WORKING PAPER

DANMARKS NATIONALBANK

22 JANUARY 2020 — NO. 151

Modeling Frailty Correlated Defaults with Multivariate Latent Factors

Benjamin Christoffersen
bch.fi@cbs.dk
COPENHAGEN BUSINESS SCHOOL

Rastin Matin
rma@nationalbanken.dk
DANMARKS NATIONALBANK

WORKING PAPER — DANMARKS NATIONALBANK
22 JANUARY 2020 — NO. 151

Modeling Frailty Correlated Defaults with Multivariate Latent Factors

Abstract

Firm-level default models are important for bottom-up modeling of the default risk of corporate debt portfolios. However, models in the literature typically have several strict assumptions which may yield biased results, notably a linear effect of covariates on the log-hazard scale, no interactions, and the assumption of a single additive latent factor on the log-hazard scale. Using a sample of US corporate firms, we provide evidence that these assumptions are too strict and matter in practice and, most importantly, we provide evidence of a time-varying effect of the relative firm size.
We propose a frailty model to account for such effects that can provide forecasts for arbitrary portfolios as well. Our proposed model displays superior out-of-sample ranking of firms by their default risk and forecasts of the industry-wide default rate during the recent global financial crisis.

Resume

Konkursmodeller på virksomhedsniveau er vigtige for bottom-up-modellering af en virksomhedsporteføljes konkursrisiko. Modeller i litteraturen har dog typisk flere strenge antagelser, der kan resultere i bias, særligt en antagelse om en lineær effekt af kovariaterne på log-hazard-skalaen, ingen vekselvirkninger, og en antagelse om en enkel additiv latent faktor på log-hazard-skalaen. Med udgangspunkt i et datasæt bestående af amerikanske virksomheder finder vi evidens for, at disse antagelser er for strenge, at de har en betydning i praksis og, endnu vigtigere, at der er evidens for en tidsvarierende effekt af den relative virksomhedsstørrelse. Vi foreslår en frailty-model til at tage højde for disse effekter, der også kan anvendes til at forecaste for arbitrære porteføljer. Blandt de undersøgte modeller udviser vores foreslåede frailty-model den bedste out-of-sample-rangering af virksomheder efter deres kreditrisiko og de bedste forecasts af konkursraten for hele økonomien under den seneste globale finansielle krise.

Key words

Credit risk; Risk management

JEL classification

C53, C55, G33, M41

Acknowledgements

The authors are grateful to Mads Stenbo Nielsen, Søren Feodor Nielsen, and seminar participants at Copenhagen Business School for helpful comments.

The authors alone are responsible for any remaining errors.

Modeling Frailty Correlated Defaults with Multivariate Latent Factors

Benjamin Christoffersen[1] Rastin Matin[2]

[1]Department of Finance, Center for Statistics, Copenhagen Business School, DK-2000 Frederiksberg, Denmark, bch.fi@cbs.dk
[2]Danmarks Nationalbank, DK-1093 Copenhagen, Denmark, rma@nationalbanken.dk

January 14, 2020

Abstract

Firm-level default models are important for bottom-up modeling of the default risk of corporate debt portfolios. However, models in the literature typically have several strict assumptions which may yield biased results, notably a linear effect of covariates on the log-hazard scale, no interactions, and the assumption of a single additive latent factor on the log-hazard scale. Using a sample of US corporate firms, we provide evidence that these assumptions are too strict and matter in practice and, most importantly, we provide evidence of a time-varying effect of the relative firm size. We propose a frailty model to account for such effects that can provide forecasts for arbitrary portfolios as well. Our proposed model displays superior out-of-sample ranking of firms by their default risk and forecasts of the industry-wide default rate during the recent global financial crisis.[1]

Modeling the default risk of a corporate debt portfolio can be accomplished by modeling the default risk of the portfolio's individual firms and then aggregate up to the portfolio level. This method is advantageous as it is easy to account for changes in the portfolio through time. It is, however, commonly known that misspecification of the firm-level model or omitted variables can lead to a large downward bias in risk measures. With this in mind, we perform an analysis of a sample of large US corporate firms from 1980 to 2016 and our results are twofold: First, our results strongly suggest that earlier models in the literature have been misspecified and, secondly, we present a model that accounts for the misspecification. With this model, we show that it is necessary to consider non-linear transformations of certain variables, interactions, and account for time-varying effects of the relative firm size. Our out-of-sample results show better ranking of firms by their default risk and a better forecast of the industry-wide default rate during the last crisis.

Typical default models in the literature use firm-specific variables and macro-variables to model the probability of a future default of a firm, see e.g. Shumway (2001), Chava and Jarrow (2004), Duffie et al. (2007), Campbell et al. (2008). These models provide quite accurate ranking of firms by their default risk. However, the predicted default rate distributions of corporate debt portfolios are too narrow for some data sets and model specifications, indicating a violation of the models' assumption of conditional independence between firms given the observable variables. Many ideas have been suggested to relax this assumption (Duffie et al. 2009, Koopman et al. 2011, 2012, Duan and Fulop 2013, Chen and Wu 2014, Qi et al. 2014, Schwaab et al. 2017). Common for these is that they all introduce one latent variable which affects all firms equally either within an industry, rating group, or across all industries and rating groups on the log-hazard scale or logit scale in

[1]The data preparation code is available at http://bit.ly/github-US_PD_data-r2 and the data analysis is available at bit.ly/github-US_PD_models-r3.

discrete time. These models are known as mixed models or frailty models when a multiplicative random factor is included in the instantaneous hazard. The addition of the random factor results in wider and more reliable prediction intervals for the default rate of a group of firms, resulting in more accurate risk measures. These models are therefore better suited for modeling risk measures of a corporate debt portfolio. Since the random factor affects groups of firms equally on the log-hazard scale, the models often do not provide better forecasts of the mean, nor do they improve the ranking of firms by their riskiness. This has been explicitly shown by Qi et al. (2014).

Within the frailty literature it is common to assume that the coefficients for observable variables are constant through time, but Lando et al. (2013) show that this assumption is too strict. Using non-parametric and semi-parametric models, they present evidence of non-constant coefficients, and not accounting for such effects may yield biased results and an invalid implied distribution for the default rate of a debt portfolio. However, the models in Lando et al. (2013) cannot directly be used for forecasting due to their non-parametric components. In this work we bridge these two approaches by presenting a frailty model that relaxes the assumption of constant coefficients, which is also able to forecast future default rates and properly account for conditional correlation.

We first ensure that any time-varying effects are not due to an invalid specification of the linear predictor. We find that additional variables are needed in the model specification of Duffie et al. (2007, 2009), since we observe a smaller difference in log-likelihood than Duffie et al. (2009) between a model with and without frailty. This is consistent with Lando and Nielsen (2010) that cannot reject the misspecification test of Duffie et al. (2007) after inclusion of additional firm-specific covariates.

Based on work by Berg (2007) and Christoffersen et al. (2018), we expect non-linear effects of some covariates on the log-hazard scale. We account for these by natural cubic splines and, unlike Lando and Nielsen (2010), we indeed find a significant non-linear effect of the idiosyncratic stock volatility of the firm, the net income over total assets, and log market value over total liabilities. Accounting for non-linear effects in this manner is rarely done in the literature even though there is no a priori reason to suspect that covariates should have a linear effect on the log-hazard scale. Our findings of a non-linear effect of the idiosyncratic stock volatility and log market value over total liabilities provide further evidence that the Merton model provides useful guidance for building default models but may need adjustments. As Lando et al. (2013), Filipe et al. (2016), Jensen et al. (2017), we also find strong evidence of a time-varying coefficient for a size measure of the firm. In this regard we note that our size variable differs from the aforementioned papers by using the market value as in Shumway (2001).

Section 1 describes in detail the hazard and frailty models we use, and in Section 2 we describe our data set. We present results for monthly hazard rate models with and without frailty in Section 3. We begin the section by providing evidence of non-linear effects and an interaction in models without frailty and then we extend these models to include frailty variables. Section 4 contains an out-of-sample test of the models, and we conclude in Section 5.

1 Model Specification

Our baseline model is a Cox proportional hazards model with a constant baseline hazard and time-varying covariates as in Duffie et al. (2007). Thus, the conditional instantaneous hazard rate of firm i at time t is

$$\lambda_i\left(t \mid \boldsymbol{x}_i(t), \boldsymbol{m}(t)\right) = R_i(t) \exp\left(\alpha + \boldsymbol{\beta}^\top \boldsymbol{x}_i(t) + \boldsymbol{\gamma}^\top \boldsymbol{m}(t)\right) \tag{1}$$

where $R_i(t)$ is a censoring indicator which is zero if the individual i is censored at time t, $\boldsymbol{x}_i(t)$ are the firm-specific covariates at time t, $\boldsymbol{m}(t)$ are the macro-variables at time t and α, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are unknown parameters which we need to estimate. We observe $R_i(t)$, $\boldsymbol{x}_i(t)$, and $\boldsymbol{m}(t)$ at discrete points so we model these as variables that are piecewise constant. Thus, the instantaneous hazard in Equation (1) is piecewise constant, resulting in a piecewise exponentially distributed arrival time. Due to the discrete hazard, we simplify the notation to

$$\lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k\right) \equiv \lambda_i\left(t \mid \boldsymbol{x}_i(t), \boldsymbol{m}(t)\right) = R_{ik} \exp\left(\alpha + \boldsymbol{\beta}^\top \boldsymbol{x}_{ik} + \boldsymbol{\gamma}^\top \boldsymbol{m}_k\right) \tag{2}$$

where $k = \lceil t \rceil$, $\lceil t \rceil$ is the ceiling of t, and we let $\boldsymbol{x}_{ik}$ be the constant value that $\boldsymbol{x}_i(t)$ takes on the interval $(k-1, k]$ and similarly for R_{ik} and $\boldsymbol{m}_k$. Letting T_i denote the default time of firm i, then the parameters are easily estimated by using that the likelihood in discrete time (i.e., the probability of default) for firm i in period $(k-1, k]$ conditional on survival up to time $k-1$ is

$$P\left(T_i \in (k-1, k] \mid T_i > k-1\right) = 1 - \exp\left(-\lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k\right)\right) \tag{3}$$

for $k \in \{0, 1, 2, \dots\}$, and in continuous time by using that the conditional density function of T_i is

$$f_i\left(k - 1 + \Delta t \mid T_i > k-1\right) = \lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k\right) \exp\left(-\lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k\right) \Delta t\right)$$

where $\Delta t \in (0, 1)$.

