

Multi-Period Corporate Failure Prediction With Stochastic Covariates*

Darrell Duffie[†] Ke Wang[‡]

First Version: August 30, 2003
Current Version: December 17, 2003

Abstract

We provide maximum likelihood estimators of term structures of conditional probabilities of bankruptcy over relatively long time horizons, incorporating the dynamics of firm-specific and macroeconomic covariates. We find evidence in the U.S. industrial machinery and instruments sector, based on over 28,000 firm-quarters of data spanning 1971 to 2001, of significant dependence of the level and shape of the term structure of conditional future bankruptcy probabilities on a firm's distance to default (a volatility-adjusted measure of leverage) and on U.S. personal income growth, among other covariates. Variation in a firm's distance to default has a greater relative effect on the term structure of future failure hazard rates than does a comparatively sized change in U.S. personal income growth, especially at dates more than a year into the future.

*We are grateful for support from Moody's, and for comments from Susan Athey, Richard Cantor, Brad Efron, Jeremy Fox, Peyron Law, Aprajit Mahajan, and especially Takeshi Amemiya.

[†]Graduate School of Business, Stanford University, Stanford, CA 94305-5015, email: duffie@stanford.edu, webpage: www.stanford.edu/~duffie/

[‡]Department of Economics, Stanford University, Stanford CA 94305. email: kewang@stanford.edu, webpage: www.stanford.edu/~kewang/ Wang gratefully acknowledges support under a grant from Moody's.

1 Introduction

We provide maximum likelihood estimators of term structures of conditional corporate bankruptcy probabilities. Our contribution over prior work is to exploit the dependence of failure intensities on stochastic covariates, as well as the time-series dynamics of the covariates, in order to estimate the likelihood of failure over several future periods (quarters or years). We estimate our model for the U.S. industrial machinery and instrument sector, using over 28,000 firm-quarters of data for the period 1971 to 2001. We find evidence of significant dependence of the level and shape of the term structure of conditional future failure probabilities on a firm's distance to default (a volatility-adjusted measure of leverage) and on U.S. personal income growth, among other covariates. Variation in a firm's distance to default has a greater relative effect on the term structure of future failure hazard rates than does a comparatively sized change in the business-cycle covariate, U.S. personal income growth, especially at dates more than one year into the future.

The estimated shape of the term structure of conditional failure probabilities reflects the time-series behavior of the covariates, namely leverage targeting by firms and mean reversion in macro-economic performance. The term structures of failure hazard rates are typically upward sloping at business-cycle peaks, and downward sloping at business-cycle troughs, to a degree that depends on corporate leverage relative to its long-run target.

In our model, a firm's stochastic failure intensity is assumed to depend on both firm-specific and macroeconomic state variables. Stochastic evolution of the combined Markov state vector X_t causes variation over time in a firm's failure intensity $\lambda_t = \Lambda(X_t)$. The firm exits for other reasons, such as merger, acquisition, privatization, with an intensity $\alpha_t = A(X_t)$. The total exit intensity is $\alpha_t + \lambda_t$.

We specify a doubly-stochastic formulation of the point process for failure and other forms of exit under which the conditional probability at time t of corporate survival (from failure or other exit) for s years is

$$p(X_t, s) = E \left(e^{-\int_t^{t+s} (\lambda(u) + \alpha(u)) du} \mid X_t \right), \quad (1)$$

and under which the conditional probability of failure within s years is

$$q(X_t, s) = E \left(\int_t^{t+s} e^{-\int_t^z (\lambda(u) + \alpha(u)) du} \lambda_z dz \mid X_t \right). \quad (2)$$

This calculation of $q(X_t, s)$, demonstrated in Section 2, reflects the fact that, in order to fail at time z , the firm must survive until time z , avoiding both failure and other forms of exit, which arrive at a total intensity of $\lambda(u) + \alpha(u)$.

While, as explained in Section 1.1, there is a significant prior literature treating the estimation of one-period-ahead bankruptcy probabilities, for example with duration models, we believe this is the first empirical study of the conditional term structure of failure probabilities over multiple future time periods. The sole exception seems to be the practice of certain banks and dealers in structured credit products of treating the credit rating of a firm as though a Markov chain, with ratings transition probabilities estimated as long-term average ratings transition frequencies.¹ It is by now well understood, however, that the current rating of a firm does not incorporate much of the influence of the business cycle on failure rates (Nickell, Perraudin, and Varotto (2000), Kavvathas (2001)), nor even the effect of prior ratings history (Behar and Nagpal (1999), Lando and Skødeberg (2002)). There is, moreover, significant heterogeneity in the short-term failure probabilities of different firms of the same current rating (Kealhofer (2003)).

