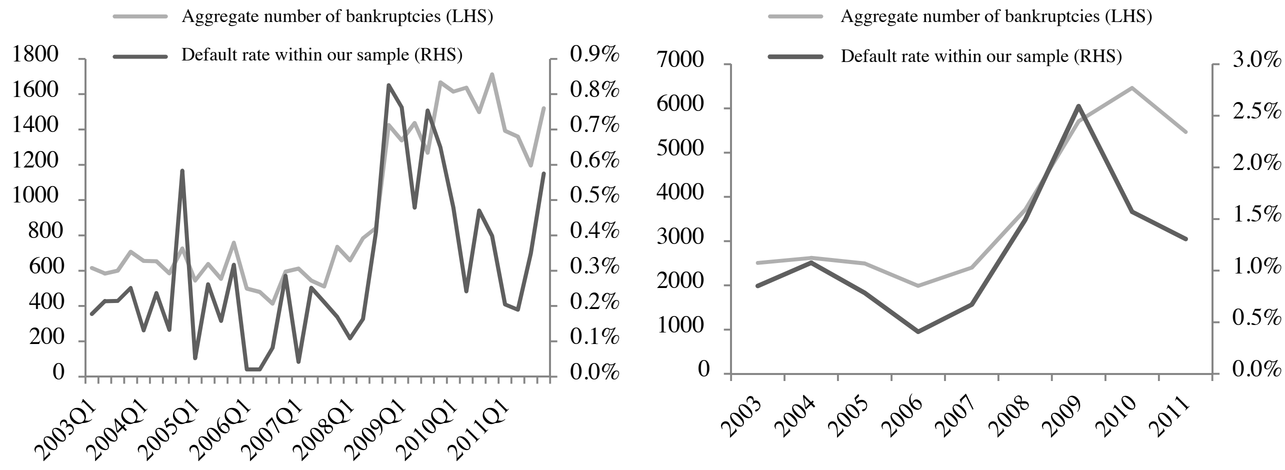

Figure 2. Default rates in the sample and the general Danish economy. The left panel shows the quarterly default rate from the sample along with the corresponding aggregate number of quarterly bankruptcies in the general Danish economy. The right panel shows the yearly default rate from the sample along with the aggregate quarterly number of default in the general Danish economy.

and firms with insufficient balance sheet information, we are left with 10,671 firms and 48,703 firm-year observations. In the cleaned dataset, a total of 633 firms experienced a default event, defined by the Basel II Accord as more than 90 days delinquency. Moreover, 54 of the 633 defaulting firms experience a second default, in the sense that they became delinquent a second time during the sample period. Other default studies have treated a firm that re-emerges from default as a new firm. In accordance with the Basel II Accord's definition of a default event as a period of delinquency, we choose to disregard multiple default events, so that only the initial default counts.

Figure 1 shows the patterns by which firms enter and potentially leave our final sample. The right panel shows the number of firms that enter the sample at each year along with an indication of the number of entries eventually corresponding to defaults and non-defaults. Despite discussions with the financial institution providing the data, the low number of firms entering the sample in 2005 remains a conundrum. It appears, however, that the firms that eventually default do not seem to differ systematically from the non-defaulting firms based on when they enter the sample. The right panel shows the number of firms at risk of defaulting, i.e. firms in the "risk set," at each quarter, along with the quarterly number of defaults. The risk set is seen to contain at least 2,000 firms at each quarter, and the 2008-09 financial crisis is readily visible from the sharp rise in the number of defaults.

In order to incorporate quarterly macroeconomic variables, we re-code the accounting variables for each firm from annual to quarterly observations. This will naturally induce persistence in the accounting variables from quarter to quarter, which we correct for by basing all inference on standard errors clustered at the firm-level. The final dataset thus consists of a total of 192,196 firm-quarter observations.

Figure 2 compares the observed default rate in our sample to the number of registered bankruptcies in Denmark. The comparison is feasible because the total number of firms at risk of default in Denmark is relatively stable over time. We see that, due to the relatively few incidences, the default rate in the sample co-moves nicely with the aggregate level in Denmark. This indicates that our results are not likely to be representative for all Danish SMEs.

4.1 Firm-specific explanatory variables

Table 1 provides an overview of the firm-specific explanatory variables which we employ in our regression analysis. Our firm-specific explanatory variables measure size, age, leverage, profitability, asset liquidity, collateralization, and (book) equity. The table also gives the accounting ratios which we use to proxy for the firm-specific explanatory variables along with the expected sign of each accounting-based variable's effect on default probabilities. All our accounting-based variables have been applied in previous default studies, and the expected signs of their effects are both intuitive and well-discussed in the literature—see, for instance, Ohlsen (1980), Shumway (2001), Duffie et al. (2007), and Lando and Nielsen (2010). We also correct for industry effects as in Chava and Jarrow (2004). The main difference between our list of firm-specific variables and the ones used in default studies of public firms is the lack of market-based measures like stock return and distance-to-default.

We control for industry effects since certain industry characteristics may prescribe a certain leverage structure, for example because of differences in the volatility of cash flows. We use the sector affiliation by Statistics Denmark to identify a firm's primary industry as either "Construction," "Manufacturing," or "Wholesale and Retail," as these have above average default rates, but are at the same time coarse enough to ensure a sufficient number of firms in each sector.

Table 2 presents summary statistics for our firm-specific explanatory variables. An analysis of the variables revealed a few miscodings and extreme values. Due to the anonymized nature of the data, we were not able to check the validity of these data points manually, and we therefore choose to winsorize all the firm-specific variables at the 1st and 99th percentile—a practice also used by Chava and Jarrow (2004), Shumway (2001), and Bonfim (2009), among others. The average firm has DKK 275 million in assets, a total debt to total assets ratio of 68%, and interest payments corresponding to 3% of total assets. Further, the average firm had a relationship with the bank for 23 years and remains in the sample for 7 out of the 9 years.

Due to Danish reporting standards, firms below a certain size may refrain from reporting rev-

Table 1. Firm-specific explanatory variables and corresponding, observable accounting-based variables. The left column shows our list of firm-specific explanatory variables, the center column shows the observable accounting-based variables which we use as proxies, and the right column shows the expected effect of each accounting-based variables on default probabilities. Industry Effects are included in the list for completeness, although we only use this variable as a control (see details in the text).

Explanatory Variable	Proxy	Expected effect on probability of default
Size	Log of book value of assets	Negative
Age	Years active in the bank	Negative
Leverage	Short term debt to total assets	Positive
	Total debt to total assets	Positive
	Interest bearing debt to total assets	Positive
	Interest payments to total assets	Positive
Profitability	Net income to total assets	Negative
	EBIT to total assets	Negative
	EBITDA to total assets	Negative
Liquidity	Current ratio	Negative
	Quick ratio	Negative
Collateralization	Fixed assets to total assets	Negative
	PPE to total assets	Negative
Negative equity	Dummy for negative equity	Positive
Industry Effects	Sector affiliation	Control variable

enue and employee count, and hence these variables are zero (or missing) for a large proportion of firms in the sample. We therefore choose not to use these two variables in our further analysis in order to retain a large sample of smaller firms. In Table 2, firm age is taken to be time since the bank recorded the first interaction with the client, entry year specifies the year at which the firm enters the sample, and duration is the number of years a firm is observed in the sample since its entry year. The negative (book) equity dummy has an unconditional mean of 0.063, meaning that just over 6% of our firm-quarters show a negative value of (book) equity.

4.2 Macroeconomic explanatory variables

Table 3 provides an overview of the macroeconomic explanatory variables which we employ in our regression analysis. Our macroeconomic explanatory variables cover the stock market, interest rates, GDP, credit supply, inflation, industrial production, as well as demand for consumer goods. The table also gives the observable time-series which we use to proxy for the macroeconomic explanatory variables along with the expected sign of each time-serie's effect on default probabil-

Table 2. Descriptive statistics for the firm-specific variables. The table shows descriptive statistics for the firm-specific variables of the cleaned sample, winsorized at the 1*st* and 99th percentile. The total number of observations is 192,196 firm-quaters. Age is time since the bank recorded the first interaction with the client. Entry is the year where the firm entered the sample. Duration is the number of years the firm remains in the sample. All other variable have standard interpretations.

Variable	**Mean**	**Std**	**1%**	**5%**	**25%**	**50%**	**75%**	**95%**	**99%**
Total assets (tDKK)	275.074	1.060.373	811	2.784	9.553	28.222	100.723	1.025.875	8.656.000
Revenue (tDKK)	242.031	885.431	0	0	0	0	70.925	1.097.486	6.733.409
Employees	113	348	0	0	1	18	64	484	2.666
Age (years)	23	20	1	3	9	18	30	71	97,25
Log(total assets) (tDKK)	10,46	1,81	6,70	7,93	9,16	10,25	11,52	13,84	15,97
Short term debt to total assets	0,51	0,28	0,01	0,09	0,30	0,49	0,69	0,96	1,58
Total debt to total assets	0,68	0,28	0,02	0,18	0,53	0,70	0,84	1,04	1,80
Interestbearing debt to total assets	0,39	0,28	0,00	0,00	0,17	0,37	0,56	0,87	1,38
Interest payments to total assets	0,03	0,03	0,00	0,00	0,01	0,02	0,03	0,07	0,17
Current ratio	1,66	2,42	0,03	0,28	0,88	1,17	1,59	3,80	20,21
Quick ratio	1,28	2,37	0,02	0,14	0,49	0,81	1,19	3,15	19,67
Fixed assets to total assets	0,40	0,29	0,00	0,01	0,14	0,36	0,63	0,93	0,99
Tangible Assets to total assets	0,30	0,28	0,00	0,00	0,06	0,23	0,49	0,87	0,97
Net Income to total assets	0,03	0,14	-0,68	-0,18	0,00	0,03	0,09	0,24	0,43
EBIT to total assets	0,06	0,15	-0,59	-0,16	0,00	0,05	0,12	0,29	0,51
EBITDA to total assets	0,10	0,15	-0,51	-0,12	0,02	0,09	0,17	0,34	0,54
Entry Year	2005	1,94	2003	2003	2003	2004	2006	2009	2010
Duration (Years)	7	2	1	2	5	7	9	9	9

ities. The macroeconomic time-series are primarily obtained from Ecowin, with additional data from Statistics Denmark, OECD, and Stoxx.