The strict assumption in Equation (1) is that firms' default intensities, λ_{ik}, are only correlated through co-movements in firm-specific covariates, $\boldsymbol{x}_{ik}$, or the macro-variables, $\boldsymbol{m}_k$. This assumption may not hold in practice for multiple reasons: Our model could be misspecified by omission of one or more macro-variables, co-movements in an omitted firm covariate, or co-movements in a variable which is modeled incorrectly. As in Duffie et al. (2009),[2] one way to account for the excess clustering of defaults is by extending the hazard in Equation (2) to

$$\lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k, A_k\right) = R_{ik} \exp\left(\alpha + \boldsymbol{\beta}^\top \boldsymbol{x}_{ik} + \boldsymbol{\gamma}^\top \boldsymbol{m}_k + A_k\right) \tag{4}$$

$$A_k = \theta A_{k-1} + \epsilon_k, \qquad \epsilon_k \sim \mathcal{N}\left(0, \sigma^2\right) \tag{5}$$

where $\mathcal{N}\left(0, \sigma^2\right)$ is a normal distribution with mean zero and variance σ^2 and the innovation terms, ϵ_k, are iid. The term A_k increases the hazard of all firms in period $(k-1, k]$ by a factor $\exp(A_k)$, and such a multiplicative factor on the hazard is formally referred to as a frailty. Large values of σ increase the probability of observing larger differences in the log-hazard between consecutive periods all else being equal, while θ controls the rate of decay of the cumulated shocks. That is, the effect of the shock ϵ_k decays towards zero with the rate $\theta^{k'-k}$ for $k' > k$. The limit $\theta \to 0$ corresponds to independent values of A_k, which has previously been employed in, e.g., Christoffersen et al. (2018) as they do not find evidence of autoregressive random effects. This could possibly be due to wider time intervals and/or fewer cross sections.

The model specification determines whether A_k is a zero-mean stationary or non-stationary process. The former would be the case when the frailty captures a stationary macro-variable. However, if the model specification, e.g., includes the log of a nominal (non-real) value variable, but the true association is linear in the log of the real value, then a non-stationary adjustment to the intercept would be needed. Examples of the latter are found in Lando and Nielsen (2010) and Chava et al. (2011), where it is unclear whether the authors use nominal or real values. As we only include real values and financial ratios where inflation adjustments do not matter, we do not expect a non-stationary intercept for the aforementioned reasons.

While the model in Equation (4) can account for conditional correlation, it may still be a poor approximation of the true dynamics. First, although we may expect a monotone effect, it is not obvious that the log of the default intensity should have a linear dependence on the covariates. We will later account for such non-linear effects by using natural cubic splines. Secondly, the

[2] Duffie et al. (2009) remark that they use an Ornstein-Uhlenbeck process, but in practice they use a discrete approximation like in Equations (4) and (5).

assumption that there is only one frailty variable which affects all firms equally on the hazard scale may not be justified. A generalization is to relax the assumption of constant proportional effect of the covariates and let some of the coefficients vary over time. The resulting model is

$$\begin{gathered}\lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k, A_k, \boldsymbol{B}_k, \boldsymbol{z}_{ik}\right) = R_{ik}\exp\left(\alpha + \boldsymbol{\beta}^\top\boldsymbol{x}_{ik} + \boldsymbol{\gamma}^\top\boldsymbol{m}_k + A_k + \boldsymbol{B}_k^\top\boldsymbol{z}_{ik}\right) \\ \begin{pmatrix} A_k \\ \boldsymbol{B}_k \end{pmatrix} = \mathbf{F}\begin{pmatrix} A_{k-1} \\ \boldsymbol{B}_{k-1} \end{pmatrix} + \boldsymbol{\epsilon}_k, \qquad \boldsymbol{\epsilon}_k \sim \boldsymbol{\mathcal{N}}\left(\mathbf{0}, \mathbf{Q}\right)\end{gathered} \tag{6}$$

where $\boldsymbol{z}_{ik}$ may contain a subset of the elements of $\boldsymbol{x}_{ik}$ and $\boldsymbol{\mathcal{N}}\left(\mathbf{0}, \mathbf{Q}\right)$ is a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{Q}$. The term $\boldsymbol{B}_k$ contains the random components of the slopes, and the interpretation of $\boldsymbol{B}_k$ is the change in log-hazard relative to the reference point ($\boldsymbol{B}_k = \mathbf{0}$) due to a unit increase in the covariates with a random slope. E.g., suppose that we only have one covariate in the model and it has a random slope. Then two firms i and j which differ by $x_{jk} = x_{ik} + 1$ will have a relative hazard in time period $(k-1, k]$ of

$$\frac{\lambda_{jk}\left(x_{jk}, \boldsymbol{m}_k, A_k, B_k, x_{jk}\right)}{\lambda_{ik}\left(x_{ik}, \boldsymbol{m}_k, A_k, B_k, x_{ik}\right)} = \frac{\exp\left(\left(\beta + B_k\right)\left(x_{ik} + 1\right)\right)}{\exp\left(\left(\beta + B_k\right)x_{ik}\right)} = \exp\left(\beta + B_k\right)$$

That is, B_k changes the effect of a unit change in the covariate by a factor $\exp(B_k)$.

There are many reasons to expect non-constant slopes. The accounting standards may change, banks may temporally change the way that variables affect their lending behavior, certain types of firms may be more risky in poor economic downturns, etc. One may again argue whether the frailty is a stationary process or not. If one expects that the frailty captures temporary excess default clustering, then a stationary process seems like a natural choice.

We will estimate the frailty models using a Monte Carlo expectation maximization algorithm, where the expectation step is approximated using a particle smoother. The details of the estimation are in Appendix A. The software we have developed to estimate the frailty models is available at the Comprehensive R Archive Network (CRAN).

2 Data and Choice of Covariates

Moody's Default Risk Service Database (MDRD) is used to get the firms' default events. We define a default event as a firm which enters into either bankruptcy, bankruptcy section 77, chapter 10, chapter 11, chapter 7, or a prepackaged chapter 11. We also regard the following as default events: A distressed exchange, a dividend omission, a grace-period default, a modification of indenture, a missed interest payment and/or a missed principal payment, payment moratorium, and a suspension of payments. These events are also included by Duffie et al. (2009) from MDRD and are nearly the same events included by Lando et al. (2013). The by far most frequent event is a missed interest payment followed by a chapter 11 bankruptcy and, as some of our events are not terminal, recurrent events can occur. Firms with multiple events typically have an intermediary period in which most would consider the firm as being in a non-normal state and thus not being at risk of entering into default. Thus, we censor a firm until the resolution date provided by MDRD or 12 months after the event if the resolution date is missing. We extend the censoring period if consecutive events fall within this default event time and the resolution date.

We only use MDRD for two reasons: First, some of the default events are closer to the point in time at which, e.g., bond holders suffer losses. Secondly, we can use the same default events for all firms in our sample. We could augment our data set with firms that are not in MDRD, but then we would track different events depending on whether the firm is tracked by Moody's. Thus, the event definition would be broader for firms in MDRD as we would likely only have legal bankruptcy events available for firms outside MDRD. Consequently, our results could reflect differences between the two groups and it would be unclear what we model.

We use CRSP and Compustat for market data and financial statements, respectively. We lag data from Compustat by 3 months to reflect the typical delay on financial statements, use quarterly data with annualized flow variables when available and otherwise we use yearly data. Data from CRSP is lagged by 1 month to reflect that we only know past market data. Summary statistics are shown in Table 4 in the appendix, and the firm-specific variables we include are:

- *Operating income to total assets*: Operating income after depreciation relative to total assets. It is a profitability measure and we expect that more profitable firms should be less likely to enter into default.

- *Net income to total assets*: Net income relative to total assets. It is similarly a profitability measure but includes all costs. Including both ratios allows one to distinguish between the partial association of the two types of costs.

- *Market value to total liabilities*: Market value from CRSP relative to total liabilities. A larger ratio should imply that the firm is further from default all else equal.

- *Total liabilities to total assets*: Total liabilities relative to total assets. It is an indicator of the firm's financial leverage and we expect that all else equal a higher ratio should imply a higher probability of default.

- *Current ratio*: Current assets relative to current liabilities. A too low ratio would imply that the firm may not be able meet its short-term debt obligations thus increasing the probability of default.

- *Working capital to total assets*: Working capital relative to total assets. Similar to the current ratio, it measures the ability to meet the short-term debt obligations but does so with a metric relative to the size of the firm.

- *Log current assets*: Log of current assets deflated with the US Government Consumer Price Index from CRSP. This is similar to the pledgeable assets used in Lando et al. (2013) but we do not add the book value of net property, plant, and equipment to the current assets. The variable captures both the size of the firm and the assets which can be quickly converted into cash.

- *Log excess return*: 1-year lagged average of monthly log return minus the value-weighted log total market return. We require at least three months of returns. While we do not have a particular effect in mind, this variable has shown to be a strong predictor in the literature (Shumway 2001, Duffie et al. 2007, 2009).

- *Relative log market size*: Log market value of the firm minus the log total market value. The total market value is the sum of market values of AMEX-, NYSE-, and NASDAQ-listed firms. As remarked by Shumway (2001), subtracting the log total market value from the log market value of the firm has the advantage that it deflates the nominal log market value. Low-valued firms should be closer to entering into default in which case any investments by investors are likely lost. Though, the variable also measures the size of the firm.

- *Distance-to-default*: Estimated 1-year distance-to-default. The drift and volatility of the underlying assets are estimated over the past year using the so-called KMV method as in Vassalou and Xing (2004), and we set the debt due in one year to be the short-term debt plus 50% of the long-term debt as is common. We require at least 60 days of market values to estimate the parameters.

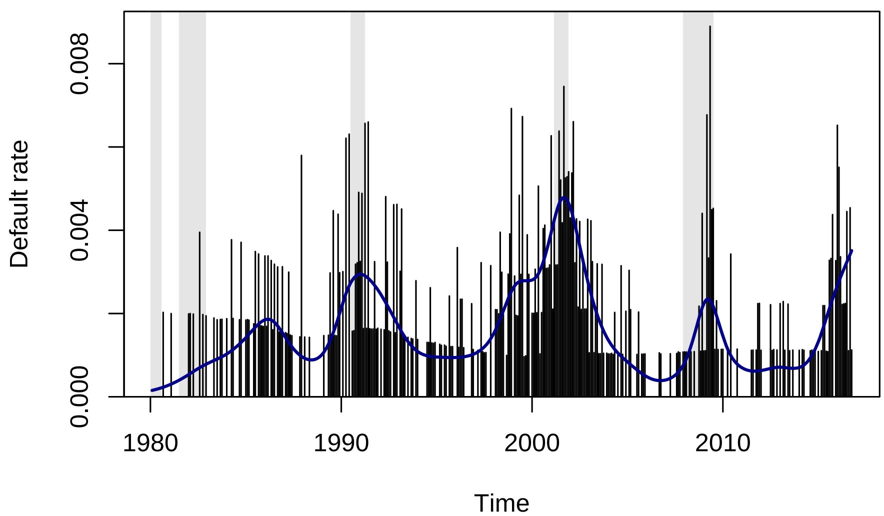


Figure 1: Monthly default rate in the sample. The blue line is a natural cubic spline with an integrated square second derivative cubic spline penalty. The penalty parameter is chosen using an un-biased risk estimator criterion. Gray areas are recession periods from National Bureau of Economic Research.