We anticipate several types of applications for our work, including *(i)* the analysis by a bank of the credit quality of a borrower over various future potential borrowing periods, for purposes of loan approval and pricing, *(ii)* the determination by banks and bank regulators of the appropriate level of capital to be held by banks, in light of the current state of their loan portfolio, especially given the upcoming Basel II accord, under which borrower default probabilities are to be introduced for the purpose of determining the capital to be held as backing for a loan to a given borrower, *(iii)* the determination of credit ratings by rating agencies, and *(iv)* the ability to shed some light on the macroeconomic links between business-cycle variables and the failure risks of corporations. Absent a model that incorporates the dynamics of the underlying covariates, it seems difficult to extrapolate prior models of one-quarter-ahead or one-year-ahead default probabilities to longer time horizons. While one could separately estimate models of fixed-horizon failure probabilities for each of various alternative time horizons, it seems natural and statistically more efficient to incorporate joint consistency conditions for failure probabilities at various time horizons within one model.

The conditional survival and default probabilities, $p(X_t, s)$ and $q(X_t, s)$, depend on:

¹See, for example, Duffie and Singleton (2003), Chapter 4.

- a parameter vector β determining the dependence of the failure and other-exit intensities, $\Lambda(X_t)$ and $A(X_t)$, respectively, on the covariate vector X_t , and
- a parameter vector γ determining the time-series behavior of the underlying state vector X_t of covariates.

The doubly-stochastic assumption, stated more precisely in Section 2, is that, conditional on the paths of the underlying state variables determining failure and other-exit intensities for all firms, these exit times are the first event times of independent Poisson processes with the same (conditionally deterministic) intensity paths.² In particular, this means that, given the path of the state-vector process, the merger and default times of different firms are conditionally independent.

A major advantage of the doubly-stochastic formulation is that it allows decoupled maximum-likelihood estimations of β and γ , which can then be combined to obtain the maximum-likelihood estimators of the survival and failure probabilities, $p(X_t, s)$ and $q(X_t, s)$, and other properties of the model, such as probabilities of joint failure of more than one firm. The maximum likelihood estimator of the intensity parameter vector β is the same as that of a conventional competing-risks duration model with time-varying covariates, because of the doubly-stochastic assumption. The maximum likelihood estimator of the time-series parameter vector γ would depend of course on the particular specification adopted for the time-series behavior of the state process X . Our approach is quite flexible in that regard. For examples, we could allow the state process X to have GARCH volatility behavior, to depend on hidden Markov chain “regimes,” or to have jump-diffusive behavior. For our specific empirical application to the U.S. machinery and instrument sector, we have adopted a simple Gaussian vector auto-regressive specification for the firm-specific leverage variables and the macroeconomic growth variables, and we use the conventional maximum-likelihood estimator for the associated parameter vector γ . A further advantage of this methodology is that it allows straightforward maximum-likelihood estimation of the term structure of failure probabilities, by simply substituting the maximum-likelihood estimators for β and γ into (2). Asymptotic confidence intervals for the term

²One must take care in interpreting this characterization when treating the “internal covariates,” those that are firm-specific and therefore no longer available after exit, as explained in Section 2.

structure of future default probabilities can then be obtained by the usual “Delta” method, as explained in the appendix.

The doubly-stochastic assumption is overly restrictive in settings for which failure or another form of exit by one firm could have an important direct influence on the failure or other-exit intensity of another firm. This influence would be anticipated to some degree if one firm plays a relatively large role in the marketplace of another. Our empirical results should therefore be treated with caution, although we hope that, because they address for the first time the impact of stochastic covariates on multi-period failure probabilities, these results may nevertheless be useful.

In our study of the U.S. industrial machinery and instrument sector between 1971 and 2001, we find that corporate failure probabilities depend significantly on both firm leverage and business-cycle covariates. We use a volatility-corrected measure of leverage, *distance to default*, that has become a standard default covariate in industry practice (Kealhofer 2003). We illustrate the degree of dependence of long-horizon failure probabilities on the time-series behavior of the covariates, principally through the effects of mean reversion, long-run means, and volatilities of distances to default and of national income growth.