The inclusion of lagged macroeconomic variables allows us to use these to compute growth rates, differences, or levels. We select the appropriate form by 1) computing the correlation between each form of the macroeconomic variable and the observed default rate, and 2) visually inspecting the relationship of each form with the observed default rate. Note, however, that some pairs of the macroeconomic variables exhibit collinearity—for example, the Danish GDP growth and the European GDP growth rate, as well as the return on the OMX index and the Stoxx index, have pair-wise correlations of 0,92 and 0,77, respectively. The high degree of collinearity should be kept in mind when interpreting the estimated regression coefficients in the following sections.

Table 3. Macroeconomic explanatory variables and corresponding observable time-series. The left column shows our list of macroeconomic explanatory variables, the center column shows the macroeconomic time-series which we use to measure each macroeconomic explanatory variable, and the right column shows the expected effect of each macroeconomic time-series on default probabilities.

Explanatory Variable	Proxy	Expected effect on probability of default
Stock return	Return of OMX index	Negative
Stock volatility	Volatility of OMX index	Unknown
Interest rates	Slope of yield curve	Negative
GDP	Real growth in Danish GDP	Negative
Loan growth	Loan growth to non-financial firms	Positive
Credit availability	Funding costs	Positive
Aggregate defaults	Danish bankruptcies	Positive
Inflation	Headline CPI	Unknown
Demand side effects	Consumer confidence	Negative
	House prices	Negative
Supply side effects	Business indicator, manufacturing	Negative
	Capacity utilization	Negative
International exposure	Exports to Danish GDP	Unknown
	Return of Stoxx50 index	Negative
	EU 27 GDP growth	Negative

5 Estimation methodology

This section present the methodology which we apply to estimate the effects of our explanatory variables on firm-specific default probabilities.

Suppose we have a sample of n levered firms observed over a time-horizon $[0, T]$, where firm i may default at a stochastic time τ_i. At each time t, the firm's financial state is determined by a vector $\mathbf{X}_{it}$ of firm-specific variables and by a vector $\mathbf{Z}_t$ of macroeconomic variables, with values common to all firms in the sample. Default at time t occurs with intensity $\lambda_{it} = \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)$, meaning that λ_{it} is the conditional mean arrival rate of default for firm i, measured in events per time unit. Intuitively, this means that, given survival and the observed covariate histories up to time t, firm i defaults in the short time-interval $[t, t + dt)$ with probability $\lambda_{it}\, dt$.[2] We assume τ_i is doubly-stochastic driven by the combined history of the internal and external covariates (see for instance Duffie et al., 2007).

We devise two different ways of testing whether defaults of smaller firms are more or less

[2]Precisely, a martingale is defined by $1_{(\tau_i \leq t)} - \int_0^t 1_{(\tau_i > s)} \lambda_{is}\, ds$ with respect to the filtration generated by the event $(\tau_i > t)$ and the combined history of the firm-specific and macroeconomic variables up to time t.

cyclical. First, we use the standard multiplicative Cox regression model to estimate the effects of firm-specific and macroeconomic variables on default intensities. Second, we use the more flexible additive Aalen regression model which allows for time-varying effects of firm-specific variables.

5.1 The Cox regression model

In our initial analysis of which accounting ratios and macroeconomic variables that significantly predict defaults, we specify the firm-specific default intensities using the popular "proportional hazards" regression model of Cox (1972). The intensity of firm i at time t is thus modeled as

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right),$$

where Y_{it} is an at-risk-indicator for firm i, taking the value 1 if firm i has not defaulted "just before" time t and 0 otherwise, while $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of regression coefficients. The effect of a one-unit increase in the jth internal covariate at time t is to multiply the intensity by the "relative risk" e^{β_j}. The same interpretation applies to the external covariates. We let the first component of the vector $\mathbf{Z}_t$ be a constant 1. This means that the first component of $\boldsymbol{\gamma}$ is a baseline intensity, corresponding to the (artificial) default intensity of firm i when all observable covariates are identically zero.[3]

Following, for instance, Andersen, Borgan, Gill, and Keiding (1992), and under the standard assumptions that late-entry, temporal withdrawal, right-censoring, and covariate distributions are uninformative on regression coefficients, the (partial) log-likehood for estimation of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ based on a sample of n firms becomes

$$l(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^T \left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) dN_{it} - \int_0^T \sum_{i=1}^{n} Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) dt,$$

where $N_{it} = 1_{(\tau_i \leq t)}$ is the the one-jump default counting process for firm i. We investigate the assumption of independent censoring and entry-pattern in Section 8, and find that our parameter estimates are robust to the exclusion of firm-years that could potentially induce bias.

Estimation, inference, and model selection for the Cox model may then be based on maximum likelihood techniques. Given maximum likelihood estimators (MLEs) of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we can judge the influence of covariates on default intensities by judging the significance of the corresponding regression coefficients, and we can predict firm-specific and aggregate default intensities by plug-

[3]Note that while the usual Cox model includes an (unspecified) time-varying baseline-intensity, thereby making it a semi-parametric survival regression model, we cannot simultaneously identify the vector $\boldsymbol{\gamma}$ of macroeconomic regression coefficients as well as a time-varying baseline-intensity—we therefore restrict to a fully parametric model with a constant baseline intensity.

ging the MLEs back into intensity specification of the Cox model. Model check may be based on the so-called "martingale residual processes,"

$$N_{it} - \int_0^t Y_{is} \exp\left(\widehat{\boldsymbol{\beta}}^\top \mathbf{X}_{is} + \widehat{\boldsymbol{\gamma}}^\top \mathbf{Z}_s\right) ds, \quad i = 1, \ldots, n, \quad t \in [0, T], \tag{2}$$

which, when the model is correct, converge to mean-zero martingales as the sample size increases. Hence, when aggregated over covariate-quantiles or sectors, the grouped residuals processes should not exhibit any systematic trends when plotted as functions of time.

5.2 The Aalen regression model

In addition to the Cox model, we will also employ the additive regression model of Aalen (1980, 1989), which specifies the default intensity of firm i as a linear function of the covariates. This allows us to estimate time-varying effects of firm-specific variables, which is a potentially important component of the cyclicality of default probabilities that is missing from our Cox models. Furthermore, due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques, in contrast to the Cox model.

Our specification of the additive model for the default intensity of firm i is given by

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = \boldsymbol{\beta}(t)^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t,$$

where $\boldsymbol{\beta}(t)$ is a vector of unspecified regression *functions of time*, while $\boldsymbol{\gamma}$ is a vector of (time-constant) regression coefficients.[4]

The linearity of the additive model allows for estimation of both time-varying and constant parameters using ordinary least squares-methods. For the time-varying coefficients, the focus is on the *cumulative regression coefficients*, $B_j(t) = \int_0^t \beta_j(s)\, ds$, which are easy to estimate non-parametrically. Further, formal tests of the significance and time-variation of regression functions is possible through resampling schemes. We refer to Aalen, Borgan, and Gjessing (2008), Martinussen and Scheike (2006), and Lando et al. (2013) for a detailed presentation of estimation and inference procedures.

[4]Note that we cannot identify time-varying effects for the macroeconomic variables, as the macroeconomic variables do not vary across firms at a given point in time.

6 Default prediction for private firms

In this section, we investigate which accounting ratios and macroeconomic variables that significantly predict defaults in our sample. We initially focus on the multiplicative Cox model, before using the addive Aalen model to judge whether allowing for time-varying effects of firm-specific variables changes our conclusions.[5]

First, we show the result from a Cox model using only firm-specific variables. We will see that this model cannot adequately predict the cyclical variation in the aggregate default rate. Second, we add macroeconomic variables to the Cox model and show that this allows the model to much more accurately predict the aggregate default rate over time. However, when judging the different Cox models' ability to correctly rank firms with respect to default likelihood, we will see that macroeconomic variables only marginally improve the ranking based on accounting ratios alone. Hence, to the capture cyclicality of default rates, it is sufficient to focus on macroeconomic variables—however, accounting variables are necessary controls for variations in firm-specific default risk not related to size. Finally, we estimate an additive Aalen using the same variables as in our preferred Cox model, but allowing for time-varying effects of the firm-specific variables. We find significant time-variation in the effects of several firm-specific variables. Furthermore, allowing for this time-variation turns the effects of the macroeconomic variables insignificant. Hence, time-variation in the effects of firm-specific variables is an important alternative way of quantifying the cyclicality of default probabilities.

6.1 Using accounting ratios alone

Initially, we fit a Cox model of firm-by-firm default intensities using only firm-specific variables. We will use this fitted model to examine to what extent macroeconomic variables add additional explanatory power to default prediction.

Table 4 presents estimation results for Cox models using only firm-specific variables. Due to the high degree of correlation among the measurements within the same categories, we perform a stepwise elimination of variables in a given category, removing the least significant variables in

[5]We focus on both the Cox and the Aalen models because they each have their advantages in the tests we consider. The Cox model has the advantages that it automatically produces nonnegative intensities and that its constant regression coefficients allow out of sample prediction. The Aalen model has the advantages of time-varying effects of firm-specific variables and easy comparison of effects across subsamples. Note that while there exist variants of the Cox model that in principle allow for time-varying effects of firm-specific variables, such variants have to apply some degree of smoothing in each estimation iteration, which might blur or distort the effects. Such smoothing is not necessary for estimating time-varying effects in the additive Aalen model. See Martinussen and Scheike (2006), Aalen et al. (2008), and Lando et al. (2013) for more on the differences between (variants of) the Cox model and the Aalen model.

Table 4. Estimation results for Cox models including only accounting ratios. The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity of firms in the sample. All variables are lagged one year to allow for one-year prediction. The full list of firm-specific variables are included in model (1). Models (2) through (4) show the stepwise elimination, keeping only the most significant measure within the groups of (1) leverage, (2) profitability, (3) liquidity, and (4) collateralization. Model (5) (shaded grey) is the preferred specification when only firm-specific variables are used as covariates. Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on standard errors clustered at the firm-level.