The statistics of our distance-to-default is comparable to that reported by Vassalou and Xing (2004), which is anticipated as we use the same method and listed US firms.[3] However, we note that a wide range of values have been reported in the literature.[4]

- *Idiosyncratic volatility*: Estimated standard deviation of 1-year past rolling window regressions of daily log return on the value-weighted log market return. We require at least 60 days of returns in the regressions. The variable is used in Shumway (2001) and one motivation is that more volatile firms should have a higher chance of entering into default (e.g., due to more volatile cash flows as argued by Shumway 2001).

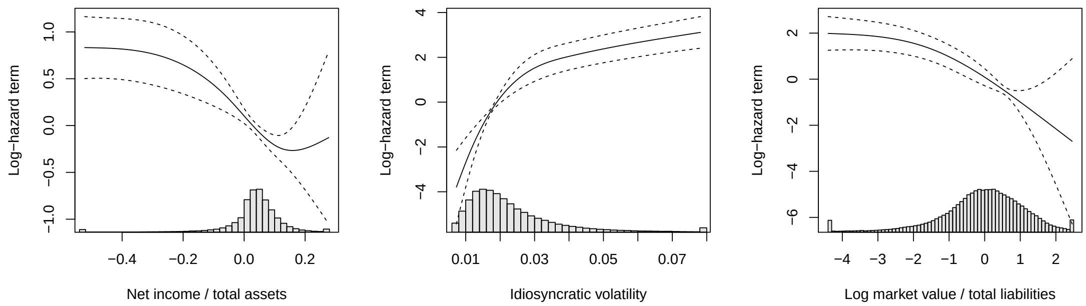
The value-weighted market return we use is the NYSE and AMEX index from CRSP. In terms of macro-variables, we include a market return and treasury bill rate like Duffie et al. (2007, 2009), specifically the value-weighted past 1-year log return of the aforementioned index and 1-year treasury bill rate. All variables are winsorized at 1% and 99% quantile and we carry forward missing covariates for up to 3 months for CRSP-based variables and 1 year for Compustat-based variables.

All of these covariates have appeared in multiple papers before (e.g., Shumway 2001, Chava and Jarrow 2004, Bharath and Shumway 2008). It is deliberate that we use covariates that have previously been used in the literature as the goal of this work is not to seek new covariates.

We include a firm in our sample as long as it is listed, we have data from Compustat and CRSP for all variables, and the firm has started being rated by Moody's or if it is less than 36 months

[3]The probabilities of default in Vassalou and Xing (2004) are available at `www.maria-vassalou.com/data/defaultdataset.zip`. Comparing our distance-to-default to theirs over the same period after truncating at a 10^{-15} and $1-10^{-15}$ probability of default as they do yields a mean and standard deviation of the distance-to-default of 4.856 and 2.739 , respectively. The corresponding figures in Vassalou and Xing (2004) are 4.391 and 2.608, respectively.

[4]Lando and Nielsen (2010), Chava et al. (2011), Duan et al. (2012), Lando et al. (2013), Qi et al. (2014) show a mean ranging from 1.867 to 16.79 and a standard deviation ranging from 2.653 to 12.83.



(a) **(b)** **(c)**

Figure 2: Splines in model $\mathcal{M}_3$. Splines for (a) net income to total assets, (b) the idiosyncratic volatility, and (c) log market value to total liabilities containing both the linear and non-linear components in Table 1. The y-axes are the effect of the covariate on the log-hazard scale and dashed lines are 95% confidence intervals. Histograms of the covariates are shown in the background to illustrate where we have data (they have no relation to the y-axis).

after the rating has been withdrawn and the firm is not rated again by Moody's. Firms outside this range have a virtually zero default rate in the MDRD database, which is likely because Moody's no longer tracks the firms or has not yet started to track them.

A delisting month counts as a default if a default event happens up to one year after the firm delists. This is similar to Shumway (2001) though he uses a five-year limit instead. An advantage of the event definition we use relative to, e.g., Shumway (2001) is that the events happen close to delisting or before delisting. Specifically, we observe that 84.7% of the events occur while we still have covariate information of the firm and the firm has not delisted or delists in the month of the event. Thus, we include the first half of 2016 in our out-of-sample test in Section 4 as we may only miss a few events since our version of MDRD was last updated in October 2016.

As is standard in the literature, we exclude firms with an SIC-code in the range 6000-6999 (financial firms) and those greater than 9000 (public administration or non-classifiable). We use the historical SIC-code from Compustat if it is available and otherwise use CRSP's historical code.

Figure 1 shows the monthly default rate in the sample. There is a visible clustering of defaults around economic crises, however, it is not clear from this plot whether the clustering can be captured by firm-specific variables or macro-variables.

3 Empirical Results

We start this section by estimating models without frailty and show that we get a better fit by adding covariates, splines, and an interaction to the model in Duffie et al. (2007). That additional covariates are needed is similar to the conclusion in Lando and Nielsen (2010), but treating non-linear effects has so far received limited attention in corporate default literature. We find a non-linear effect of variables that are related to the Merton model and discuss this finding and its relation to previous work. We then turn to frailty models, where we estimate a model with a single frailty that affects all firms equally on the log-hazard scale (cf., Equation (4)), and subsequently extend it to include a frailty that depends on the relative log market size of the firm as in Equation (6). We end by comparing our results to previous work.

Table 1 presents the parameter estimates for the models without frailty. Column $\mathcal{M}_1$ is similar to the specification used by Duffie et al. (2007) and Duffie et al. (2009) when the frailty variable is not included. A one standard deviation change in distance-to-default (log excess return) is

Table 1: Coefficient estimates from monthly default models without frailty. All covariates are centered in the shown models. The χ^2 test statistics from the Wald tests are given in parentheses: *** implies significance at the 1% level, ** at the 5% level, and * at the 10% level. The spline rows are from a three-dimensional natural cubic spline basis, which is restricted to a two-dimensional space that is orthogonal to the linear term. Thus, the linear term can be interpreted as the linear part of the spline. The two coefficients for each basis function in the splines are omitted and a "✓" indicates that the spline term is included in the model.

	$\mathcal{M}_1$		$\mathcal{M}_2$		$\mathcal{M}_3$	
Intercept	-9.938^{***}	(2379.1)	-10.265^{***}	(1264.2)	-10.436^{***}	(911.8)
Distance-to-default	-0.546^{***}	(227.6)	0.001	(0.0)	0.148^{***}	(7.9)
Log excess return	-2.331^{***}	(420.7)	-1.744^{***}	(207.7)	-1.729^{***}	(205.5)
T-bill rate	1.827	(1.2)	4.459^{**}	(6.1)	4.751^{***}	(6.9)
Log market return	1.132^{***}	(18.3)	0.877^{***}	(10.0)	0.890^{***}	(10.1)
Log current assets			0.178^{***}	(18.4)	0.158^{***}	(10.7)
Working capital / total assets			-1.060^{**}	(4.1)	-0.888^{*}	(2.8)
Operating income / total assets			-1.643^{***}	(9.7)	-1.180^{**}	(4.7)
Market value / total liabilities			-1.302^{***}	(24.9)		
Log market value / total liabilities					-0.806^{***}	(11.6)
Net income / total assets			-1.156^{***}	(16.2)	-2.041^{***}	(15.2)
Total liabilities / total assets			0.595^{***}	(7.4)	0.411^{*}	(3.3)
Current ratio			-0.188^{*}	(3.7)	-0.816^{***}	(20.3)
Idiosyncratic volatility			26.844^{***}	(60.9)	108.095^{***}	(41.1)
Relative log market size			-0.356^{***}	(47.6)	-0.293^{***}	(25.0)
Net income / total assets (spline)					$\checkmark^{**}$	(6.2)
Idiosyncratic volatility (spline)					$\checkmark^{***}$	(27.0)
Log market value / total liabilities (spline)					$\checkmark^{***}$	(19.3)
Current ratio · Idiosyncratic volatility					15.293^{***}	(19.8)
AIC	5511.7		5067.0		5004.4	
log-likelihood	-2750.8		-2519.5		-2481.2	
Number of firms	3020		3020		3020	

associated with an exp(−2.263) (exp(−0.932)) factor change in the hazard. The former shows that distance-to-default is a good predictor of default as changes in distance-to-default are associated with large changes in the instantaneous hazard. However, the latter shows that distance-to-default is not a sufficient explanatory covariate on its own, as is also observed in Bharath and Shumway (2008).

The signs of the coefficient estimates for the two macro-variables are similar to Duffie et al. (2009) and like them we do not have an intuitive explanation for the log market return slope estimate that makes sense marginally. However, we agree that a univariate interpretation may be invalid and a plausible explanation is instead that the distance-to-default could be "too large" on average when entering as a linear effect on the log-hazard scale in good periods on the stock market. For this case one would expect a positive slope on the market return. We remark that our estimates are not directly comparable to those of Duffie et al. (2007, 2009) as they only (among other things) consider industrial firms, have a different time period, and include events from other databases.

The AICs of $\mathcal{M}_1$ and $\mathcal{M}_2$ show that $\mathcal{M}_2$ is a much better fit. Distance-to-default is not significant in model $\mathcal{M}_2$, which may not be that surprising as we include both the market value to liabilities and the idiosyncratic volatility, which have strong associations with the value and volatility of the underlying asset in the Merton model (we remark on this further in the next subsection). All signs of the coefficients are as expected except for the log current assets. As we show in the table and in Figure 10 in the appendix, we find that larger current assets are associated with a higher default hazard for fixed relative market size. We note that removing the relative market size from $\mathcal{M}_2$ yields a negative (but small in absolute terms) coefficient on the log current asset. Further, none of the variance inflation factors in $\mathcal{M}_2$ are larger than four, so there is no severe multicollinearity.

The differences between columns $\mathcal{M}_2$ and $\mathcal{M}_3$ are three spline terms and an interaction between the current ratio and the idiosyncratic volatility. The latter shows that the partial effect of the current ratio increases when the idiosyncratic volatility increases. This implies a lower partial association with a measure of capability to pay short-term debt (the current ratio) when the firm's equity value is more volatile, since the main effect is negative. This seems plausible.

The three splines in $\mathcal{M}_3$ are shown in Figure 2. The splines are subject to a typical sum-to-zero constraint, and a three-dimensional natural cubic spline basis is used. The left plot shows a tilted "hockey-stick" curve, which flattens for large negative loss to total assets, similar to what is observed in Christoffersen et al. (2018). The partial effect of the idiosyncratic volatility only differs for firms with small to moderate idiosyncratic volatility and flattens thereafter. Due to a relatively small number of events for firms with a moderate or large market value to total liabilities, we log-transformed the ratio to obtain a more well-posed problem. The low number of observed events in the right tail of the ratio is reflected in wider confidence bounds of the spline.