Our methods also lead to a calculation at time t of the conditional probability $P(\{T \in [t, u]\} \cup \{q(X_u, s) > \bar{q}\} | X_t)$ that the failure time T of a given firm is before some given future time u , or that the firm’s s -year failure probability at time u will exceed a given level \bar{q} . This and related calculations could play a role in credit rating, risk management, and regulatory applications. The estimated model can be further used to calculate probabilities of joint failure of groups of firms, or other properties related to failure correlation. In our doubly-stochastic model setting, failure correlation between firms arises from correlation in their failure intensities due to (i) common dependence of these intensities on macro-variables and (ii) correlation across firms of changes in leverage.

Our econometric methodology may be useful in other subject areas requiring estimators of multi-period survival probabilities under exit intensities that depend on covariates with pronounced time-series dynamics. Examples might include the timing of real options such as technology switch, mortgage prepayment, securities issuance, and labor mobility. We are unaware of previously available econometric methodology for multi-period event prediction under stochastic covariates. That is, there has been extensive research on multi-period event prediction (for example, baseline-hazard duration mod-

els), and there is a separate literature on stochastic event intensity estimation, but we are aware of no prior work that estimates multi-period event probabilities under intensities depending on stochastic covariates.

1.1 Related Literature

A standard structural model of bankruptcy timing assumes that a corporation fails when its assets drop to a sufficiently low level relative to its liabilities. For example, the prototypical models of Black and Scholes (1973), Merton (1974), Fisher, Heinkel, and Zechner (1989), and Leland (1994), take the asset process to be a geometric Brownian motion. In these models, a sufficient statistic for conditional failure probabilities is the distance to default, which is the number of standard deviations of annual asset growth by which the current asset level exceeds the firm's liabilities. This failure covariate, using market equity data and accounting data for liabilities, has been adopted in industry practice by Moody's KMV, a leading provider of estimates of failure probabilities for essentially all publicly traded firms. (See Crosbie and Bohn (2002) and Kealhofer (2003).) Based on this theoretical foundation, it seems natural to include distance to default as a covariate.

In the context of a standard structural default model of this type, Duffie and Lando (2001) show that if the distance to default cannot be accurately measured, then a filtering problem arises, and the failure intensity depends on the measured distance to default and also on other covariates that may reveal additional information about the firm's conditional failure probability.

More generally, a firm's financial health may have multiple influences over time. For example, firm-specific, sector-wide, and macroeconomic state variables may all influence the evolution of corporate earnings and leverage. Given the usual benefits of parsimony, the preliminary model of long-horizon failure probabilities analyzed in this paper adopts two failure covariates, only, distance to default and U.S. personal income growth. (Other macroeconomic performance measures might serve approximately as well.) Given distance to default, the choice of a second covariate calls for a tradeoff between variables that are more directly tied to the firm's marketplace (such as sector performance measures), and variables that capture information that is not largely explained by distance to default.

Prior empirical models of corporate failure probabilities, reviewed by Jones (1987) and Hillegeist et al. (2003), have relied on many types of covariates, both fixed and time-varying. Prior work has not, however, attempted

to estimate failure probabilities over multiple time periods that exploits the time-series behavior of the covariates.

The first generation of empirical corporate failure analysis, originating with Altman (1968), was based on multivariate discriminant analysis. Among the covariates used in Altman’s ‘*Z*-score’ is a measure of leverage, defined as the market value of equity divided by the book value of total debt. Our distance to default covariate is essentially a volatility-corrected measure of leverage.

A second generation of empirical work is based on qualitative-response models, such as logit and probit. Among these, Ohlson (1980) used an “*O*-score” method in his year-ahead failure prediction model.

The latest generation of modeling is dominated by duration analysis. Early in this literature is the work of Lane, Looney, and Wansley (1986) on bank failure prediction, using time-independent covariates.³ These models typically apply a Cox proportional-hazard model. Lee and Urrutia (1996) used a duration model based on a Weibull distribution of failure time. They compare duration and logit models in forecasting insurer insolvency, finding that, for their data, a duration model identifies more significant variables than does the logit model.

Duration models based on time-varying covariates include those of McDonald and Van de Gucht (1999), in a model of the timing of high-yield bond failures and call exercises, taking as time-varying covariates the bond’s age and treating certain macroeconomic variables as time-varying covariates.⁴ Related duration analysis by Shumway (2001), Kavvathas (2001), and Hillegeist, Keating, Cram, and Lundstedt (2003) predicts bankruptcy.⁵

Shumway (2001) uses a discrete duration model with time-dependent covariates. Computationally, this is equivalent to a multi-period logit model with an adjusted-standard-error structure. In predicting one-year failure, Hillegeist, Keating, Cram, and Lundstedt (2003) also exploit a discrete duration model. By taking as a covariate the theoretical probability of failure implied by the Black-Scholes-Merton’s model, based on distance to default, Hillegeist, Keating, Cram, and Lundstedt (2003) find, at least in this model setting, that distance to default is not a sufficient statistic for predicting failure. Accounting-based and macroeconomic variables are also relevant. Our

³Whalen (1991) and Wheelock and Wilson (2000) also used Cox proportional-hazard models for bank failure analysis.