	Dependent variable: Default (0/1)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	Variables (all lagged 1 year)	**Coef**	**Coef**	**Coef**	**Coef**	**Coef**
	Intercept	-6,788 ***	-6,494 ***	-6,668 ***	-6,684 ***	-6,729 ***
	Years active in the bank	-0,011 ***	-0,011 ***	-0,011 ***	-0,011 ***	-0,011 ***
	Log of total assets	0,043 *	0,014	0,020	0,019	0,016
(1)	Short term debt to total assets	-0,572 **				
	Total debt to total assets	0,644 **				
	Interest bearing debt to total assets	0,661 ***	1,226 ***	1,262 ***	1,261 ***	1,255 ***
	Interest payments to total assets	7,843 ***				
(2)	Net income to total assets	-0,961 **	-1,790 ***	-2,012 ***	-2,025 ***	-2,015 ***
	EBIT to total assets	1,403 *	1,963 **			
	EBITDA to total assets	-2,615 ***	-2,514 ***			
(3)	Quick Ratio	-0,078	-0,045	-0,057	-0,212 **	-0,202 **
	Current ratio	-0,192	-0,179	-0,161		
(4)	Fixed assets to total assets	-0,455	-0,288	-0,308	-0,280	
	PPE to total assets	0,529 **	0,571 **	0,473 **	0,466 **	0,255
	Negative equity, dummy	0,467 **	0,547 ***	0,555 ***	0,563 ***	0,570 ***
	Construction, dummy	0,926 ***	0,885 ***	0,928 ***	0,921 ***	0,951 ***
	Wholesale and retail trade, dummy	0,214	0,216 *	0,267 **	0,250 *	0,275 **
	Manufacturing, dummy	0,404 ***	0,399 ***	0,420 ***	0,406 ***	0,417 ***
	Number of observations	192.196	192.196	192.196	192.196	192.196
	Number of firms	10.671	10.671	10.671	10.671	10.671
	Number of events	633	633	633	633	633
	Sector effects	YES	YES	YES	YES	YES
	QIC	7.677,2	7.716,4	7.724,8	7.722,8	7.721,9
	QICu	7.669,2	7.710,5	7.719,2	7.717,8	7.717,5

each step. The outcome is that interest bearing debt to total assets, net income to total assets, quick ratio, and tangible assets (PPE) to total assets remain in the model, along with age of banking relationship, log of total assets, and a negative equity dummy.

Interpreting the preferred model (Model 5 in Table 4) the effect of age is negative, implying that the longer a firm has had a relationship with the bank, the less likely it is that the firm will default. The effect of size, as measured by book assets, appears insignificant in the specification.

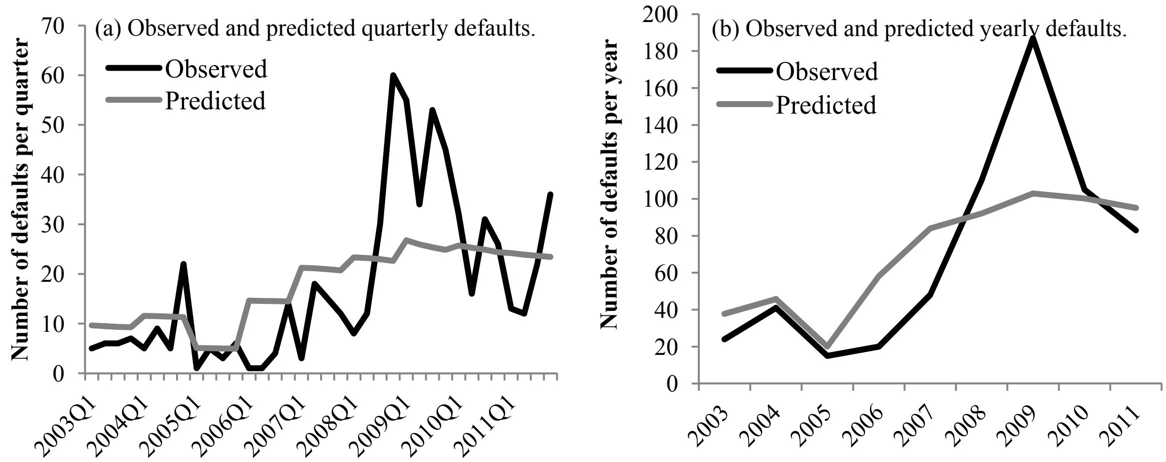

Figure 3. Default prediction based on the preferred Cox model only including accounting ratios. Panel (a) shows the observed number of quarterly defaults in the sample along with the predicted number of defaults based on the preferred Cox model only including accounting ratios (Model (5) in Table 4). Panel (b) is similar, except that the aggregation is done on a yearly basis.

This might potentially be explained by the sample pertaining to only the largest corporate clients, where size is less relevant as an explanation of default. The leverage ratio of interest bearing debt to total assets is, as expected, positively related to default probability. Likewise, past profitability is negatively related to default probability. The quick ratio enters with a significant negative sign confirming the hypothesis that the more liquidity a firm has, the higher its ability to service unexpected cash shortfalls which would otherwise have resulted in a default. Tangible assets, measured as Plant Property and Equipments (PPE) to total assets, does not appear to have a significant effect, confirming the findings of Bonfim (2009) that tangible assets remain insignificant in explaining corporate defaults. The negative (book) equity dummy enters with a positive sign in all specifications, confirming that negative (book) equity is in fact a sign of a firm in trouble and at increased risk of default. The sign of the sectoral dummies are all positive and significant, confirming that these sectors have above average default rates.

Using the results of Table 4, we calculate a predicted quarterly default intensity for each firm in the sample, and then aggregate these to get a predicted aggregate intensity for each quarter. Figure 3 shows the observed number of quarterly defaults in the sample along with the predicted number of defaults based on the preferred Cox model only including accounting ratios. As evident in panel (a), the model based on accounting ratios alone is unable to explain the cyclical nature of the observed defaults. However, acknowledging that the firm-specific data can only change yearly through annual financial statements, it may be more appropriate to aggregate the predicted and observed number of defaults on a yearly basis. This is shown in panel (b), and the conclusion is the same: The model based solely on firm-specific variables is not capable of capturing the cyclical variation in defaults.

Table 5. Estimation results for Cox models with both accounting- and macroeconomic variables. The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity of firms in the sample. All variables are lagged one year to allow for one-year prediction. The full list of firm-specific and macroeconomic variables are included in model (6), and model (7) is the preferred specification after stepwise elimination of variables. Model (8) is the preferred specification in Model (7) excluding the firm-specific variables. Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on standard errors clustered at the firm-level.

Dependent variable: Default (0/1)	**(6)**		**(7)**		**(8)**	
Variables (all lagged 1 year)	**Coef**		**Coef**		**Coef**	
Intercept	-13,341	***	-8,051	***	-7,206	***
Years active in the bank	-0,011	***	-0,011	***		
Log of total assets	0,007		0,007			
Interest bearing debt to total assets	1,232	***	1,231	***		
Net income to total assets	-1,877	***	-1,877	***		
Quick Ratio	-0,200	**	-0,200	**		
PPE to total assets	0,282	*	0,281	*		
Dummy for negative equity	0,592	***	0,591	***		
Aggregate quarterly number of Danish bankruptcies	0,005					
Danish Real GDP growth	-0,027					
Export / GDP	9,633	*				
Inflation, pct point	-0,124					
OMX stock market return	-0,052	***	-0,045	***	-0,038	***
OMX stock market volatility	-0,119	***	-0,101	***	-0,099	***
Difference CIBOR - policy rate, pct. Point	1,837	***	2,006	***	2,117	***
Yield curve slope 10y - 3m, pct. Point	0,986	***	0,588	***	0,634	***
Growth in house prices	-0,114	***	-0,104	***	-0,115	***
Change in consumer confidence Indicator	-0,110	***	-0,099	***	-0,110	***
Change in cyclical indicator, construction	0,002					
Change in capacity utilization in the industrial sector	-0,338	***	-0,300	***	-0,292	***
Loan growth to non-financial institutions	0,023	**				
Stoxx50 stock market return	0,031	***	0,034	***	0,029	***
EU27 Real GDP growth	0,250	***	0,227	***	0,213	***
Number of observations	192.196		192.196		192.196	
Number of firms	10.671		10.671		10.671	
Number of events	633		633		633	
Sector effects	YES		YES		YES	
QIC	7.513		7.513,6		8.265,0	
QICu	7.518		7.508,6		8.264,7	

6.2 Including macroeconomic variables

Given that firm-specific variables are unable to explain the cyclical nature of defaults in our sample, this section incorporates macroeconomic effects in our Cox model. In order to assess if macroeconomic variables add explanatory power beyond what is implied by the firm-specific variables, the preferred model of the firm-specific variables is used as the basis of the covariate specification.

Table 5 presents estimation results for Cox models incorporating macroeconomic variables. The selection procedure has been to perform a stepwise elimination of insignificant variables until only significant macroeconomic variables remain in the model. Model (7) is the preferred model including both firm-specific and microeconomic variables, while Model (8) is this preferred model excluding the firm-specific variables.

The effects of the firm-specific variables remain robust to the inclusion of the macroeconomic variables. In the preferred model (Model (7) of Table 5), the significant macroeconomic variables are as follows: The return of the OMX stock market index, the volatility of OMX index, the difference between CIBOR and the policy rate, slope of the yield curve, change in consumer confidence, change in the capacity utilization, the return of the Stoxx 50 index, and, finally, the European GDP growth rate. On the other hand, the aggregate number of defaults, the Danish real GDP growth, exports as a fraction of GDP, inflation, changes in the cyclical indicator for construction, as well as the loan growth to non-financials are all insignificant.

When interpreting the coefficients of the macroeconomic variables in multivariate intensity regression models, one should bear in mind that it would be unrealistic to obtain a complete *ceteris paribus* effect of one macroeconomic variable, as this variable cannot be viewed in isolation from other macroeconomic variables. While not done here, an appropriate interpretation would involve testing the model from the perspective of internally consistent scenarios of macroeconomic variables. For instance, a further analysis shows that the volatility of the stock market, the slope of the yield curve, the return of the Stoxx 50 index, and the European GDP growth rate appear with an opposite sign in the preferred model compared to a model where they enter separately.

Nonetheless, a positive return of the OMX stock market would, controlling for other macroeconomic effects, imply a lower number of default occurrences one year after. An increased spread between CIBOR and the policy rate would be associated with an increased number of default occurrences, thereby supporting the notion that the higher funding costs of the banks would generally be passed through to clients. Both growth in house prices and changes in the consumer confidence index tend to be negatively linked to defaults, illustrating the importance of demand side effects. Capacity utilization is also negatively associated with default occurrences, meriting the interpretation that higher level of idle capacity could result in price competition that would ultimately lead a number of firms to default.