3.1 Distance-to-Default

Distance-to-default has received a lot of attention in the literature. Many noticeable papers cover one or more probability of default models which include distance-to-default in some way (Hillegeist et al. 2004, Duffie et al. 2007, Bharath and Shumway 2008, Campbell et al. 2008, Duffie et al. 2009). The distance-to-default comes from the Merton model and assumes that the underlying firm value follows a geometric Brownian motion and that the firm has issued a single zero-coupon bond. Both assumptions are potentially restrictive. In particular, the ad hoc practice of setting the debt maturing in one year to the short-term debt and 50% of the long-term debt suggests a violation of the latter assumption.

While Duffie et al. (2007, 2009) show that distance-to-default is a strong predictor, Campbell et al. (2008) only find smaller improvements when including distance-to-default in a model that

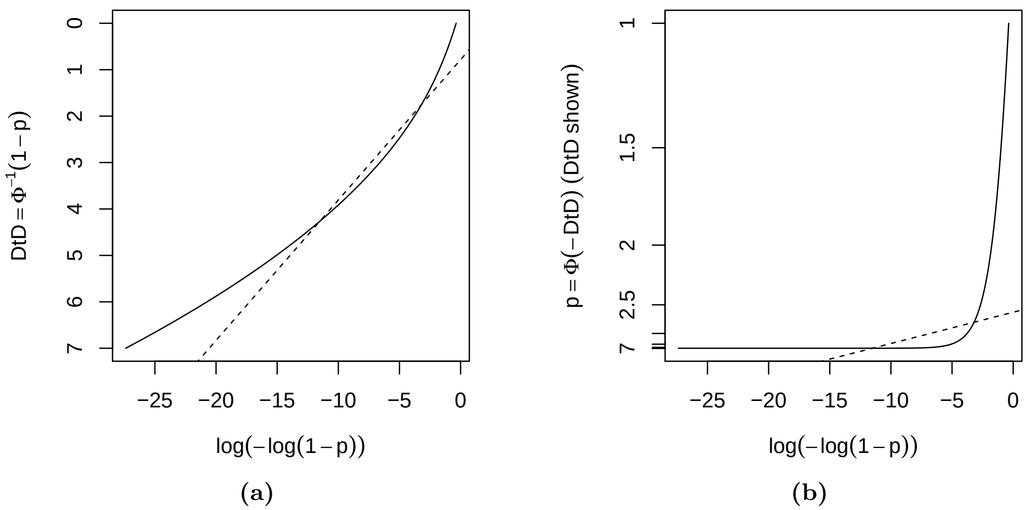


Figure 3: Comparison of output from the Merton model. Plots of (a) inverse standard normal cumulative distribution, Φ^{-1}, versus the complementary log-log function and (b) uniform distribution versus the complementary log-log function both for varying quantiles, p. The lines goes through the 0.00001 and 0.04 quantile to emphasize the region where we may expect to have data. The y-axis is decreasing from north to south to get a positive slope due to the negative relationship between the distance-to-default and the probability of default.

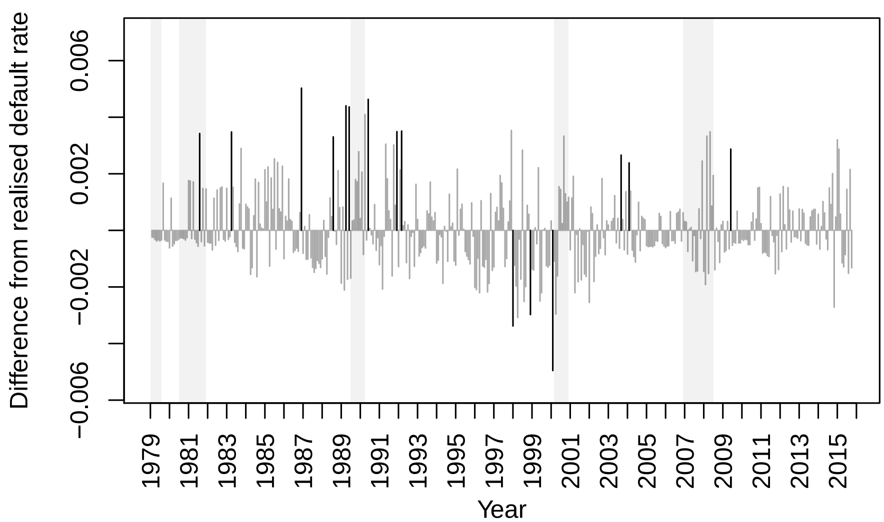


Figure 4: Difference between realized monthly default rate and in-sample predicted default rate of model $\mathcal{M}_3$. Black bars indicate that a 90% point-wise confidence interval does not cover the realized default rate. The scale is large compared to the monthly default rate (see Figure 1). Gray areas are recession periods from National Bureau of Economic Research.

also includes volatility and leverage. The latter is similar to our findings in that we cannot reject a zero slope in $\mathcal{M}_2$. Bharath and Shumway (2008) show that distance-to-default is not sufficient on its own and that a simpler and highly correlated metric performs equally, if not better, in ranking firms by their default risk. However, our results are not directly comparable to Bharath and Shumway (2008), as they include the probability of default from the Merton model on the log-hazard scale whereas we include the distance-to-default as in Duffie et al. (2007, 2009). That is, if DtD_{it} denotes the distance-to-default of firm i at time t and Φ is the standard normal cumulative distribution function, then Bharath and Shumway (2008) assume that the probability of default from the Merton model

$$p_{it} = \Phi^{-1}(-\mathrm{DtD}_{it}) \tag{7}$$

has a linear association on the log-hazard scale while we assume that DtD_{it} has linear association on the log-hazard scale. The central question is what to expect if Equation (7) is approximately true.

Figure 3 illustrates the inverse standard normal cumulative distribution and uniform distribution versus the complementary log-log function (inverse of Equation (3)). The inverse standard normal cumulative distribution and complementary log-log seem to match in the low to rather high probability of default region, suggesting that including the distance-to-default on the log-hazard scale is not unreasonable if Equation (7) is approximately true. This is not the case for the uniform distribution. Moreover, we have tried to replace the distance-to-default in model $\mathcal{M}_1$ with the probability from Equation (7), resulting in an AIC of 5633 which is worse than the original model that included the distance-to-default. Further, performing a likelihood ratio test of the original $\mathcal{M}_1$ against the nested model which also includes the probabilities from Equation (7) yields a p-value of 0.083. In conclusion, we find no arguments or evidence that the probability should be used on the log-hazard scale rather than or simultaneously with distance-to-default.[5]

Finally, we turn to our $\mathcal{M}_3$ model. As in the online appendix of Duffie et al. (2009), we find limited evidence of a non-linear relation on the log-hazard scale with distance-to-default. However, we do find a significant non-linear relationship with two related variables, namely the idiosyncratic volatility and log market value to total liabilities. The regressions we run to estimate the idiosyncratic volatility typically do not fit well so the idiosyncratic volatility is close to the estimated volatility of the equity. The volatility of the equity is related to the volatility of underlying asset in the Merton model by

$$\begin{aligned} d_{it} &= \frac{\log V_{it}/F_{it} + (r + \sigma_{iV}^2/2)}{\sigma_{iV}} \\ \sigma_{iE} &= \frac{V_{it}}{E_{it}}\Phi(d_{it})\sigma_{iV} \end{aligned}$$

where V_{it}, E_{it}, and F_{it} is the value of the underlying asset, the value of the equity, and the value of debt maturing in one year of firm i at time t, respectively, and σ_{iE} and σ_{iV} is the volatility of the equity and underlying value of firm i, respectively. Thus, when d_{it} is large then σ_{iE} is merely a rescaling of σ_{iV}, implying that the idiosyncratic volatility is a good proxy for the underlying volatility. Consequently, adding a spline to the idiosyncratic volatility can be seen as a relaxation of the effect of one of the key components in the Merton model. Secondly, for large E_{it}/F_{it} and low σ_{iV}

$$\log\frac{V_{it}}{F_{it}} \approx \log\frac{E_{it}+F_{it}}{F_{it}} \approx \log\frac{E_{it}}{F_{it}}$$

implying that the log market value to total liabilities is a close proxy for another key component

[5]Though, Bharath and Shumway (2008) remark that they find inferior performance by including the log of the distance-to-default. However, it is unclear to us how negative distance-to-default values are handled. It also seems to be more common to include to the untransformed distance-to-default on the log-hazard or log-odds scale.

in the Merton model. Thus, two of our splines can be seen as a relaxation of the assumptions in the Merton model.

Our choice of covariates is based on previous literature which is reflected in our preference for the idiosyncratic volatility and market value over total liabilities instead of the volatility of the equity and the market value plus total liabilities over total liabilities. All the splines we include are based on standard residual diagnostic plots. In this regard, it is interesting that we find evidence of non-linear effects for two variables that are this closely related to the Merton model. As Bharath and Shumway (2008) conclude, the Merton model seems to provide guidance to default models but we find that relaxations seem to be needed.

To motivate the next section, Figure 4 shows the in-sample difference between the predicted default rate (percentage of firms that experience an event in one month) and the realized monthly rate for $\mathcal{M}_3$, and reveals that the model may have issues with excess clustering of defaults. There is a tendency of co-occurrences of too large or too small predicted monthly default rates, which suggests a frailty model with a temporal dependence as in Equation (4).

3.2 Frailty Models

We will present results for two frailty models in this section. One with a random intercept as in Equation (4) and one where we add a random slope for the relative log market size. Table 2 shows the estimated parameters. Column $\mathcal{M}_4$ shows estimates for the model with splines and the interaction ($\mathcal{M}_3$) with an added frailty effect for the intercept as in Equation (4). The parameter estimates are similar to $\mathcal{M}_3$. The estimated loading θ is close to what Duffie et al. (2009) find. However, the difference Δ in twice the log-likelihood is only 7.0 (the log-Bayes factor mentioned in Duffie et al. 2009). Though, Δ between the model similar to Duffie et al. (2009) ($\mathcal{M}_1$) with and without a frailty term is 14.8, which is closer to the 22.6 reported in Duffie et al. (2009). In conclusion, our finding suggests that the additional variables, non-linear effects, and the interaction capture some of the temporal heterogeneity.

Figure 5 shows the predicted frailty variable, A_k, conditional on the observed data. There are some periods where the predicted value of the frailty-term is 0.2, yielding $\exp(0.2) \approx 1.22$ factor higher instantaneous hazard for all firms at the same point in time. The small difference in log-likelihood is reflected in the wide prediction intervals.