⁴Meyer (1990) used a similar approach in a study of unemployment duration.

⁵Kavvathas (2001) also analyzes the transition of credit ratings.

results confirm this conclusion for our data, and extend the analysis to multiple period prediction. Further discussion of the selection of covariates for corporate failure prediction may be found in Section 3.2.

Moving from the empirical literature on corporate failure prediction to the statistical methods available for this task, typical econometric treatments of stochastic intensity models include Lancaster (1990) and Kalbfleisch and Prentice (2002), which provide likelihood functions in settings similar to ours.⁶ In their language, our macro-covariates are “external,” and our firm-specific covariates are “internal,” that is, cease to be generated once a firm has failed. These sources do not treat large-sample properties, nor indeed do such properties appear to have been developed in a form suitable for our application. For example, Berman and Frydman (1999) do provide asymptotic properties for maximum-likelihood estimators of stochastic intensity models, including a version of Cramèr’s Theorem, but treat only cases in which the covariate vector X_t is fully external (with known transition distribution), and in which event arrivals continue to occur, repeatedly, at the specified parameter-dependent arrival intensity. This clearly does not treat our setting, for a firm typically disappears once it fails.⁷

2 Econometric Model

This section outlines our probabilistic model for corporate survival, and the estimators that we propose. The following section applies the estimator to data on the U.S. industrial machinery and instrument sector.

2.1 Conditional Survival and Failure Probabilities

Fixing a probability space (Ω, \mathcal{F}, P) and an information filtration $\{\mathcal{G}_t : t \geq 0\}$ satisfying the usual conditions,⁸ let $X = \{X_t : t \geq 0\}$ be a time-homogeneous

⁶For other textbook treatments, see Andersen, Borgan, Gill, and Keiding (1992), Miller (1981), Cox and Isham (1980), Cox and Oakes (1984), Daley and Vere-Jones (1988), and Therneau and Grambsch (2000).

⁷For the same reason, the autoregressive conditional duration framework of Engle and Russell (1998) and Engle and Russell (2002) is not suitable for our setting, for the updating of the conditional probability of an arrival in the next time period depends on whether an arrival occurred during the previous period, which again does not treat a firm that disappears once it fails.

⁸See Protter (1990) for technical definitions.

Markov process in \mathbb{R}^d , for some integer $d \geq 1$. The state vector X_t is a covariate for a given firm's exit intensities, in the following sense. Let (M, N) be a doubly-stochastic non-explosive two-dimensional counting process driven by X , with intensities $\alpha = \{\alpha_t = A(X_t) : t \in [0, \infty)\}$ for M and $\lambda = \{\lambda_t = \Lambda(X_t) : t \geq 0\}$ for N , for some non-negative real-valued measurable functions $A(\cdot)$ and $\Lambda(\cdot)$ on \mathbb{R}^d . Among other implications, this means that, conditional on the path of X , the counting processes M and N are independent Poisson processes with conditionally deterministic time-varying intensities, α and λ , respectively. For details on these definitions, one may refer to Karr (1991) and Appendix I of Duffie (2001).

We suppose that a given firm exits (and ceases to be observable) at $\tau = \inf\{t : M_t + N_t > 0\}$, which is the earlier of the first event time of N , corresponding to failure, and the first event time of M , corresponding to exit for some other reason. In our application to the U.S. industrial machinery and instrument sector, the portion of exits for reasons other than failure is far too substantial to be ignored.

The main idea is that, so long as the firm has not exited for some reason, its failure intensity is $\Lambda(X_t)$ and its intensity of exit for other reasons is $A(X_t)$.

It is important to allow the state vector X_t to include firm-specific failure covariates that cease to be observable when the firm exits at τ . For simplicity, we suppose that $X_t = (U_t, Y_t)$, where U_t is firm-specific and Y_t is macroeconomic. Thus, we consider conditioning by an observer whose information is given by the smaller filtration $\{\mathcal{F}_t : t \geq 0\}$, where \mathcal{F}_t is the σ -algebra generated by

$$\{(U_s, M_s, N_s) : s \leq \min(t, \tau)\} \cup \{Y_s : s \leq t\}.$$

We now verify that the observer's time- t conditional probabilities $p(X_t, s)$ and $q(X_t, s)$ of survival for s years, and of failure within s years, respectively, are as shown in (1) and (2). The firm's failure time is $T = \inf\{u > t : N_u > N_t, M_u = M_t\}$.