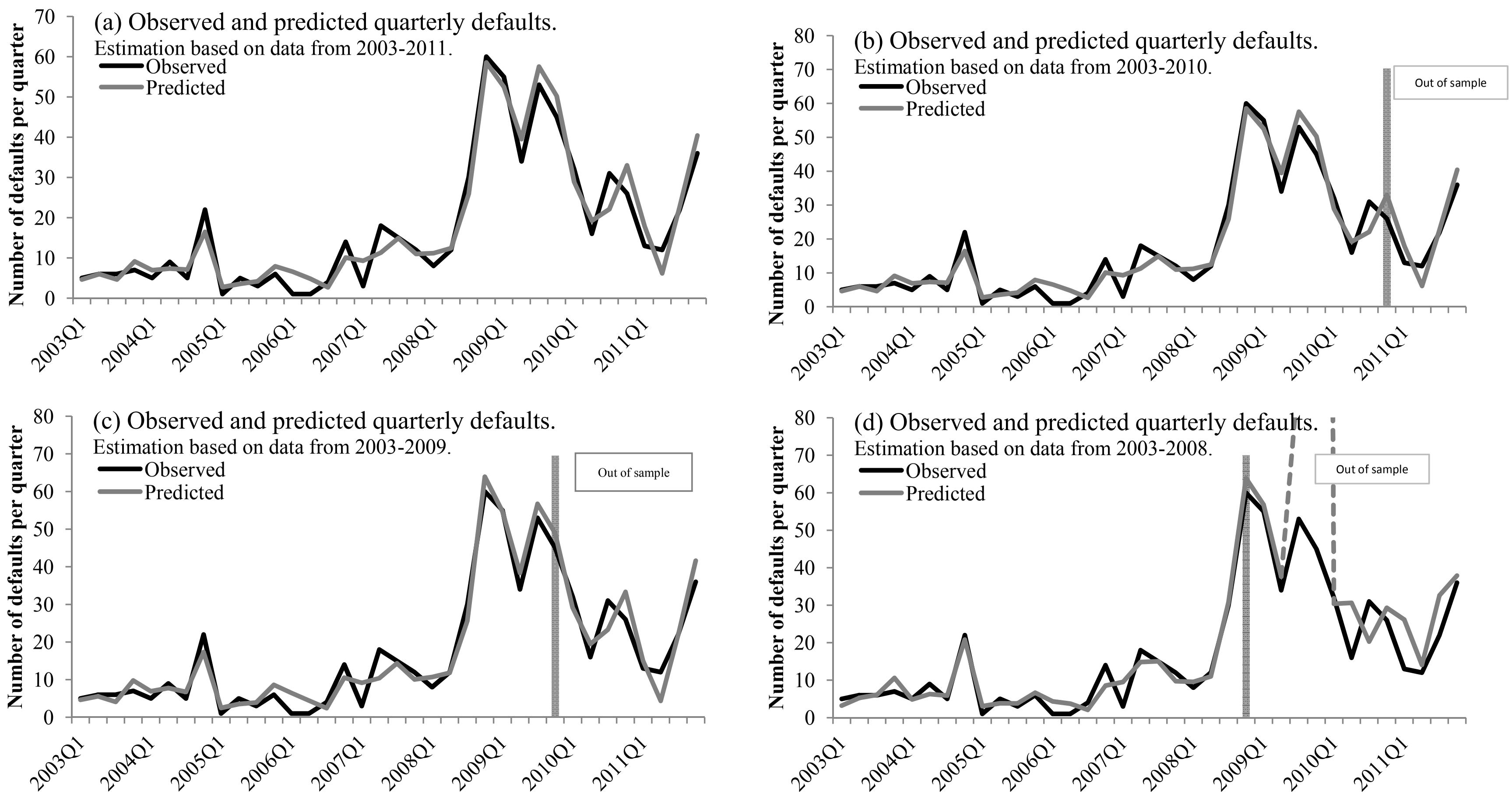

Figure 4. Default prediction based on the preferred Cox model with both accounting ratios and macroeconomic variables. All panels show the observed quarterly number of defaults in the sample along with a predicted number of defaults based on the preferred Cox model with both firm-specific and macroeconomic covariates (Model (7) in Table 5). In Panel (a), the model is estimated using the full sample from 2003 to 2011. In Panels (b), (c), and (d), the model is estimated on shorter subsamples, allowing in each case out-of-sample prediction on the remaining part of the full sample.

Figure 4 illustrates the relationship between the observed and predicted number of defaults taking into account both the firm-specific variables and the macroeconomic variables in Model (7) of Table 5. Adding macroeconomic variables as explanatory factors improves the model's ability to predict the cyclical variation in quarterly default occurrences. To rule out that the good fit of the preferred model is merely a result of over-fitting, the out of sample prediction obtained from estimating the same model on only part of the data supports the chosen model. Panel (b) and (c) of the figure estimates the model on the sample excluding observations from 2010 and both 2010 and 2011 respectively. The obtained coefficients from the models estimated on the reduced samples are then used to estimate the aggregate intensities for all 36 quarters, thereby generating out of sample predictions. Hence, Panel (b) of the figure shows one and Panel (c) shows two years of out of sample prediction. The out of sample prediction based on the reduced sample estimation adequately captures both the level and cyclical variation in default rates.

Panel (d) of Figure 4 shows prediction based on excluding the years 2009, 2010 and 2011 from the estimation. For the out of sample prediction in Panel (d), large deviations occur in 2009 (which pertains to 2008 covariates observations because of the one year lag). However, it should be emphasized that the latter model has been fitted to a period of economic expansion, and therefore it is of little surprise that the model cannot be used to predict future defaults in a period of economic contraction. This finding also highlights the importance of estimating default predicting occurrences on a full business cycle.

6.3 Ranking firms with respect to default likelihood

The out-of-sample estimation results presented in Figure 4 showed that the preferred Cox model including both firm-specific and macroeconomic variables adequately captures the level and cyclicality of defaults. A common way of illustrating the predictive power of different models based on the same data is to plot the Receiver Operating Curve (ROC), as shown in Figure 5. The curve illustrates the percentage of defaults that are correctly classified as defaults on the vertical axis against the percentage of non-defaults that are mistakenly classified as defaults on the horizontal axis for all possible cutoff points. The area under the curve (AUC) is then used as measure of the model goodness of fit where a value of 1.0 implies a model with perfect discriminatory ability and a value of 0.5 is a completely random model.

In terms of discriminatory power, the addition of macroeconomic variables does not improve the model's ability to effectively determine which firms eventually default beyond what is implied by the accounting ratios. From the ROC curves, we see that the model with both macroeconomic and firm-specific variables is only marginally better in correctly determining defaults compared to

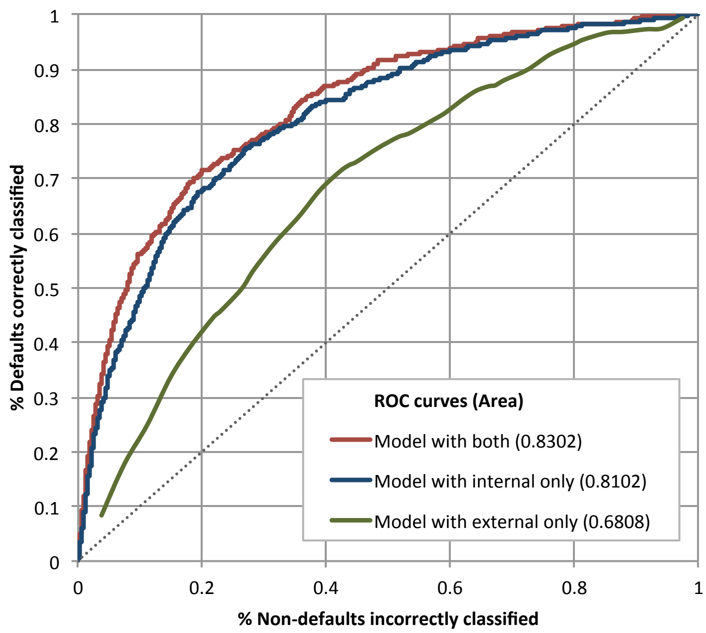

Figure 5. Comparison of firm-ranking accuracy for different covariate-specifications. The figure shows receiver operating characteristic (ROC) curves for Cox models with different covariate-specifications fitted to the sample. Each curve illustrates the model's ability to correctly discriminate between defaults and non-defaults, and plots the percentage correctly classified defaults (true positives) against the percentage incorrectly classified non-defaults (false positives) at all possible cut-off points of default intensity. The area under each curve serves as a goodness-of-fit measure, where a value of 1 means a model with perfect discriminatory ability, while a value of 0.5 means a model that discriminates based on a random guess.

the model with just firm-specific variables. This confirms that it is the firm-specific characteristics that provide the ordinal ranking of firms, and therefore also ultimately determine *which* firms that actually default. Including the macroeconomic factors only improves the model's ability to capture the cyclicality in the aggregate default rate, which is related to *when* defaults occur.

6.4 Allowing for time-varying effects of firm-specific variables

We now use the additive Aalen model to judge whether allowing for time-varying effects of firm-specific variables changes our conclusions regarding which firm-specific and macroeconomic variables significantly affect default intensities.

We initially fit an additive model for our entire sample of firms, including the same covariate specification as our final Cox model (Model (7) of Table 5), and with time-varying coefficients for the firm-specific covariate. The null hypothesis of a time-constant marginal effect is rejected at standard significance levels for the following four firm-specific variables: Firm-size (log of total assets), interest bearing debt to total assets, quick ratio, and the construction sector indicator. The time-varying marginal effects of these variables are shown as cumulative regression coefficients with 95% pointwise confidence bands in Figure 6. When interpreting these effects, one should