The frailty model in Equation (6) is denoted $\mathcal{M}_5$. We model the temporal dependence between the frailty variables as

$$\begin{pmatrix} A_k \\ B_k \end{pmatrix} = \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix} \begin{pmatrix} A_{k-1} \\ B_{k-1} \end{pmatrix} + \boldsymbol{\epsilon}_k, \qquad \boldsymbol{\epsilon}_k \sim \boldsymbol{\mathcal{N}}\left(\mathbf{0}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}\right) \tag{8}$$

where B_k is the zero-mean term on the slope of the relative log market size. There is a large difference between twice the log-likelihood of model $\mathcal{M}_5$ and $\mathcal{M}_4$ of 60.0, providing strong evidence in favour of the former model.

Figure 6 shows the predicted value of B_k. The predicted value of A_k and B_k are very similar as θ_1 and θ_2 are almost equal and the correlation coefficient, ρ, is high. There is an increase in the relative log market size slope at each of the crisis in 1990s and 2001, implying that large firms tend to be relatively more risky during a crisis all else equal. A similar time-varying effect of the size-variable is shown in Lando et al. (2013), Jensen et al. (2017) for a broad sample of Danish firms, and Filipe et al. (2016) for SMEs in Europe.[6] We remark that the size variable differs between the aforementioned papers and our work. For completeness we tried to use the log of the real value of total assets similar to Filipe et al. (2016) and Jensen et al. (2017) instead of the relative log market size in $\mathcal{M}_2$, but this resulted in a worse fit.

[6] See the log-size coefficient in the robustness check of Filipe et al. (2016).

Table 2: Estimated monthly frailty models. Both models correspond to the model with splines and an interaction, $\mathcal{M}_3$, with additional frailty variables: $\mathcal{M}_4$ has a random intercept and $\mathcal{M}_5$ has a random intercept and slope for the relative log market size. All covariates are centered in the shown models. The χ^2 test statistics from the Wald tests are given in parentheses: *** implies significance at the 1% level, ** at the 5% level, and * at the 10% level. The spline rows are from a three-dimensional natural cubic spline basis, which is restricted to a two-dimensional space that is orthogonal to the linear term. Thus, the linear term can be interpreted as the linear part of the spline. The two coefficients for each basis function in the splines are omitted and a "✓" indicates that the spline term is included in the model. The estimated splines are shown in Figure 11 in the appendix.

	$\mathcal{M}_4$		$\mathcal{M}_5$	
Intercept	-10.507^{***}	(878.2)	-10.789^{***}	(290.2)
Distance-to-default	0.137^{***}	(6.7)	0.107^{**}	(4.0)
Log excess return	-1.734^{***}	(201.8)	-1.673^{***}	(187.0)
T-bill rate	5.671^{**}	(6.3)	-0.665	(0.0)
Log market return	0.945^{***}	(7.2)	0.981^{***}	(9.4)
Log current assets	0.207^{***}	(15.1)	0.249^{***}	(23.2)
Working capital / total assets	-0.981^{*}	(3.4)	-1.096^{**}	(4.2)
Operating income / total assets	-1.073^{*}	(3.8)	-1.108^{**}	(4.1)
Log market value / total liabilities	-0.731^{***}	(9.5)	-0.633^{***}	(7.1)
Net income / total assets	-2.019^{***}	(14.9)	-1.929^{***}	(13.7)
Total liabilities / total assets	0.447^{**}	(3.8)	0.521^{**}	(5.2)
Current ratio	-0.847^{***}	(21.2)	-0.809^{***}	(19.0)
Idiosyncratic volatility	109.290^{***}	(42.1)	126.717^{***}	(48.2)
Relative log market size	-0.340^{***}	(29.1)	-0.415^{***}	(7.1)
Net income / total assets (spline)	✓*	(5.9)	✓*	(5.2)
Idiosyncratic volatility (spline)	✓***	(28.3)	✓***	(34.8)
Log market value / total liabilities (spline)	✓***	(17.3)	✓***	(18.1)
Current ratio · Idiosyncratic volatility	16.074^{***}	(21.2)	15.108^{***}	(18.6)
θ_1	0.897		0.972	
θ_2			0.974	
σ_1	0.116		0.269	
σ_2			0.064	
ρ			0.991	
AIC	5001.4		4947.4	
log-likelihood	-2477.7		-2447.7	
Number of firms	3020		3020	

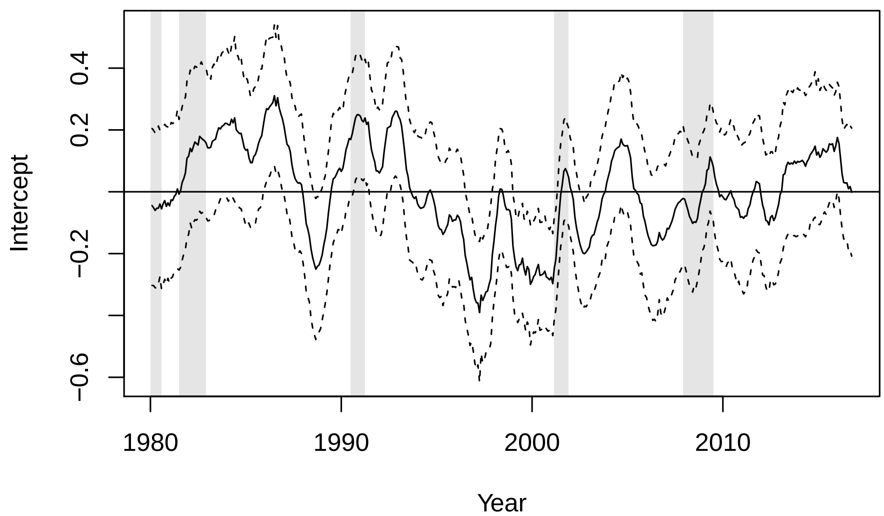


Figure 5: Predicted (smoothed) frailty variable A_k of $\mathcal{M}_4$ conditional on the observed data. The dashed lines are 68.3% point-wise prediction intervals (similar to the ±1 standard deviation in Duffie et al. 2009). Gray areas are recession periods from National Bureau of Economic Research.

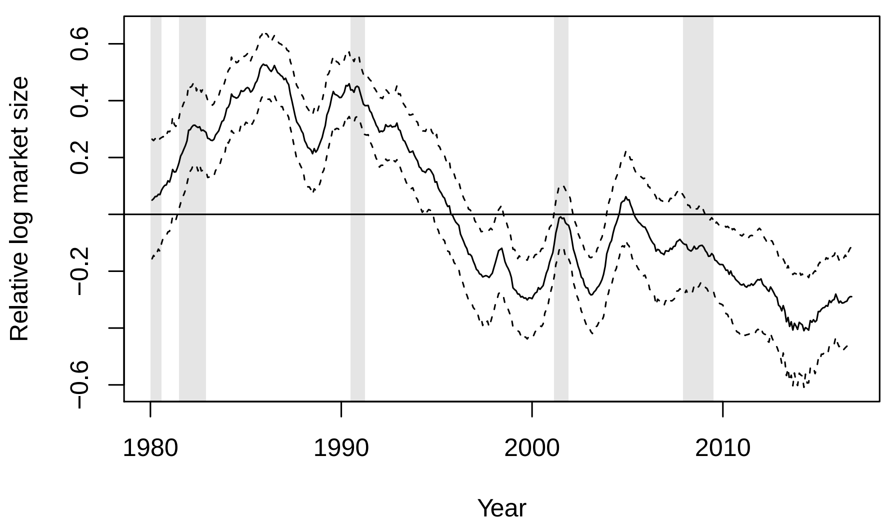


Figure 6: Predicted (smoothed) frailty variable B_k of $\mathcal{M}_5$ conditional on the observed data. The dashed lines are 68.3% point-wise prediction intervals (similar to the ±1 standard deviation in Duffie et al. 2009). A_k is very similar and roughly $0.264/0.064 \approx 4.1$ times greater in magnitude due to the very similar decay-rate (θ) estimate and high correlation between the two random effects. The plot does not include the fixed slope estimate, i.e., -0.415 needs to be added to get the slope on relative log market size at any point in time. Gray areas are recession periods from National Bureau of Economic Research.

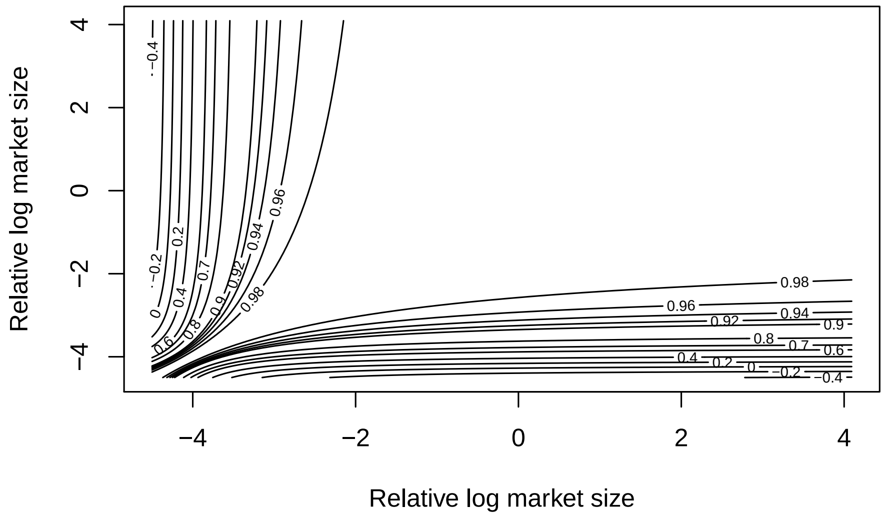


Figure 7: Correlation in innovation term for different values of relative log market size. Contours of correlation between $\epsilon_{1t} + \epsilon_{2t} z_{it}$ for different values of the log relative market size, z_{it}. The log relative market size is centered as in the $\mathcal{M}_5$ model.

We note that the slope of the risk-free rate is virtually zero in model $\mathcal{M}_5$, while all other coefficients are similar to those of model $\mathcal{M}_4$. This smaller association with macro-variables is also observed in Lando et al. (2013).

Lastly, we consider the estimated random effect. Each firm's random effect component on the log-hazard scale is given by

$$A_t + B_t z_{it} = \epsilon_{1t} + \epsilon_{2t} z_{it} + \theta_1 A_{t-1} + \theta_2 B_{t-1} z_{it}$$

where z_{it} is the relative log market size. Figure 7 shows the correlation between the $\epsilon_{1t} + \epsilon_{2t} z_{it}$ term for two values of z_{it}. The conclusion is that firms of equal size have a highly correlated random effect term, while low-valued firms tend to have a lower correlation with the random effect term of moderately- and high-valued firms. This is in contrast to the model $\mathcal{M}_4$ where the random effect term for all firms is perfectly correlated by construction.