Proposition 1. *On the event $\{\tau > t\}$ of survival to t , the \mathcal{F}_t -conditional probability of survival to time $t + s$ is*

$$P(\tau > t + s | \mathcal{F}_t) = p(X_t, s),$$

where $p(X_t, s)$ is given by (1), and the \mathcal{F}_t -conditional probability of failure by $t + s$ is

$$P(T < t + s | \mathcal{F}_t) = q(X_t, s),$$

where $q(X_t, s)$ is given by (2).

Proof: We begin by conditioning instead on the larger information set \mathcal{G}_t , and later show that this does not affect the result.

We first calculate that, on the event $\{\tau > t\}$,

$$P(\tau > t + s | \mathcal{G}_t) = p(X_t, s), \quad (3)$$

and

$$P(T < t + s | \mathcal{G}_t) = q(X_t, s). \quad (4)$$

The first calculation (3) is standard, using the fact that $M + N$ is a doubly-stochastic counting process with intensity $\alpha + \lambda$. For the second calculation (4), we use the fact that, conditional on the path of X , the (improper) density, evaluated at any time $z > t$, of the failure time T , exploiting the X -conditional independence of M and N is, with the standard abuse of notation,

$$P(T \in dz | X) = P(\inf\{u : N_u \neq N_t\} \in dz, M_z = M_t | X) \quad (5)$$

$$= P(\inf\{u : N_u \neq N_t\} \in dz | X) P(M_z = M_t | X) \quad (6)$$

$$= e^{-\int_t^z \lambda(u) du} \lambda(z) dz e^{-\int_t^z \alpha(u) du} \quad (7)$$

$$= e^{-\int_t^z (\alpha(u) + \lambda(u)) du} \lambda_z dz. \quad (8)$$

From the doubly-stochastic property, conditioning also on \mathcal{G}_t has no effect on this calculation, so

$$P(T \in [t, t + s] | \mathcal{G}_t, X) = \int_t^{t+s} e^{-\int_t^z (\alpha(u) + \lambda(u)) du} \lambda_z dz. \quad (9)$$

Now, taking the expectation of this conditional probability given \mathcal{G}_t only, using the law of iterated expectations, leaves (4).

On the event $\{\tau > t\}$, the conditioning information in \mathcal{F}_t and \mathcal{G}_t coincide. That is, every event contained by $\{\tau > t\}$ that is in \mathcal{G}_t is also in \mathcal{F}_t . The result follows.

One can calculate $p(X_t, s)$ and $q(X_t, s)$ explicitly in certain settings, for example if the state vector X is affine and the exit intensities have affine dependence on X , as shown in various cases by Duffie, Pan, and Singleton (2000), Duffie, Filipović, and Schachermayer (2003). In our eventual application, however, in order to simplify the maximum-likelihood problem, our specification of the dependence of the intensities on X_t is non-linear, which calls for numerical solutions of $p(X_t, s)$ and $q(X_t, s)$, as we shall see in Section 3.

2.2 Maximum Likelihood Estimator

We turn to the problem of inference from data.

For each of n firms, we let $T_i = \inf\{t : N_{it} > 0, M_{it} = 0\}$ denote the failure time of firm i , and let $S_i = \inf\{t : M_{it} > 0, N_{it} = 0\}$ denote the censoring time for firm i due to other forms of exit, subscripting by i in the obvious way. We let U_{it} be the firm-specific vector of variables that are observable for firm i until its exit time $\tau_i = \min(S_i, T_i)$, and let Y_t denote the vector of environmental variables (such as business-cycle variables) that are observable at all times. We let $X_{it} = (U_{it}, Y_t)$, and assume, for each i , that $X_i = \{X_{it} : t \geq 0\}$ is a Markov process. (This means that, given Y_t , the transition probabilities of U_{it} do not depend on U_{jt} for $j \neq i$, a simplifying assumption.) Because, in our current implementation of the model, we observe these covariates X_{it} only quarterly, we take $X_{it} = X_{i,k(t)} = Z_{i,k(t)}$, where $k(t)$ denotes the last (integer) discrete time period before t , and where Z_i is the time-homogeneous discrete-time Markov process of covariates for firm i . This means that X_i is constant between periodic observations, a form of time-inhomogeneity that involves only a slight extension of our basic theory of the Section 2.1. We continue to measure time continuously, however, because we wish to allow the use of information associated with the intra-period timing of exits.