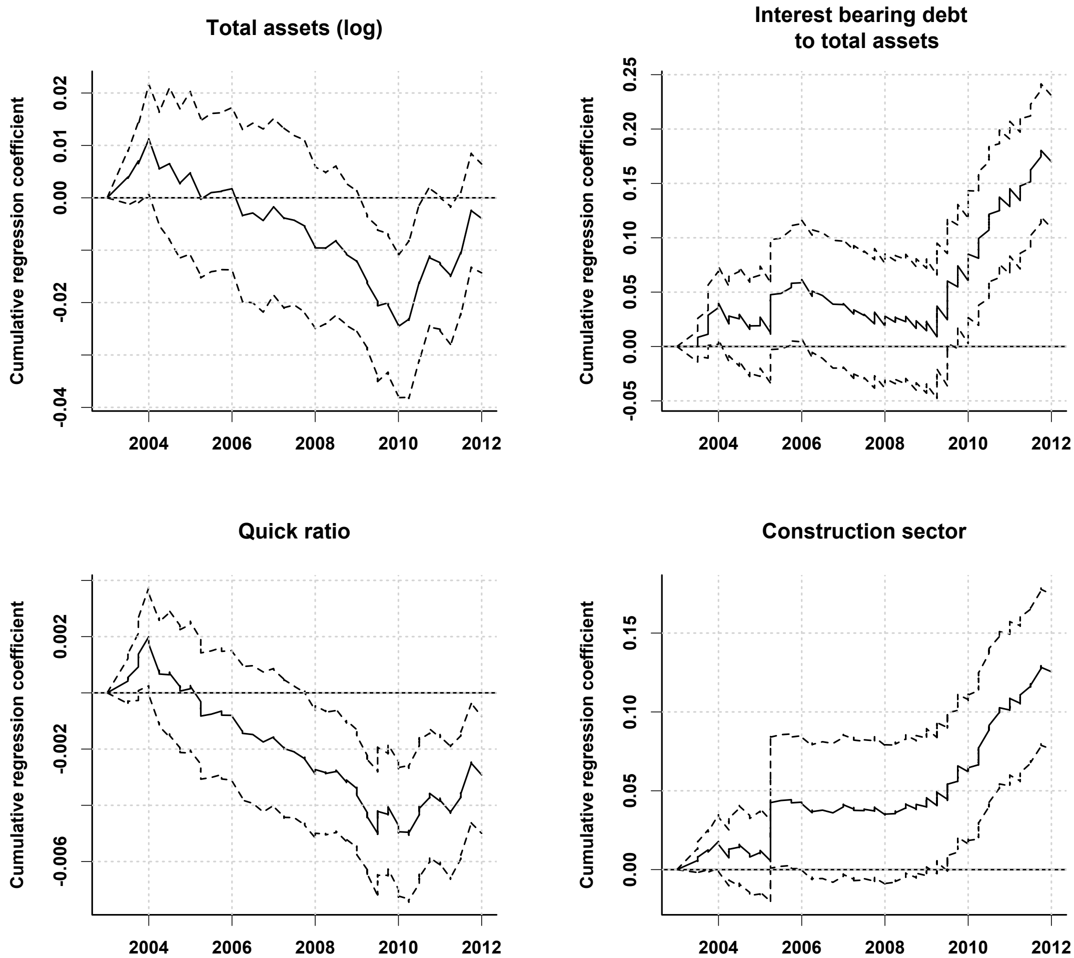

Figure 6. Cumulative regression coefficients from Aalen analysis of full sample. The panels show (cumulative) estimation results for the significantly time-varying firm-specific covariates from an analysis based on the additive Aalen model including the same covariate specification as our final Cox model (Model (7) of Table 5). All variables are lagged one year to allow for one-year prediction. The dotted lines are asymptotic 95% pointwise confidence bands.

focus on the *slopes* of the cumulative coefficients, which estimate the regression coefficients themselves. The plots show that smaller firms, firms with higher interest bearing debt to total assets, firms with lower quick ratio, and firms within the construction sector have higher default probabilities, and particularly so during the financial crisis of 2008-2010. Importantly, firm-size has a mostly negligible marginal effect for most of our sample period (consistent with the results from our analysis based on the Cox model), except for around the financial crisis of 08-10, where the effect becomes significantly negative. Hence, contrary to the assumption in the Basel Accords, we find that larger firms were safer, not riskier, during the most recent recession.

The estimation results for the time-constant regression coefficients from the additive model fitted to the entire sample are given in Table 6. We see that all coefficients corresponding to firm-specific variables have the same sign as in our final Cox model (Model (7) of Table 5) and roughly

Table 6. Constant regression coefficients from Aalen analysis of full sample. The table shows the estimation results for the time-constant regression coefficients for the firm-specific and macroeconomic variables from an analysis based on the additive Aalen model including the same covariate specification as our final Cox model (Model (7) of Table 5). All variables are lagged one year to allow for one-year prediction. Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on robust standard errors.

Variables (all lagged 1 year)	Coef		Se
Intercept	0.0025		0.0407
Years active in the bank	-2.05×10^{-5}	***	5.34×10^{-6}
Net income to total assets	−0.0140	***	0.0017
PPE to total assets	−0.0014	**	0.0006
Dummy for negative equity	0.0078	***	0.0012
Wholesale and retail trade, dummy	0.0005		0.0004
Manufacturing, dummy	0.0012	**	0.0004
OMX stock market return	−0.0036	**	0.0016
OMX stock market volatility	−0.0064		0.0080
Difference CIBOR - policy rate, pct. point	0.0059		0.0510
Yield curve slope 10y - 3m, pct. point	0.0110		0.0111
Growth in house prices	−0.0006		0.0024
Change in consumer confidence indicator	−0.0105	**	0.0047
Change in capacity utilization	−0.0077		0.0085
Stoxx50 stock market return	0.0042	***	0.0012
EU27 Real GDP growth	0.0050		0.0077

the same significance level. The macroeconomic variables, however, appear to have lost much of their importance compared to the analysis based on the Cox models. In the additive setting, only OMX stock market return, change in consumer confidence indicator, and Stoxx50 stock market return have significant marginal effects. The latter is of the reversed sign compared to intuition, but is nonetheless consistent with results for public-firms found by Duffie et al. (2007); Duffie, Eckner, Horel, and Saita (2009), Lando and Nielsen (2010), and Figlewski et al. (2012), amongst others.[6] Importantly, we see that allowing for time-varying effects for firm-specific variables leaves little room for additional explanatory power for macroeconomic variables. This is a consequence of the fact that the time-variation in the effects shown in Figure 6 is highly correlated with indicators of macroeconomic conditions, leaving little room for additional explanatory power of macroeconomic variables.

In sum, the analysis based on the additive model suggests that allowing for time-variation in the effects of firm-specific variables is, indeed, an important component of the cyclicality of default probabilities that is missing from our preferred Cox model.

[6] Giesecke, Lando, and Medhat (2013) show that univariately significant but multivariately insignificant or even reversed effects may be observed for macroeconomic variables if these have indirect effects mediated through other covariates included in the models. This is in particular the case for stock market returns.

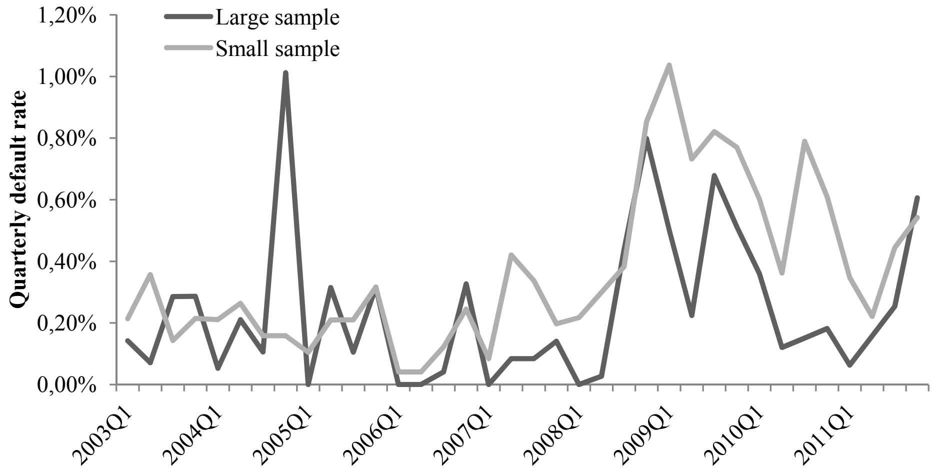

Figure 7. Aggregate quarterly default rate in each of the subsamples of small and large firms. This figure shows the aggregate quarterly default rate in each of the subsamples of small and large firms. Based on the year a firm enters the sample, it is classified as "large" if its first-year asset level is above the median asset level that year. Similarly, a firm is classified as "small" if its asset value at the time of entry is below the median asset level that year.

7 The macroeconomy's impact on small and large firms' default risk

The results of the previous section show that our final Cox model specification including both firm-specific and macroeconomic variables is able to both accurately rank firms and predict the aggregate default rate over time. In this section, we first use this Cox model to investigate whether our data supports the assumption underlying the Basel II and III Accords, that the default probabilities of smaller firms are less impacted by the macroeconomy. We then redo our analysis using the Aalen model, in order to judge whether allowing for time-varying effects of firm-specific variables changes our conclusions.

Our tests examine the extent to which the default probabilities of small and large firms are impacted differently by macroeconomic variables. To do so, we choose to we split the sample into two subsamples, corresponding to "small" and "large" firms. Based on the year that a paticular firm enters the sample, it is classified as "small" ("large") if its first-year asset level is below (above) the median asset level that year. We use assets instead of revenue as our measure of size because, as discussed in Section 2.1, Danish reporting standards allow firms below a certain size to refrain from reporting revenue. Note, however, that our use of asset level as the measure of size is in line with Article 274 of The Basel Committee on Banking Supervision's (2006) report.[7]

Our choice to divide the sample based on the median asset level at a firm's entry year is done

[7]The article's wording is as follows: "Subject to national discretion, supervisors may allow banks, as a failsafe, to substitute total assets of the consolidated group for total sales in calculating the SME threshold and the firm-size adjustment. However, total assets should be used only when total sales are not a meaningful indicator of firm size."

to ensure approximately equal sample sizes with a sufficient amount of default events in each subsample, as well as to allow for the classification of firms as "small" or "large" in a predicable manner. However, we later show (in the section on robustness and model checks) that our results are robust to a division into four subsamples based on asset value quartiles for the entry year.

Figure 7 shows the aggregate quarterly default rate in each of the subsamples of small and large firms. The two default rates are seen to exhibit a high degree of comovement throughout our entire sample period. In the beginning of our sample period, the two default rates are seen to be virtually indistinguishable except for the spike in the default of large firms around 2005. However, near the end of our sample period, the default rate of small firms tends to be systematically higher than the default rate of large firms, indicating that small firms were hit the hardest by the financial crises of 2008-10.

7.1 Macro-sensitivity analysis based on the Cox model

Table 7 shows estimation results for the final Cox model specification including both firm-specific and macroeconomic variables fitted to each of the two subsamples. With the exception of OMX stock market volatility, the magnitude of all macroeconomic effects are larger for large firms compared to small firms. The difference between the coefficients of macroeconomic factors for the two subsamples is significant for the slope of the yield curve, growth in house prices, and European GDP growth, and all are larger in magnitude in the subsample of large firms.[8] Hence, if we suppose there exists a large and a small firm whose only difference is their size (which is in principle possible since all accounting ratios in our models are relative to total assets), the apparent interpretation of these results is that the small firm's default intensity is less exposed to macroeconomic fluctuations.