3.3 Comparison with Other Work

Omitting covariates with co-movements through time for all firms or groups of firms can yield evidence of a single shared frailty (i.e., a time-varying intercept). An example of a covariate with a co-movement through time is shown in Figure 8, showing density estimates of the winsorized idiosyncratic volatility through time. Many of the covariates in our model have time-varying distributions, which may explain the smaller Bayes factor we observe between a model with a time-varying intercept versus a model without in comparison to Duffie et al. (2009). The results presented by Lando and Nielsen (2010) are yet another indication of this effect, since they fail to reject the misspecification test in Das et al. (2007) after including additional covariates.

A different approach to default modeling is to consider aggregate defaults. Some examples are Koopman et al. (2011, 2012), Schwaab et al. (2017), Azizpour et al. (2018) who aggregate to different levels, which are either total default counts or default counts in rating and industry (and region or age cohort) groups. Koopman et al. (2011), Schwaab et al. (2017), Azizpour et al. (2018)

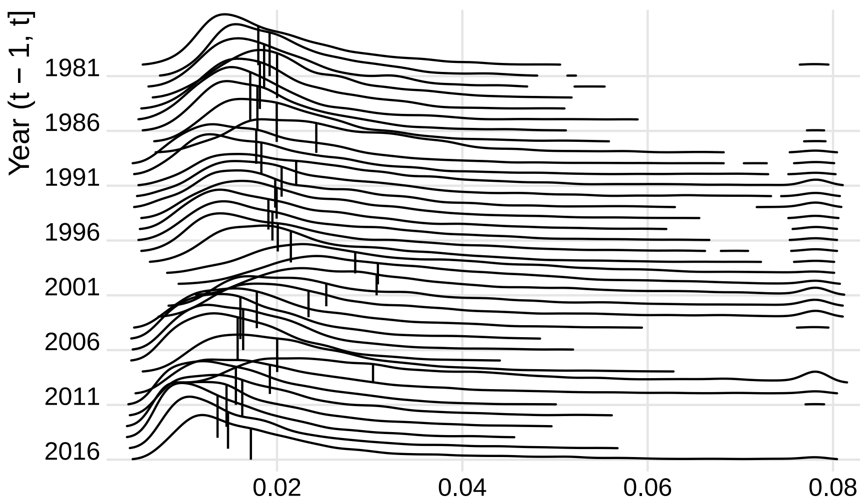


Figure 8: Density estimates of idiosyncratic volatility. The plot shows density estimates of the winsorized idiosyncratic volatility through time. The middle vertical lines are the medians. Densities lower than 1% of the maximum are omitted. A Gaussian kernel is used with a bandwidth of 0.00137.

use cross-sectional averages or medians of firm-level covariates and either explicitly or implicitly assume that firms are homogeneous given these variables at the level of default that they model. This is a strict assumption and may not be valid in practice when covariates have time-varying distributions as in Figure 8. Time-varying distributions of firm-level covariates can yield evidence of a macro effect, frailty variable, or contagion variable as the following example will show. We have 1000 firms which we observe over 65 periods. Each firm has a randomly distributed incorporations date which is uniformly distributed on $\{0, \dots, 64\}$ and a single time-varying covariate X_{it} which is drawn from the mixture given by

$$X_{it} \mid U_{it} = u \sim \begin{cases} \mathcal{N}(0,1) & u = 0 \\ \mathcal{N}(g(t),1) & u = 1 \end{cases}, \qquad g(t) = 4 \cdot \left| \frac{t-33}{32} \right| - 1$$

where U_{it} is Bernoulli distributed with probability 0.2. That is, the covariate is either drawn from a time-invariant distribution or from a distribution with a time-varying mean. Defaults are terminal, we observe defaults in discrete time, and define the default intensity as

$$\lambda_{it} = \exp(\beta_0 + \beta_1 x_{it})$$

where $(\beta_0, \beta_1) = (-4, 1)$. We perform 1000 simulations and fit a model for aggregate defaults for each simulated data set using the mean at time t, $\bar{X}_t = |R_t|^{-1} \sum_{i:R_t} X_{it}$, as a covariate where R_t is the risk set at time t. Furthermore, we fit a second model where we include time as a second order polynomial. The latter model can be seen as a model which includes either macro-variables, a contagion factor, or a shared frailty variable. We reject the likelihood ratio test between the two nested models in 618 of the 1000 simulations at the conventional 5% level, which could be interpreted as evidence of a macro, a contagion, or a frailty effect when employing cross-sectional averages of firm covariates. More importantly, both models have the undesirable feature that they miss substantial cross-sectional variation, yielding incorrect results for a corporate debt portfolio with a non-random sample of the population. For completeness, Koopman et al. (2012), Schwaab et al. (2017) do state that a firm's rating may not be sufficient statistics for default. Further, there is evidence that ratings alone may be poor proxies of risk (e.g., see Hamilton and Cantor 2004).

Our results in Figure 6 suggest that larger firms are partially more risky in some periods than others, whereas Azizpour et al. (2018) show that periods with large amounts of defaulted debt are followed by a higher aggregate default rate. The advantage of our model in this context is that it can be applied to an arbitrary portfolio of firms and can distinguish between an overall change in default rates and a change in default rates for a subset of firms. The latter is key, e.g., for regulators that want to evaluate the risk exposure of banks that provide loans to a subset of firms that are not a random subset of the entire population.

We remark that we do not attempt to infer causal effects. The observed frailty effect may be either "true" frailty (i.e., temporary shocks that affect all or groups of firms), a proxy for a contagion effect (i.e., the default of one firm spreads to other firms), causal associations or non-causal associations. However, we provide a model which is an accurate firm-level as well as joint default model. Such models are needed to perform bottom-up modeling of the default risk of a corporate debt portfolio. The model can easily be extended to relax the assumption that other coefficients are constant and exploit information of all defaults through time.

4 Out-of-sample

In this section, we will test the performance of the models out-of-sample through time. Our goal is to test how well the models perform on the firm-level and aggregate level. The former is important as we want to be able to use the models on a portfolio which contains an arbitrary subset of the firms in our sample. The latter is important as any bias at a point in time can affect the overall predicted default rate of a portfolio and subsequent modeling of other quantities using the default model.

As described in Duffie et al. (2009), our models are so-called doubly stochastic Poisson processes conditional on the frailty variables. That is, conditional on the covariates (and frailty variable path), we have piecewise exponentially distributed arrival times. However, we do not know the future covariates when we forecast apart from the value in the present month. We also do not know whether, and if so when, the firm will exit the sample due to other reasons than a default. What we will do is take both covariates and exits as given, i.e., in our risk set we include the firms up to and including the month where they exit due to a default or for other reasons. During this period we treat the firm as if it can default unconditional on that the firm will exit within our forecasting period. E.g., if a firm exits at the end of month 4 (due to exit or a default), then the firm is at risk for 4 months. This allows us to solely test our default models' performance and not how well we can model the covariates in our model.[7]

We estimate $\mathcal{M}_1$ and $\mathcal{M}_3$-$\mathcal{M}_5$ up to January or July of a given year and then use the estimated models to forecast defaults for the following half-year. We quantify the model performance in two ways: First, we use the area under the receiver operator characteristic curve (AUC). It is a commonly used metric in the default literature and the AUC allows us to assess the firm-level performance. The interpretation of AUC is the fraction of correctly ordered firms in terms of whether or not they default within the six month period. Thus, it is the probability that a random firm in our sample with a default has a higher hazard than a random firm without a default. A value of 0.5 is random guessing and a value of 1 means that all firms with a default have a higher

[7] Evidence presented in Duan et al. (2012) suggests that a first order vector autoregression as in Duffie et al. (2007) may be inappropriate when modeling covariates. As remarked, modeling the high dimensional and non-fixed size set of covariate vectors (varying since the set of firms at-risk changes through time) is interesting, but not what we pursue in this paper.

Table 3: Number of breaches of the upper bounds of the prediction intervals in Figure 9(b). The first row shows the number of strictly greater realized default rates and the second row shows greater than or equal. The large difference between the two rows is due to low realized and expected default rates in some periods. The last row shows the number of half-years in our out-of-sample test.

	$\mathcal{M}_1$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$
# breaches (>)	3	3	1	0
# breaches (≥)	7	7	5	4
# periods	36	36	36	36

hazard than firms without a default. We compute the AUC of each model using the mean hazard rates, which follow from the predicted default probabilities.

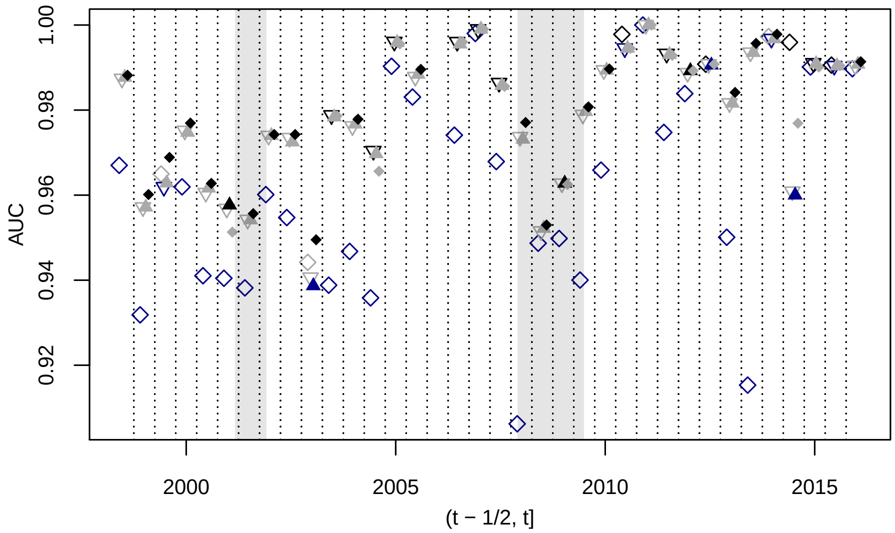
Secondly, we simulate the events conditional on the predicted default probabilities for each firm using the models without frailty and use these to compute the industry-wide default rate for the following half-year. Repeating this multiple times gives us a distribution for the predicted default rate, and a correctly specified model should have decent coverage of the prediction interval of this rate. To assess this we use that the upper bound of the prediction interval is similar to a 95% Value-at-Risk but for the whole industry's default rate. For the frailty models, we first have to sample a (A_k, B_k)-pair from the so-called particle cloud (see Appendix A) of the last month of the estimation period, simulate $(A_{k+1}, B_{k+1}, \ldots, A_{k+6}, B_{k+6})$ conditional on the sampled (A_k, B_k)-pair, and then simulate the outcomes as we do for the models without frailty conditional on the sampled path of the frailty variables. We do so as we need the entire paths of the frailty variables over the following six months to forecast the likelihood of an event for a given firm. We emphasize that the Monte Carlo EM algorithm, particle filter, and particle smoother only use the data available up to and including the last month of the estimation period (time k) when estimating parameters and forming the particle cloud.