Extending our notation from Section 2.1, for all i , we let $\Lambda(X_{it}, \beta)$ and $A(X_{it}, \beta)$ denote the failure and other-exit intensities of firm i , where β is a parameter vector, common to all firms, to be estimated. This homogeneity across firms allows us to exploit both time-series and cross-sectional data, and is traditional in duration models of failure such as Shumway (2001). This leads to inaccurate estimators to the degree that the underlying firms are actually heterogeneous in this regard. We do, however, allow for heterogeneity across firms with respect to the probability transition distributions of the Markov covariate processes Z_1, \dots, Z_n of the n firms. For example, some firms may have different target leverage ratios than others.

We assume that the exit-counting process $(M_1, N_1, \dots, M_n, N_n)$ of the n firms is doubly-stochastic driven by $X = (X_1, \dots, X_n)$, in the sense of Section 2.1, so that the exit times τ_1, \dots, τ_n of the n firms are X -conditionally independent. There is some important loss of generality here, for this implies that the exit of one firm has no direct impact on the failure intensity of another firm. Their failure times are correlated only insofar as their exit intensities are correlated. Similarly, an exit by one firm due to a merger or acquisition has no direct impact on the failure intensity of another firm, and

so on.

The econometrician's information set \mathcal{F}_t at time t is that generated by

$$\mathcal{I}_t = \{Y_s : s \leq t\} \cup \mathcal{J}_{1t} \cup \mathcal{J}_{2t} \cdots \cup \mathcal{J}_{nt},$$

where the information set for firm i is

$$\mathcal{J}_{it} = \{(1_{S_i < s}, 1_{T_i < s}, U_{is}) : t_i^0 \leq s \leq \min(S_i, T_i, t)\},$$

and where t_i^0 is the time of first appearance of firm i in the data set. For simplicity, we take t_i^0 to be at the end of a discrete time period and deterministic, but our results would extend to treat left-censoring of each firm at a stopping time, under suitable conditional independence assumptions. It would be enough, for example, that the left-censoring times are (\mathcal{F}_t) -stopping times, which leaves our likelihood functions unaffected.

In order to simplify the estimation of the time-series model of covariates, we suppose that the environmental discrete-time covariate process $\{Y_1, Y_2, \dots\}$ is itself a time-homogeneous (discrete-time) Markov process.

Conditional on $Z_k = (Z_{1k}, \dots, Z_{nk})$, we suppose that Z_{k+1} has some joint density $f(\cdot | Z_k; \gamma)$, for some parameter vector γ to be estimated. Despite our prior Markov assumption on the covariate process $\{Z_{ik} : k \geq 1\}$ for each firm i , this allows for conditional correlation between $U_{i,k+1}$ and $U_{j,k+1}$ given (Y_k, U_{ik}, U_{jk}) . We emphasize that this transition density $f(\cdot)$ is not conditioned on survivorship.

As a notational convenience, whenever $K \subset L \subset \{1, \dots, n\}$ we let $f_{KL}(\cdot | Y_k, \{U_{ik} : i \in L\}; \gamma)$ denote the joint density of $(Y_{k+1}, \{U_{i,k+1} : i \in K\})$ given Y_k and $\{U_{ik} : i \in L\}$, which is a property of (in effect, a marginal of) $f(\cdot | Z_k; \gamma)$. In our eventual application, we will further assume that $f(\cdot | z; \gamma)$ is a joint-normal density, which makes the marginal density function $f_{KL}(\cdot | y, \{u_i : i \in L\})$ an easily-calculated joint normal.

For additional convenient notation, let $R(k) = \{i : \tau_i > k\}$ denote the set of firms that survive to at least period k , let $\tilde{U}_k = \{U_{ik} : i \in R(k)\}$, $S_i(t) = \min(t, S_i)$, $S(t) = (S_1(t), \dots, S_n(t))$, and likewise define $T_i(t)$ and $T(t)$. Under our doubly-stochastic assumption, the likelihood for the information set \mathcal{I}_t is

$$\mathcal{L}(\mathcal{I}_t; \gamma, \beta) = \mathcal{L}(\tilde{U}, Y; \gamma) \times \mathcal{L}(S(t), T(t); Y, \tilde{U}, \beta), \quad (10)$$

where

$$\mathcal{L}(\tilde{U}, Y; \gamma) = \prod_{k=0}^{k(t)} f_{R(k+1), R(k)}(Y_{k+1}, \tilde{U}_{k+1} | Y_k, \tilde{U}_k; \gamma), \quad (11)$$