On the other hand, the estimation results for the two subsamples also show substantial differences with regards to the coefficients of the firm-specific variables: Firm size appears with a significant positive coefficient for small firms, but an insignificant (yet negative) coefficient for large firms; neither quick ratio nor the ratio of tangible assets to total assets have significant effects for small firms, whereas they have significant effects for large firms; and, finally, negative equity has a significant effect for small firms, but not for large.

[8]In an unreported analysis, we find that the importance of the European GDP growth for the large firms is merited by the tendency of large firms in the sample to engage more actively in exports.

Table 7. Estimation results for Cox models fitted to the subsamples of small and large firms. The table shows parameter estimates, standard errors, model summary statistics, and comparison criteria for Cox models of the quarterly default intensity for small and large firms in the sample. The covariate list corresponds to the preferred model including both firm-specific and macroeconomic variables (Model (7) in Table 5). All variables are lagged one year to allow for one-year prediction. In each year, an entering firm is deemed "small" if its book value of assets is under the median assets level for that particular year—otherwise, it is deemed "large." Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on standard errors clustered at the firm level. "Same sign" indicates whether or not the estimated coefficients are of the same sign for small and large firms; "Magnitude" indicates which type of firm has the largest estimated coefficient; and "Sig. Diff." indicates whether or not the coefficients for small and large firms are significantly different at the 5% level using a Welch t-test. Finally, "PEA" indicates whether the partial effect at the average is largest for the large firms or the small firms, while "APE" indicates whether the average partial effect is largest for the large firms or the small firms.

	Small firms			Large firms			Comparison				
Variables (all lagged 1 year)	Coef		Se	Coef		Se	Same sign	Magnitude	Sig. diff.	PEA	APE
Intercept	−8.290	***	0.676	−9.057	***	0.765	Yes	Large	No	Small	Small
Years active in the bank	−0.008	**	0.004	−0.012	***	0.004	Yes	Large	No	Small	Small
Log of total assets	0.122	**	0.056	−0.006		0.043	No	NA	No	NA	NA
Interest bearing debt to total assets	1.048	***	0.224	1.628	***	0.351	Yes	Large	No	Small	Small
Net income to total assets	−1.788	***	0.302	−2.432	***	0.422	Yes	Large	No	Small	Small
Quick ratio	−0.112		0.074	−0.462	***	0.137	Yes	Large	Yes	Large	Large
PPE to total assets	0.040		0.197	0.508	**	0.255	Yes	Large	No	Large	Large
Dummy for negative equity	0.782	***	0.199	0.269		0.291	Yes	Small	No	Small	Small
OMX stock market return	−0.045	***	0.015	−0.045	**	0.018	Yes	Large	No	Small	Small
OMX stock market volatility	−0.102	***	0.038	−0.100	**	0.049	Yes	Small	No	Small	Small
Difference CIBOR - policy rate, pct. point	1.655	***	0.447	2.463	***	0.553	Yes	Large	No	Small	Small
Yield curve slope 10y - 3m, pct. point	0.372	***	0.130	0.869	***	0.154	Yes	Large	Yes	Small	Large
Growth in house prices	−0.040		0.032	−0.200	***	0.047	Yes	Large	Yes	Large	Large
Change in consumer confidence indicator	−0.083	***	0.026	−0.122	***	0.036	Yes	Large	No	Small	Small
Change in capacity utilization	−0.194	**	0.086	−0.420	***	0.106	Yes	Large	No	Small	Large
Stoxx50 stock market return	0.027	***	0.009	0.043	***	0.013	Yes	Large	No	Small	Small
EU27 Real GDP growth	0.101	*	0.055	0.410	***	0.070	Yes	Large	Yes	Large	Large
PEA scaling factor	0.0023			0.0009							
APE scaling factor	0.0041			0.0025							
Number of observations	91,182			101,014							
Number of firms	5,333			5,338							
Number of defaults	376			257							
Sector effects	Yes			Yes							

While the results for the macroeconomic variables may corroborate the lower asset correlation adopted in Basel II and III for SMEs, the direct comparison of coefficients in the two subsamples ignores the fact that a covariate's marginal effect in a non-linear model, like a Cox regression, depends on the values of all the other covariates. This implies that comparing the coefficients for the macroeconomic variables in the two subsamples is potentially problematic, because such a comparison fails to take into account that the firm-specific characteristic for the small and large firms are generally different and have different effects on default intensity.

To elaborate, note that the marginal effect in a Cox regression of a change in the jth macroeconomic variable on the default intensity of firm i is given by

$$\frac{\partial\lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial Z_{jt}} = \gamma_j \, Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) = \gamma_j \, \lambda(\mathbf{X}_{it}, \mathbf{Z}_t),$$

which depends on all the characteristics of firm i through $\mathbf{X}_{it}$, as well as all other macroeconomic variables through the dependence on $\mathbf{Z}_t$.

A somewhat crude way to facilitate comparison between subsamples in non-linear models, like Cox regression, is to compute the marginal effect of a covariate at "average levels" in each of the subsamples. This gives rise to the *partial effect at the average* (PEA) and the *average partial effect* (APE)—see, for instance, Wooldridge (2009, p. 582-83). In the setting of an intensity model, the PEA plugs a subsample's average covariate values into the subsample's estimated intensity, while the APE takes the average across the estimated intensity values for each subsample. Due to the non-linearity of the intensity, Jensen's inequality implies that the two ways of averaging will generally produce different results.

In our analysis, the PEA is a measure of a covariate's marginal effect for the "average firm" and at "average macroeconomic levels" in each of the two subsamples of small and large firms. We thus compute the PEA for the jth macroeconomic variable in subsample k as

$$\text{PEA}_{kj} = \gamma_{kj} \underbrace{\exp\left(\widehat{\boldsymbol{\beta}}_k \overline{\mathbf{X}}_k + \widehat{\boldsymbol{\gamma}}_k^\top \overline{\mathbf{Z}}_k\right)}_{= s_k^{\text{PEA}}},$$

where $k \in \{\text{small}, \text{large}\}$, $\overline{\mathbf{X}}_k$ is the average firm-specific covariate vector for firms in subsample k, $\overline{\mathbf{Z}}_k$ is the average macroeconomic covariate vector in subsample k, $\widehat{\boldsymbol{\beta}}_k$ and $\widehat{\boldsymbol{\gamma}}_k$ are the estimated regression coefficients in subsample k, while s_k^{PEA} is a subsample-specific scaling factor for each PEA. On the other hand, the APE is a measure of a covariate's marginal effect at the "average intensity level" across firms and time in each of the two subsamples. The APE for the jth macroeconomic variable in subsample k is thus computed as

$$\mathrm{APE}_{kj} = \gamma_{kj} \underbrace{\frac{1}{T}\int_0^T \frac{1}{|k(t)|}\sum_{i\in k(t)} \exp\left(\widehat{\boldsymbol{\beta}}_k^\top \mathbf{X}_{it} + \widehat{\boldsymbol{\gamma}}_k^\top \mathbf{Z}_t\right) dt}_{=s_k^{\mathrm{APE}}},$$

where $k(t)$ denotes the firms belonging to subsample $k \in \{\text{small}, \text{large}\}$ at time t, and s_k^{APE} is again a subsample-specific scaling factor for each APE.

The two right-most columns of Table 7 show the PEAs and APEs for each covariate in each of the two subsamples. Focusing on the effects of the macroeconomic variables, the PEA suggests that most macro effects are, on average, stronger in the sample of small firms, while the APEs suggest that it is entirely dependent on the macro variable at hand whether its average effect is stronger for small or large firms.

In sum, while the direct comparison of regression coefficients indicates that smaller firms are less cyclical than larger firms, the more refined analysis based on the PEA and APE, which takes the non-linearity of the Cox model into account, indicates that small firms may "on average" be as cyclical, or perhaps even more cyclical, than large firms. To rule out that this conclusion somehow hinges on the Cox specification, we now investigate macro-sensitivity of default intensities using an additive intensity regression model.

7.2 Macro-sensitivity analysis based on the additive Aalen model

Due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques.

Recall that our specification of the additive model for the default intensity of firm i is given by

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = \boldsymbol{\beta}(t)^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t.$$

It follows that the marginal effect of the jth firm-specific variable is given by

$$\frac{\partial \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial X_{ij,t}} = \beta_j(t),$$

which is time-varying but otherwise independent of other variables and their effects. Similarly, the marginal effect of the jth macroeconomic variable is

$$\frac{\partial \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial Z_{j,t}} = \gamma_j,$$

Table 8. Constant regression coefficients from Aalen analysis of small and large firms. The table shows the estimation results for the time-constant regression coefficients for the firm-specific and macroeconomic variables from an analysis based on the additive Aalen model for the two subsamples of small and large firms. All variables are lagged one year to allow for one-year prediction. Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on robust standard errors.

	Small firms			Large firms		
Variables (all lagged 1 year)	Coef		Se	Coef		Se
Intercept	0.0597		0.1883	0.0600		0.1918
Years active in the bank	-1.83×10^{-5}	*	1.14×10^{-5}	-1.70×10^{-5}		1.18×10^{-5}
Net income to total assets	−0.0160	***	0.0025	−0.0168	***	0.0025
PPE to total assets	−0.0021	**	0.0010	−0.0021	**	0.0010
Dummy for negative equity	0.0082	***	0.0016	0.0085	***	0.0016
Wholesale and retail trade, dummy	0.0003		0.0006	0.0003		0.0006
Manufacturing, dummy	0.0011		0.0007	0.0011		0.0007
OMX stock market return	-0.0084		0.0066	−0.0135	*	0.0069
OMX stock market volatility	0.0192		0.0324	0.0089		0.0317
Difference CIBOR - policy rate, pct. point	−0.3409		0.2317	−0.3048		0.2416
Yield curve slope 10y - 3m, pct. point	0.0003		0.0524	0.0134		0.0560
Growth in house prices	0.0136		0.0105	0.0098		0.0115
Change in consumer confidence indicator	−0.0324	*	0.0191	−0.0283	*	0.0196
Change in capacity utilization	−0.0625		0.0402	−0.0713	*	0.0404
Stoxx50 stock market return	0.0111	**	0.0050	0.0146	**	0.0054
EU27 Real GDP growth	0.0049		0.0351	0.0131		0.0370

which is time-constant, but again independent of other variables and their effects. Hence, in contrast to the multiplicative effects in the Cox model, the linear effects in the additive model are easy to interpret and compare across subsamples without the need for averaging techniques. We will therefore use the additive model to check the assumption that the default probabilities of smaller firms are less sensitive to macroeconomic cycles.