The out-of-sample results are shown in Figure 9. The AUC in Figure 9(a) shows that the model like in Duffie et al. (2007) ($\mathcal{M}_1$) performs poorly in terms of firm-level performance relative to the other models. The difference in AUC sometimes exceeds 0.05 compared to the other models, which means that the latter models have more than a 5% higher fraction of correctly ordered firms by whether they default or not. This is substantial. The second conclusion is that our final frailty model ($\mathcal{M}_5$), which allows for a time-varying slope of the relative log market size, does best 19 of the 35 periods.

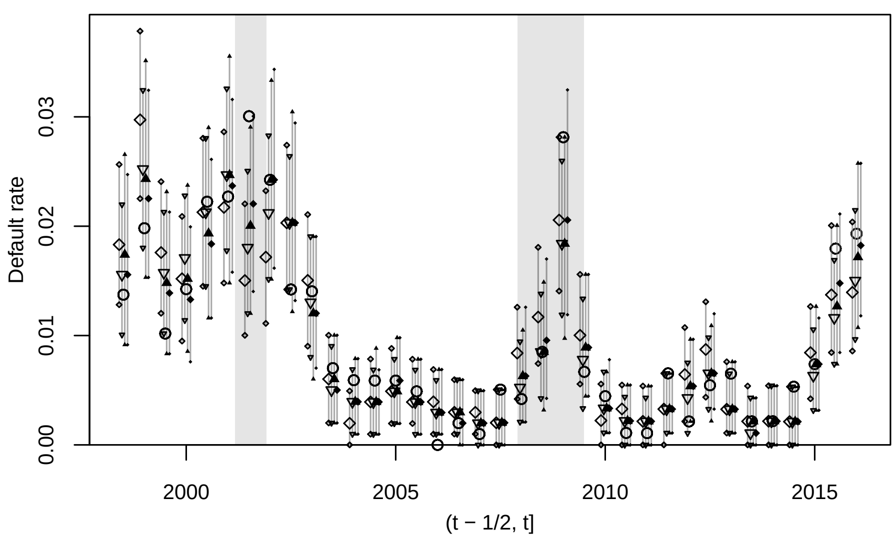
Next, we turn to the aggregate level performance. The industry-wide predicted default rates show that $\mathcal{M}_5$ is better in some periods in the sense that the median of the predicted rate is closer to the realized, and in particular we notice the crisis in 2009 and last period from January to June in 2016. However, it is not always true that the median of $\mathcal{M}_5$ is closest to the realized default rate. Table 3 shows the frequency of breaches of the upper bound in the prediction interval for the industry-wide default rate, and from these figures we observe that both of the two random effect models, $\mathcal{M}_4$ and $\mathcal{M}_5$, have coverage close to the 95% coverage level.

5 Conclusion

We have extended the hazard model of Duffie et al. (2007) by including additional covariates, a non-linear effect for the idiosyncratic volatility, net income relative to total assets of the firm, and log market value over total liabilities, and by adding an interaction between the idiosyncratic volatility



(a)



(b)

Figure 9: Out-of-sample performance. Figure (a) shows the out-of-sample AUC where $\diamond$ is model $\mathcal{M}_1$, $\triangledown$ is $\mathcal{M}_3$, $\blacktriangle$ is $\mathcal{M}_4$, and $\blacklozenge$ is $\mathcal{M}_5$. The points are the values over the past half-year. The vertical bars are used to separate different time periods, the model with the highest AUC is in black, the model with the lowest AUC is in blue, and the rest are gray. Results for the half-year $t = 2006$ are absent as the half-year contains no defaults. Plot (b) shows the out-of-sample default rate along with 90% prediction intervals, where the upper bound is like a 95% Value-at-Risk but for the industry default rate. The symbols in the middle denote the median value and $\circ$ denotes the realized default rates. Gray areas are recession periods from National Bureau of Economic Research.

Table 4: Summary statistics for the covariates used in the monthly hazard models in Section 3. The right-most columns show the 1% and 99% quantiles, and means and standard deviations are computed after winsorizing. The current assets are deflated by the US Consumer Price Index.

	Mean	Median	Standard deviation	1%	99%
Distance-to-default	5.834	5.294	4.144	-1.631	18.756
Log excess return	-0.057	-0.020	0.400	-1.516	0.964
Working capital / total assets	0.153	0.129	0.164	-0.168	0.607
Operating income / total assets	0.090	0.087	0.088	-0.220	0.362
Market value / total liabilities	1.657	1.035	1.927	0.012	11.686
Net income / total assets	0.032	0.042	0.103	-0.523	0.278
Total liabilities / total assets	0.631	0.615	0.189	0.237	1.358
Current ratio	1.836	1.602	1.042	0.421	6.432
Log current assets	5.117	5.076	1.546	1.686	9.026
Idiosyncratic volatility	0.023	0.020	0.013	0.007	0.078
Relative log market size	-8.945	-8.908	1.792	-13.445	-4.852

and the current ratio. This yields much better out-of-sample ranking of firms by their riskiness. Despite these additions, the model still has issues with excess default clustering although we observe less evidence for excess clustering with a random effect as in Duffie et al. (2009) compared to the model without our additions. We show that this clustering cannot be modeled by adding a single frailty effect affecting all firms equally on the log-hazard scale, as otherwise argued by Duffie et al. (2009).

Instead, we add a time-varying random slope to the relative log market size of the firm, similar to the size effect in Lando et al. (2013). However, unlike the semi-parametric and non-parametric models used by Lando et al. (2013), our model is an extension of previous frailty models and thus it can be used for forecasting. We show that our frailty model fits much better in-sample and, furthermore, our out-of-sample test shows superior ranking of firms by riskiness.

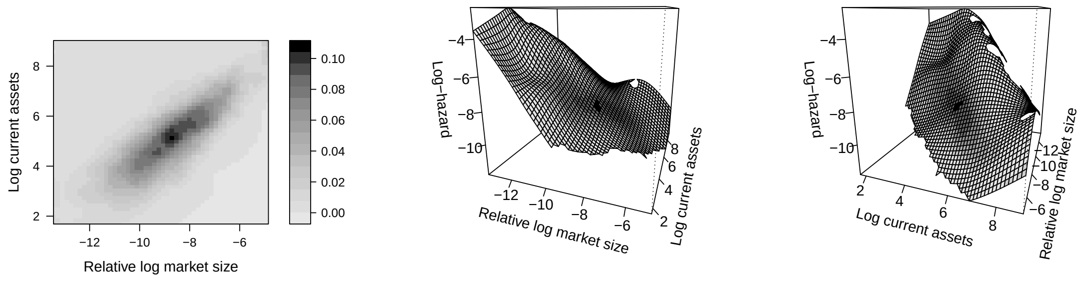
We also present evidence for two non-linear effects of variables that are closely related to the Merton model. This finding corroborates the conclusion of Bharath and Shumway (2008) that the Merton model provides useful guidance for building default models but it is not sufficient.

We remark that our list of covariates may not be complete, and this may also be true for the covariates we model with non-linear association on the log-hazard scale, included interactions and frailty variables, and the assumed distribution of frailty variables. Despite this, our study highlights that the traditional assumption of linearity on the log-hazard scale, the assumption of no interactions, and the assumption of constant slopes within corporate default modeling are too strict in our sample. With this work we show how to easily relax these assumptions in the presented models, and the software we have developed is readily available for practitioners. See the appendix for details.

Time-varying size effects like the one we show have been observed in other data sets by Jensen et al. (2017), Filipe et al. (2016), but it is yet to be determined if this is a more general effect within corporate default models.

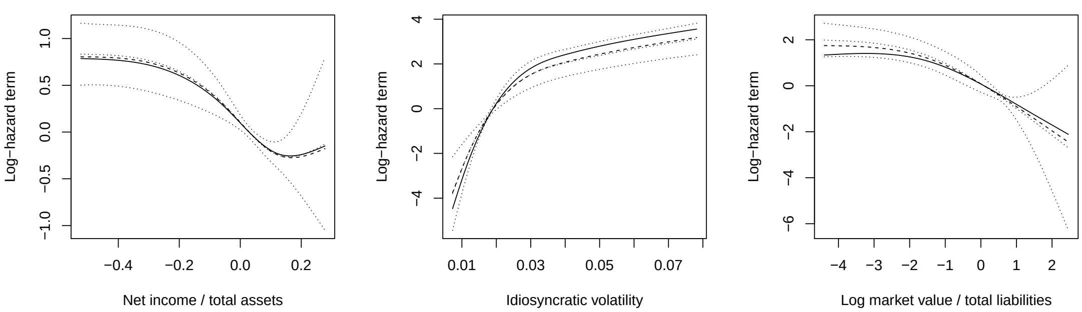
Acknowledgement

The authors are grateful to Mads Stenbo Nielsen, Søren Feodor Nielsen, and seminar participants at Copenhagen Business School for helpful comments.



(a) **(b)** **(c)**

Figure 10: Marginal effect of relative log market size and log current assets. Figure (a) shows the density of the relative log market size and log current assets. It shows that the two are correlated and the majority of the observations are along the diagonal. The density plot also shows that the two are not nearly perfectly correlated. Figures (b) and (c) show a tensor product spline from a generalized additive model with only the two variables included. The z-axis shows the log-hazard rate of default. The integral of the estimated density over the plotted surface is 99% and the colors of Figure (a) are added to the surface. We observe that along the diagonal in the xy-plane (the area where we have data) the log-hazards are higher for firms with higher current assets. The penalty parameter in the generalized additive model is chosen with an un-biased risk estimator criterion.



(a) **(b)** **(c)**

Figure 11: Estimated splines for the monthly frailty models. The inner dotted lines are from the model without frailty, $\mathcal{M}_3$, the dashed lines are from the model with a random intercept, $\mathcal{M}_4$, and the solid lines are from the model with a random intercept and random relative log market size slope, $\mathcal{M}_5$. The splines are for the following covariates: (a) net income to total assets, (b) the idiosyncratic volatility, and (c) log market value to total liabilities. The y-axis is the effect of the covariate on the log-hazard scale and outer dotted lines are 95% confidence intervals for the model without frailty.