and

$$\mathcal{L}(S(t), T(t); Y, \tilde{U}; \beta) = \prod_{i=1}^n G_{it}(\beta), \quad (12)$$

for

$$\begin{aligned} G_{it}(\beta) &= \exp \left(- \int_{t_i^0}^{H_i} (A(Z_{i,k(s)}; \beta) + \Lambda(Z_{i,k(s)}; \beta)) ds \right) \\ &\quad \times (1_{H_i=t} + A(Z_{i,S_i}; \beta) 1_{S_i(t)<t} + \Lambda(Z_{i,T_i}; \beta) 1_{T_i(t)<t}), \end{aligned}$$

where $H_i = \min(S_i(t), T_i(t)) = \min(\tau_i, t)$.

Because of this structure for the likelihood, we can decompose the overall maximum likelihood estimation problem into the separate problems

$$\sup_{\gamma} \mathcal{L}(\tilde{U}, Y; \gamma) \quad (13)$$

and

$$\sup_{\beta} \mathcal{L}(S, T; Y, \tilde{U}, \beta). \quad (14)$$

Further simplification is obtained by taking the parameter vector β determining intensity dependence on covariates to be of the decoupled form $\beta = (\mu, \nu)$, with

$$\lambda_{it} = \Lambda(X_{it}; \mu); \quad \alpha_{it} = A(X_{it}; \nu). \quad (15)$$

(This involves a slight abuse of notation.) This means that the form of dependence of the failure intensity on the covariate vector X_{it} does not restrict the form of the dependence of the other-exit intensity, and *vice versa*. An examination of the structure of (14) reveals that this decoupling assumption allows problem (14) to be further decomposed into the pair of problems

$$\sup_{\mu} \prod_{i=1}^n e^{-\int_{t_i^0}^{H_i} \Lambda(X_i(u); \mu) du} (1_{H_i=t} + \Lambda(X_i(T_i); \mu) 1_{T_i(t)<t}) \quad (16)$$

and

$$\sup_{\nu} \prod_{i=1}^n e^{-\int_{t_i^0}^{H_i} A(X_i(u); \nu) du} (1_{H_i=t} + A(X_i(S_i); \nu) 1_{S_i(t)<t}). \quad (17)$$

We have the following result, which summarizes our parameter-fitting algorithm.

Proposition 2. *Solutions γ^* and β^* of the respective maximum-likelihood problems (13) and (14) collectively form a solution to the overall maximum-likelihood problem*

$$\sup_{\gamma, \beta} \mathcal{L}(\mathcal{I}_t; \gamma, \beta). \quad (18)$$

Under the parameter-decoupling assumption (15), solutions μ^ and ν^* to the maximum-likelihood problems (16) and (17), respectively, form a solution $\beta^* = (\mu^*, \nu^*)$ to problem (14).*

In our particular empirical application, as explained in Section 3, each of these optimization problems is solved numerically. The decomposition of the MLE optimization problem given by Proposition 2 allows a significant degree of tractability.

Under the usual technical regularity conditions, given a maximum-likelihood estimator (MLE) $\hat{\theta}$ for some parameter θ , the maximum-likelihood estimator (MLE) of $h(\theta)$, for some smooth function $h(\cdot)$, is $h(\hat{\theta})$. Thus, under these technical conditions, the maximum likelihood estimators of the survival probability $q(X_t, s)$ and the failure probability $p(X_t, s)$ are obtained by (1) and (2), using the maximum likelihood estimators for $\beta = (\mu, \nu)$ and γ to determine the distributions underlying these expectations.

Under further technical conditions, an MLE is consistent, efficient, and asymptotically normal, in the sense that the difference between the maximum-likelihood estimator and the “true” data-generating parameter, scaled by the square root of the number of observations, converges weakly to a vector whose distribution is joint normal with mean zero and with a well-known covariance matrix (Amemiya 1985). In our case, it is apparent that a consistency result will require allowing both the number n of firms and the number $k(t)$ of periods of data to become large in this sense. In Section 3, we will compare asymptotic standard errors with standard errors based on a typical bootstrap simulation procedure (Efron and Tibshirani 1993).

3 Empirical Analysis

The section describes our data set, specific parameterizations of our covariate processes and intensity models, our parameter estimates, some of their

properties, and some of the substantive conclusions regarding the behavior of conditional term structures of failure hazard rates. We are particularly interested in the sensitivity of these term structures of failure hazard rates to firm-specific and macroeconomic variables.