We now re-do our macro-sensitivity analysis by fitting an additive model including the preferred firm-specific and macroeconomic variables to each of the two subsamples of small and large firms. The definition of small and large firms is the same as for the subsample analysis based on the Cox model in Section 4.1. Even though the macroeconomic variables did not have much explanatory power in the additive model fitted to the entire sample, one could still imagine that some macroeconomic variables had significant additive effects in the subsample of large firms, but not in the subsample of small firms. This would be evidence that macro-dependence differs for small and large firms, as is the working assumption in the Basel Accords.

Estimating the additive model in each of the two subsamples did not change the conclusions regarding which firm-specific variables have significantly time-varying effects. Furthermore, the time-varying effects within the two subsamples were virtually indistinguishable from their counterparts for the full sample shown in Figure 6, and are therefore omitted here. In particular, in both subsamples, firm-size has a mostly negligible marginal effect except during the financial crisis of 2008-2010, where larger firms (within each subsample) have significantly lower default probabilities.

With regards to the time-constant effects, Table 8 shows the estimation results for the time-constant regression coefficients in the subsamples of small and large firms. As previously mentioned, the linearity of the additive model allows us to directly compare coefficients across subsamples without the need for averaging techniques. Doing so, we see no difference in sign and virtually no difference in significance or magnitude in the two subsamples, and this is regardless of whether we consider the firm-specific variables or the macroeconomic variables.

In conclusion, our subsample analysis based on the additive model shows no evidence that the default probabilities of smaller firms are less sensitive to macroeconomic cycles. Our findings based on the additive model thus generally reiterate the ones based on the Cox model, but they do so in a framework that allows for cyclicality through the effects of firm-specific variables as well as direct comparison of effects across subsamples.

8 Robustness and model check

In this section, we perform robustness checks and examine model fits. Because our main analysis is based on the Cox model, we will focus on this model. However, the results presented here are in general very similar when done based on the additive model.

8.1 Alternative criteria for defining small and large firms

In our analysis of small and large firms in Table 7, the criterion defining the two subsamples of small and large firms was the median asset level of the year where a particular firm enters the sample. We now investigate alternative criteria for defining the subsamples of small and large firms and how they affect the effects of our variables within the two subsamples.

Table 9 shows estimation results for Cox models within subsamples of small and large firms, where the subsamples are using defined different, time-varying size-criteria. All the models use the covariate list corresponding to our preferred Cox model including both firm-specific and macroeconomic variables, i.e. Model (7) in Table 5.

Table 9. Estimation results for Cox models fitted to subsamples of small and large firms using different, time-varying size-criteria. The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity for small and large firms in the sample. The covariate list corresponds to the preferred model including both firm-specific and macroeconomic variables (Model (7) in Table 5). All variables are lagged one year to allow for one-year prediction. The firm-specific variables are kept in all models to focus on the added explanatory effect of macroeconomic variables. In models (A-1) and (A-2), a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 25th percentile of the full sample's book asset values (which is DKK 9,553,000). In models (A-3) and (A-4), the cut-off is instead the 50th percentile of the full sample's book asset values (which is DKK 28,222,000). Finally, in models (A-5) and (A-6), the cutoff is the 75th percentile of the full sample's book asset values (which is DKK 100,723,000). This classification is time-varying in the sense that the same firm can be classified as small in some years but large in other years. Significance of parameters is indicated at the 10% (*), 5% (**), and 1% (***) levels. Parameter significance is based on standard errors clustered at the firm level.

Dependent variable: Default (0/1)	Cutoff: 25th percentile						50th percentile						75th percentile					
Variables (all lagged 1 year)	(A.1)		Se	(A.2)		Se	(A.3)		Se	(A.4)		Se	(A.5)		Se	(A.6)		Se
Intercept	-9,00	***	1,08	-8,62	***	0,63	-8,32	***	0,71	-8,58	***	0,81	-8,14	***	0,56	-5,80	***	0,56
Years active in the bank	-0,01		0,01	-0,01	***	0,00	-0,01	**	0,00	-0,01	**	0,00	-0,01	***	0,00	0,00		0,00
Log of total assets	0,28	**	0,11	-0,03		0,04	0,12	**	0,06	-0,06		0,05	0,06		0,04	-0,29	***	0,04
Interest bearing debt to total assets	0,87	***	0,26	1,67	***	0,27	0,89	***	0,22	2,15	***	0,36	1,09	***	0,21	2,25	***	0,21
Net income to total assets	-1,65	***	0,34	-2,54	***	0,34	-2,02	***	0,28	-1,98	***	0,51	-1,96	***	0,26	-1,96	**	0,26
Quick Ratio	-0,09		0,07	-0,36	**	0,15	-0,12	*	0,07	-0,53	***	0,14	-0,16	*	0,09	-0,55	**	0,09
PPE to total assets	-0,01		0,26	0,29		0,20	0,23		0,19	0,11		0,25	0,31	*	0,17	-0,18		0,17
Dummy for negative equity	0,93	***	0,23	0,29		0,22	0,79	***	0,19	0,19		0,31	0,67	***	0,17	0,24		0,17
Construction	0,41	**	0,20	1,34	***	0,19	0,65	***	0,17	1,56	***	0,24	0,79	***	0,15	1,81	***	0,15
Wholesale and retail trade	0,01		0,19	0,42	**	0,17	0,19		0,15	0,34		0,21	0,24	*	0,14	0,30		0,14
Manufacturing	0,17		0,23	0,61	***	0,17	0,29	*	0,18	0,65	***	0,21	0,42	***	0,15	0,50		0,15
OMX stock market return	-0,05	***	0,02	-0,04	***	0,02	-0,04	***	0,01	-0,05	**	0,02	-0,05	***	0,01	-0,04		0,01
OMX stock market volatility	-0,12	**	0,05	-0,09	**	0,04	-0,10	***	0,04	-0,11	**	0,05	-0,11	***	0,03	-0,09		0,03
Difference CIBOR - policy rate, pct. point	1,45	***	0,54	2,46	***	0,44	1,79	***	0,44	2,32	***	0,55	1,94	***	0,38	2,10	***	0,38
Yield curve slope 10y - 3m, pct. point	0,21		0,16	0,87	***	0,13	0,36	***	0,13	0,96	***	0,16	0,47	***	0,11	1,08	***	0,11
Growth in House prices	-0,03		0,04	-0,16	***	0,04	-0,03		0,03	-0,25	***	0,05	-0,06	**	0,03	-0,33	***	0,03
Change in Consumer Confidence Indicator	-0,08	**	0,03	-0,11	***	0,03	-0,09	***	0,03	-0,11	***	0,04	-0,10	***	0,02	-0,10	*	0,02
Change in Capacity utilization in the industrial sector	-0,14		0,11	-0,41	***	0,09	-0,16	*	0,08	-0,50	***	0,11	-0,24	***	0,07	-0,48	***	0,07
Stoxx50 stock market return	0,04	***	0,01	0,03	***	0,01	0,03	***	0,01	0,04	***	0,01	0,03	***	0,01	0,04	*	0,01
EU27 Real GDP growth	0,07		0,07	0,35	***	0,06	0,10	*	0,05	0,44	***	0,07	0,16	***	0,05	0,56	***	0,05
Number of observations	48.049			144.147			96.083			96.113			144.150			48.046		
Number of firms	3.817			8.291			6.420			5.530			8.728			2.762		
Number of events	242			391			396			237			523			110		
Sector effects	YES			YES			YES			YES			YES			YES		
QIC	2.739,2			4.733,6			4.586			2.881,4			6.119,9			1.365,5		
QICu	2.736,3			4.729,9			4.582			2.880,1			6.115,3			1.363,6		

Models (A-1) and (A-2) show the estimation results when the full sample is divided into subsamples of small and a large firms According to the following criterion: A particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 25th percentile of the full sample's book asset values. This classification is time-varying in the sense that the same firm can be classified as small in some years but large in other years. Similarly, models (A-3) and (A-4) show the estimation results when a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 50th percentile of the full sample's book asset values. Lastly, models models (A-5) and (A-6) show the estimation results when a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 75th percentile of the full sample's book asset values.

A comparison of the estimation results within the three subsamples of small firms—i.e. models (A-1), (A-3), and (A-5)—shows that the signs, magnitudes, and significance levels of the effects of both the firm-specific and the macroeconomic variables remain largely unaltered as the size-criterion defining the subsample of small firms increases. The only notable changes for the firm-specific variables are for the size variable itself, which becomes smaller and insignificant as the size-criterion defining a small firm increases, and for the manufacturing sector dummy, which becomes larger and significant. The only notable changes for the macroeconomic variables are for the slope of the yield curve and the change in capacity utilization, which gain magnitude and significance as the size-criterion defining a small firm increases. Furthermore, a comparison of models (A-1), (A-3), and (A-5) with the model for small firms in our main analysis in Table 7 shows that using a time-varying classification of small firms entails no material changes to the signs, magnitudes, or significance levels of the effects of our variables within the subsamples of small firms.

Similarly, a comparison of the estimation results within Table 9's three subsamples of large firms—i.e. models (A-2), (A-4), and (A-6)—shows that the signs, magnitudes, and significance levels of the effects of both the firm-specific and the macroeconomic variables remain largely unaltered as the size-criterion defining the subsample of small firms increases. Finally, a comparison of models (A-2), (A-4), and (A-6) with the model for large firms in our main analysis in Table 7 shows that using a time-varying classification of large firms entails no material changes to the signs, magnitudes, or significance levels of the effects of our variables within the subsamples of large firms.

8.2 Independent censoring and entry-pattern

Given that data is only available from 2003 onwards, the existing stock of firms entering the sample in 2003 may potentially be of better average quality than the firms entering at a later point in time. This bias would violate the assumption of independent censoring. To address this issue, estimation was done on a reduced sample that excludes firms entering the sample in 2003 (where a considerable part of these entries ties to the existing stock of the bank clients). The results (not presented here, but available upon request) are that all estimated coefficients remain significant and of the same sign as the final model (Model (7) in Table 5). To address the concern that the very low number of entries in 2005 might have an impact on the results, the final model specification was re-estimated using two samples: One that exclude entries from 2005, and another that excludes all entries up until 2006. The estimates from these model fits (not presented here, but available upon request) are still all significant and of the same sign as the model estimated on the full sample.