A Estimating Frailty Models

We will describe how the frailty models are estimated in this section. To do this, we start by defining the log-likelihood term for each firm. We will focus here on the continuous case where we observe the exact event times and the most general frailty model which is shown in Equation (6). Let y_{ik} be one if firm i has an event in time period $(k-1, k]$ and zero otherwise. Furthermore, let Δt_{ik} be the time at risk of firm i in interval $(k-1, k]$, i.e., this will be 1 in the monthly hazard models if firm i does not have an event in the interval and otherwise the time until the event, $T_i - k + 1$. Then the log-likelihood terms from firm i conditional on the frailty variables are

$$l_i\left(A_{0:d}, \boldsymbol{B}_{0:d}\right) = \sum_{k=1}^{d} \sum_{i:\, R_{ik}=1} y_{ik} \log \lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k, A_k, \boldsymbol{B}_k, \boldsymbol{z}_{ik}\right) - \lambda_{ik}\left(\boldsymbol{x}_{ik}, \boldsymbol{m}_k, A_k, \boldsymbol{B}_k, \boldsymbol{z}_{ik}\right) \Delta t_{ik}$$

where we observe d periods, $A_{0:d} = (A_0, A_1, \ldots, A_d)$, and $\boldsymbol{B}_{0:d} = (\boldsymbol{B}_0, \boldsymbol{B}_1, \ldots, \boldsymbol{B}_d)$. The complete data log-likelihood where we observe the frailty variables is

$$\mathcal{L}\left(\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{F}, \mathbf{Q}\right) = \phi\left(\begin{pmatrix} A_0 \\ \boldsymbol{B}_0 \end{pmatrix}; \mathbf{0}, \mathbf{Q}_0\right) + \sum_{k=1}^{d} \phi\left(\begin{pmatrix} A_k \\ \boldsymbol{B}_k \end{pmatrix}; \mathbf{F}\begin{pmatrix} A_{k-1} \\ \boldsymbol{B}_{k-1} \end{pmatrix}, \mathbf{Q}\right) + \sum_{i=1}^{n} l_i\left(A_{0:d}, \boldsymbol{B}_{0:d}\right) \quad (9)$$

where we have n firms, $\phi\left(\cdot; \boldsymbol{v}, \mathbf{V}\right)$ is the multivariate normal distribution density function with mean $\boldsymbol{v}$ and covariance matrix $\mathbf{V}$, and $\mathbf{Q}_0$ is the time-invariant covariance matrix which is given by $\mathbf{Q}_0 = \mathbf{F}\mathbf{Q}_0\mathbf{F}^\top + \mathbf{Q}$. Direct maximization would require that we integrate $A_{0:d}$ and $\boldsymbol{B}_{0:d}$ out of Equation (9) which is infeasible. An alternative is to employ an expectation maximization (EM) algorithm (Dempster et al. 1977). This is done by starting at some value $\boldsymbol{\theta}^{(0)} = \left(\alpha^{(0)}, \boldsymbol{\beta}^{(0)}, \boldsymbol{\gamma}^{(0)}, \mathbf{F}^{(0)}, \mathbf{Q}^{(0)}\right)$ and then computing

$$H\left(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(0)}\right) = E\left(\mathcal{L}\left(\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{F}, \mathbf{Q}\right) \mid \boldsymbol{y}_{1:d}, \boldsymbol{\theta}^{(0)}\right) \quad (10)$$

where $\boldsymbol{y}_{1:d} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_d)$, $\boldsymbol{y}_k = (y_{1k}, y_{2k}, \ldots y_{nk})$, and the expectation is w.r.t. $A_{0:d}$ and $\boldsymbol{B}_{0:d}$. This is referred to as the E-step. Then, we find a new set of parameters by

$$\boldsymbol{\theta}^{(1)} = \arg\max_{\boldsymbol{\theta}} H\left(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(0)}\right)$$

which is referred to as the M-step. The process is then repeated with $\boldsymbol{\theta}^{(1)}$ in place of $\boldsymbol{\theta}^{(0)}$ until a convergence criterion is reached.

In our case, Equation (10) has no closed-form solution. Instead, we use a Monte Carlo expectation maximization algorithm where the E-step is approximated by the particle smoother suggested by Fearnhead et al. (2010). A multivariate t-distribution approximation at a mode is used at each step of the algorithm as described by Pitt and Shephard (1999). The particle smoother uses an auxiliary particle filter which is also used to get the log-likelihood approximations shown in Table 2. The auxiliary particle filter also yields a discrete approximation of the density of $(A_d, \boldsymbol{B}_d)$ given the observed data. The approximation is a so-called particle cloud consisting of K $(A_d, \boldsymbol{B}_d)$-pairs where each pair has a weight of its conditional probability relative to the other pairs in the cloud. The Wald tests are computed with an approximation of the observed information matrix obtained with the method suggest by Poyiadjis et al. (2011) with a particle filter as in Lin et al. (2005). The R (R Core Team 2018) package `dynamichazard` (Christoffersen 2019) contains implementations of the methods described above and the "Particle filters in the dynamichazard package" vignette in the package covers the methods in more details.

References

Azizpour, S., Giesecke, K., Schwenkler, G., 2018. Exploring the sources of default clustering. Journal of Financial Economics 129, 154 – 183. URL: http://www.sciencedirect.com/science/article/pii/S0304405X1830103X, doi:https://doi.org/10.1016/j.jfineco.2018.04.008.

Berg, D., 2007. Bankruptcy prediction by generalized additive models. Applied Stochastic Models in Business and Industry 23, 129–143.

Bharath, S.T., Shumway, T., 2008. Forecasting Default with the Merton Distance to Default Model. The Review of Financial Studies 21, 1339–1369. URL: https://dx.doi.org/10.1093/rfs/hhn044, doi:10.1093/rfs/hhn044.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. The Journal of Finance 63, 2899–2939.

Chava, S., Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. Review of Finance 8, 537–569.

Chava, S., Stefanescu, C., Turnbull, S., 2011. Modeling the loss distribution. Management Science 57, 1267–1287. URL: https://doi.org/10.1287/mnsc.1110.1345, doi:10.1287/mnsc.1110.1345.

Chen, P., Wu, C., 2014. Default prediction with dynamic sectoral and macroeconomic frailties. Journal of Banking & Finance 40, 211 – 226. URL: http://www.sciencedirect.com/science/article/pii/S0378426613004573, doi:https://doi.org/10.1016/j.jbankfin.2013.11.036.

Christoffersen, B., 2019. dynamichazard: Dynamic Hazard Models using State Space Models. URL: https://github.com/boennecd/dynamichazard. r package version 0.6.4.

Christoffersen, B., Matin, R., Mølgaard, P., 2018. Can machine learning models capture correlations in corporate distresses? Available at SSRN: https://ssrn.com/abstract=3273985.

Das, S.R., Duffie, D., Kapadia, N., Saita, L., 2007. Common failings: How corporate defaults are correlated. The Journal of Finance 62, 93–117. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2007.01202.x, doi:10.1111/j.1540-6261.2007.01202.x.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39, 1–38. URL: http://www.jstor.org/stable/2984875.

Duan, J.C., Fulop, A., 2013. Multiperiod corporate default prediction with the partially-conditioned forward intensity. Working paper.

Duan, J.C., Sun, J., Wang, T., 2012. Multiperiod corporate default prediction – A forward intensity approach. Journal of Econometrics 170, 191 – 209. URL: http://www.sciencedirect.com/science/article/pii/S0304407612001145, doi:https://doi.org/10.1016/j.jeconom.2012.05.002.

Duffie, D., Eckner, A., Horel, G., Saita, L., 2009. Frailty correlated default. The Journal of Finance 64, 2089–2123.

Duffie, D., Saita, L., Wang, K., 2007. Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics 83, 635 – 665. URL: http://www.sciencedirect.com/science/article/pii/S0304405X06002029, doi:https://doi.org/10.1016/j.jfineco.2005.10.011.

Fearnhead, P., Wyncoll, D., Tawn, J., 2010. A sequential smoothing algorithm with linear computational cost. Biometrika 97, 447–464. URL: https://dx.doi.org/10.1093/biomet/asq013, doi:10.1093/biomet/asq013.

Filipe, S.F., Grammatikos, T., Michala, D., 2016. Forecasting distress in European SME portfolios. Journal of Banking & Finance 64, 112 – 135. URL: http://www.sciencedirect.com/science/article/pii/S0378426615003611, doi:https://doi.org/10.1016/j.jbankfin.2015.12.007.

Hamilton, D.T., Cantor, R., 2004. Rating transition and default rates conditioned on outlooks. The Journal of Fixed Income 14, 54–70. URL: https://jfi.iijournals.com/content/14/2/54, doi:10.3905/jfi.2004.439837.

Hillegeist, S.A., Keating, E.K., Cram, D.P., Lundstedt, K.G., 2004. Assessing the probability of bankruptcy. Review of Accounting Studies 9, 5–34. URL: https://doi.org/10.1023/B:RAST.0000013627.90884.b7, doi:10.1023/B:RAST.0000013627.90884.b7.

Jensen, T., Lando, D., Medhat, M., 2017. Cyclicality and firm size in private firm defaults. International Journal of Central Banking 13, 97–145.

Koopman, S.J., Lucas, A., Schwaab, B., 2011. Modeling frailty-correlated defaults using many macroeconomic covariates. Journal of Econometrics 162, 312–325.

Koopman, S.J., Lucas, A., Schwaab, B., 2012. Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: The credit crisis of 2008. Journal of Business & Economic Statistics 30, 521–532. URL: https://doi.org/10.1080/07350015.2012.700859, doi:10.1080/07350015.2012.700859.

Lando, D., Medhat, M., Nielsen, M.S., Nielsen, S.F., 2013. Additive intensity regression models in corporate default analysis. Journal of Financial Econometrics 11, 443–485.

Lando, D., Nielsen, M.S., 2010. Correlation in corporate defaults: Contagion or conditional independence? Journal of Financial Intermediation 19, 355 – 372. URL: http://www.sciencedirect.com/science/article/pii/S1042957310000070, doi:https://doi.org/10.1016/j.jfi.2010.03.002. risk Transfer Mechanisms and Financial Stability.

Lin, M.T., Zhang, J.L., Cheng, Q., Chen, R., 2005. Independent particle filters. Journal of the American Statistical Association 100, 1412–1421. URL: http://www.jstor.org/stable/27590681.

Pitt, M.K., Shephard, N., 1999. Filtering via simulation: Auxiliary particle filters. Journal of the American Statistical Association 94, 590–599. URL: http://www.jstor.org/stable/2670179.

Poyiadjis, G., Doucet, A., Singh, S.S., 2011. Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. Biometrika 98, 65–80. URL: http://www.jstor.org/stable/29777165.

Qi, M., Zhang, X., Zhao, X., 2014. Unobserved systematic risk factor and default prediction. Journal of Banking & Finance 49, 216 – 227. URL: http://www.sciencedirect.com/science/article/pii/S0378426614003094, doi:https://doi.org/10.1016/j.jbankfin.2014.09.009.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Schwaab, B., Koopman, S.J., Lucas, A., 2017. Global credit risk: World, country and industry factors. Journal of Applied Econometrics 32, 296–317. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2521, doi:10.1002/jae.2521.

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model. The Journal of Business 74, 101–124.

Vassalou, M., Xing, Y., 2004. Default risk in equity returns. The Journal of Finance 59, 831–868. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2004.00650.x, doi:10.1111/j.1540-6261.2004.00650.x.

DANMARKS NATIONALBANK
HAVNEGADE 5
DK-1093 COPENHAGEN K
WWW.NATIONALBANKEN.DK

DANMARKS NATIONALBANK