8.3 Lag length

We have throughout chosen to focus on a lag length of one year for the covariates employed in our intensity models. One may, however, believe that for macroeconomic variables, this is not the appropriate lag, as aggregate changes may take longer to impact firms. This may, for instance, be because firms operate with a cash buffer that allow them to operate though an extended period of time before a default is observed. To address this concern, we re-estimated the preferred model with all macroeconomic variables lagged eight quarters instead of four. The results (not presented here, but available upon request) showed that the macroeconomic variables are generally less able to explain defaults when lagged eight quarters, as indicated by the loss in significance for most of the coefficients.

Still, this analysis does not consider the possibility that different macroeconomic variables are operating though different lag periods. Considering all possible combinations of lag periods would result in an extensive number of permutations of the model for us to check. Instead, we constructed a correlation matrix for the observed default rate and each of the macroeconomic time series lagged from zero to eight quarters. It generally shows that, while the lag length of four quarters does not provide the highest correlation with the default rate for all macroeconomic variables, it appears that a unified lag period of four quarters is at least a very appropriate choice.

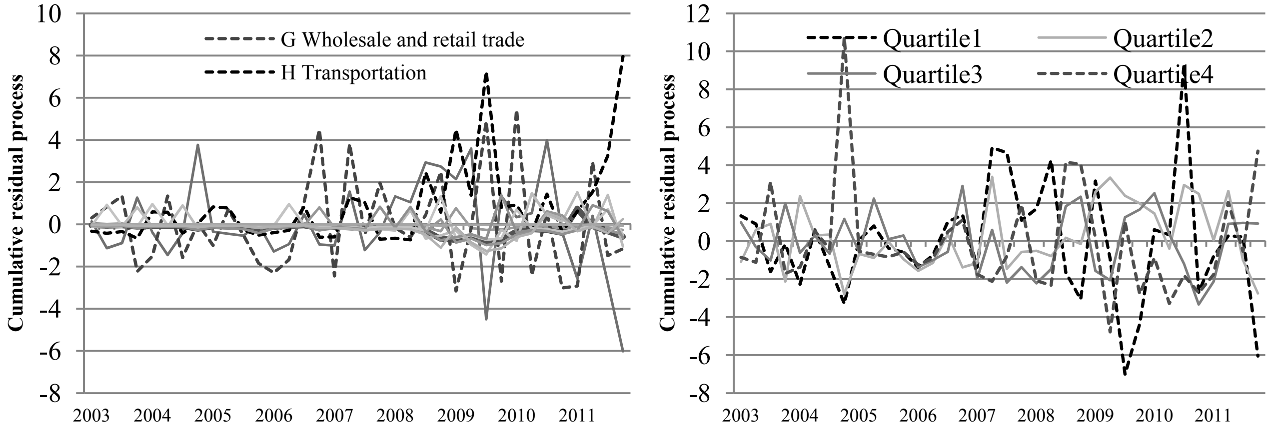

Figure 8. Model check based on grouped martingale residual processes. The figure shows cumulative martingale residual processes for the preferred Cox model including both firm-specific and macroeconomic variables (Model (7) in Table 5). In the left panel, grouping corresponds to each firm's sector, which is stationary across the sample period, while in the right panel, the grouping is time-varying and done According to asset quartiles.

8.4 Grouped martingale residual processes

We check the fit of our final model using the martingales residual processes (2). Specifically, we consider to what extent the model is systematically over-or underestimating the default frequency in different sectors and size-groups.

By definition, the martingale residual processes are the difference between the observed default frequency and the default frequency predicted by the model. Since a single firm can at most have one default event in our estimation setup, the firm-specific processes contain too little information. However, when grouped in sufficiently large clusters, the increments of the grouped processes should not be systematically positive or negative if the model fit is adequate. An increasing grouped residual process would imply that the model is under-predicting the number of defaults this particular group, whereas a decreasing grouped residual process would imply that the model is predicting too many defaults for this group.

The left panel of Figure 8 shows grouped residual processes by sector as a function of time. We see that the residual processes fluctuate around zero with both positive and negative increments for all sectors. This is support for the model performing equally well for all sectors. However, noting that the sectors "wholesale and retail trade" and "transportation" have the largest deviances, we re-estimated our final model excluding firms in these two sectors – the results (not presented here, but available upon request) do not change.

The right panel of the figure shows grouped residual processes by asset quartiles as a function of time. We again observe no truly systematic deviations. We note, however, that the largest quarterly deviances occur in the largest and smallest quartiles, further motivating the point that default prediction models should take firm size into account.

9 Conclusion

The Basel II and III Accords impose smaller capital charges on bank loans to SMEs by assuming a lower correlation in the formula for unexpected loss in a loan portfolio. The reduction in correlation for SMEs corresponds to an assumption that these firms have default probabilities that are less sensitive to the economy-wide factor. This paper investigates whether there is empirical support for this assumption. Using a default intensity regression framework, our results indicate that solely discriminating with respect to firm-size, the default probabilities of small firms do in fact exhibit less sensitivity to macroeconomic cyclicality compared to the default probabilities of large firms, in the sense that the effects of macroeconomic variables are of smaller magnitude for smaller firms. However, when we account for differences in firm-characteristics other than size, our results indicate that the default probability of the average small firm is as cyclical or even more cyclical than the default probability of the average large firm. The results are robust to different regression models and different divisions of our sample into small and large firms. Our findings suggest that a reduction in capital charges based solely on firm-size may imply a higher risk for banks with a high exposure to the SME-segment.

References

Aalen, O. O. (1980). A model for non-parametric regression analysis of life times. In J. R. W. Klonecki, A. Kozek (Ed.), *Mathematical Statistics and Probability Theory*, Volume 2 of *Lecture Notes in Statistics*, New York, USA, pp. 1–25. Springer-Verlag.

Aalen, O. O. (1989). A linear regression model for the analysis of life times. *Statistics in medicine 8*, 907–925.

Aalen, O. O., Ø. Borgan, and H. K. Gjessing (2008). *Survival and Event History Analysis: A Process Point Of View* (1st ed.). New York, USA: Springer-Verlag.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance 23*, 589–609.

Andersen, P. K., Ø. Borgan, R. D. Gill, and N. Keiding (1992). *Statistical models based on counting processes*. Springer-Verlag.

Beaver, W. (1966). Financial ratios and the prediction of failure. *Journal of Accounting Research. Supplement : Empirical research in accounting : Selected studies 1966 4*, 77–111.

Bharath, S. T. and T. Shumway (2008). Forecasting Default with the Merton Distance to Default Model. *Review of Financial Studies 21*(3), 1339–1369.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–654.

Bonfim, D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *The Journal of Banking and Finance 28*, 281–299.

Carling, K., T. Jacobson, J. Lindé, and K. Roszbach (2007). Corporate credit risk modeling and the macroeconomy. *Journal of Banking and Finance 31*(3), 845–868.

Chava, S. and R. Jarrow (2004). Default risk and diversification : Theory and empirical implications. *Review of Finance 8*(4), 537–569.

Chava, S., C. Stefanescu, and S. Turnbull (2011). Modeling the loss distribution. *Management Science 57*(7), 1267–1287.

Chionsini, G., J. Marcucci, and M. Quagliariello (2010, March). The Treatment of Small and Medium Enterprises in Basel 2: So Right, So Wrong? Technical report, Center for Economic Policy Research (CEPR), http://dev3.cepr.org/meets/wkcn/1/1737/papers/Marcucci.pdf.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 34*, 187–220.

Dietsch, M. and J. Petey (2004). Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs. *Journal of Banking and Finance 28*(4), 773–788.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009, October). Frailty correlated default. *The Journal of Finance LXIV*(5), 2089–2123.

Duffie, D. and D. Lando (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica 58*, 635–665.

Duffie, D., L. Saita, and K. Wang (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics 83*, 635–665.

European Commission (2016). Annual Report on European SMEs 2014/2015. Technical Report Ref. Ares(2016)1791252 - 15/04/2016, European Commission.

Figlewski, S., H. Frydman, and W. Liang (2012). Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics & Finance 21*(1), 87–105.

Giesecke, K., D. Lando, and M. Medhat (2013, September). The Macroeconomy and Credit Risk: Direct and Indirect effects. *Working Paper, Copenhagen Business School and Stanford University*.

Jacobson, T., J. Linde, and K. Roszbach (2005). Credit risk versus capital requirements under Basel II: are SME loans and retail credit really different? *Journal of Financial Services Research 28*(1), 43–75.

Lando, D., M. Medhat, M. S. Nielsen, and S. F. Nielsen (2013). Additive Intensity Regression Models in Corporate Default Analysis. *Journal of Financial Econometrics 11*(3), 443–485.

Lando, D. and M. S. Nielsen (2010). Correlation in Corporate Defaults: Contagion or Conditional Independence? *Journal of Financial Intermediation 19*(3), 335–372. http://ssrn.com/paper=1338381.

Leland, H. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance 49*, 1213–1252.

Lopez, J. A. (2004). The empirical relationship between average asset correlation, firm probability of default, and asset size. *Journal of Financial Intermediation 13*(2), 265–283.

Martinussen, T. and T. H. Scheike (2006). *Dynamic Regression Models for Survival Data* (1. ed.). New York, USA: Springer-Verlag.

McDonald, C. G. and L. M. V. de Gucht (1999). High-Yield Bond Default and Call Risks. *The Review of Economics and Statistics 81*(3), 409–419.

Merton, R. C. (1974). On the pricing of coroporate debt: The risk structure of interest rates. *Journal of Finance 29*, 449–470.

Ohlsen, J. S. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Account. Res. 19*, 109–31.

Peseran, M. H., T. Schuermann, B.-J. Treutler, and S. M. Weiner (2006). Macroeconomic Dynamics and Credit Risk: A Global Perspective. *Journal of Money, Credit, and Banking 38*, 1211–1262.

Shumway, T. (2001). Forecasting bankruptcy more efficiently : A simple hazard model. *Journal of Bus. 74*, 101–124.

The Basel Committee on Banking Supervision (2006, June). International Convergence of Capital Measurement and Capital Standards – A Revised Framework (Comprehensive Version). Technical report, Bank for International settlements (BIS).

The European Parliament and the Council of the European Union (2013, June). Regulation (eu) no 575/2013 of the european parliament and of the council of 26 june 2013 on prudential requirements for credit institutions and investment firms and amending regulation (eu) no 648/2012. Technical report, The European Parliament and the Council of the European Union, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:176:0001:0337:EN:PDF.

Wooldridge, J. M. (2009). *Introductory Economics—A Modern Approach* (4rth ed.). South-Western.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Account. Res. 22*, 59–82.