3.1 Data

We use three main data sources. Quarterly balance sheets and income statements for each firm in our sector are from the Compustat database. Stock price and market capitalization data are from the Center of Research in Security Prices (CRSP). Both of these databases are available online from Wharton Research Data Services. The third source is the National Economic Accounts of The Bureau of Economic Analysis, from which we get quarterly business-cycle variables.

Our target set of firms consists of those publicly-traded companies that are recorded in the Compustat industrial database as a member of the industrial machinery and instrument sector (2-digit SIC code 35). This sector is chosen mainly for the fact that it has the largest number of bankruptcies recorded in the Compustat database during our sample period, 1971 to 2001. Also, failures in this sector are not concentrated within a short time period, as is the case, for examples, in the banking and oil-and-gas sectors. A concentration of bankruptcies within a short time period would limit our ability to identify dependence of failure intensities on macroeconomic variables.

Specifically, we include all firms of this sector for which the data necessary to construct our covariates are available from Compustat and CRSP, with exceptions to be noted. In all, 870 such firms existed during our sample period, of which 332 remain “active” as public firms as of the end of our sample period, the end of 2001.

Of the remaining 538 firms, 70 failed, filing for bankruptcy under Chapter 11 or Chapter 7 of the U.S. Bankruptcy Code. Coverage by Computstat of the remaining 468 firms was discontinued for other reasons, such as merger, acquisition, leveraged buyout, privatization, or ceasing to provide data.⁹ Exits are recorded by month in the Compustat annual file (item AFTNT33 and item AFTNT34), and are assumed for purposes of our estimation to occur at the end of the relevant month. Approximately 80% of the non-failure exits in our sample were due to merger and acquisition activities.

⁹The alternative forms of exit can be determined from Compustat item AFTNT35.

Among our sample of 870 firms, 50 firms existed before the January, 1971 beginning of our sample period. As explained earlier, start dates are non-informative in our model. That is, with our probabilistic specification of exit times, “left-censoring” does not call for an adjustment to the likelihood function. In practice, however, there is likely to be some influence on failure intensities, given our other covariates, on survival time. For example, one might estimate this effect with a baseline hazard rate, which would call for a left-censorship adjustment for the firms that existed before our sample period.

Our sample period begins at the first quarter of 1971, and ends¹⁰ at the fourth quarter of 2001. Although CRSP stock-price data are available from 1925, and Compustat quarterly coverage of public firms begins in 1962, only since 1972 has Compustat reported quarterly data on short-term liabilities, a relatively important determinant of our distance-to-default covariate.¹¹ In any case, of the 1,283 failures in all industries recorded in Compustat, only 9 are reported to have occurred before 1971. Our decision to include only post-1971 data was also followed by Shumway (2001) and Vassalou and Xing (2003).

3.2 Covariates

We have examined the dependence of estimated failure and other-exit intensities on several types of firm-specific, sector-wide, and macroeconomic variables. These include:

1. *Distance to default*, which, roughly speaking, is the number of standard deviations of quarterly asset growth by which current assets exceed a standard measure of current liabilities. As explained Section 1.1, this covariate has theoretical underpinnings in the Black-Scholes-Merton structural model of default probabilities. Our method of construction of this covariate, based on market equity data and Compustat book liability data, is along the lines of that used by Vassalou and Xing (2003), Crosbie and Bohn (2002), and Hillegeist, Keating, Cram, and Lundstedt (2003). Details are given in Appendix A.

¹⁰We include Compustat and CRSP variables from the fourth quarter of 1970 in order to estimate the failure intensity during the first quarter of 1971. Likewise, lagged-quarter variables are used to fit the covariate time-series.

¹¹The quality of pre-1971 Compustat data seems somewhat unreliable. For example, pre-1971 liability data for many companies are missing.

2. *Personal income growth.* As a measure of macroeconomic performance, we use U.S. personal income growth. Data on quarterly national personal income (seasonally adjusted) is obtained from the Bureau of Economic Analysis' National Economic Accounts database. Personal income growth is measured in terms of quarterly percentage changes. McDonald and Van de Gucht (1999) used quarterly industrial production growth in the U.S. as a covariate for high-yield bond failure. Hillegeist, Keating, Cram, and Lundstedt (2003) exploit the national rate of corporate bankruptcies, in a baseline-hazard-rate model of default. Our selection of personal income growth is pragmatic; we found it to be more contemporaneous with failure rates than GDP growth rates, which seems to have lagged effects. This may be because, like bankruptcy, personal income growth lags business-cycle movement.¹²
3. *Sector earnings performance,* measured as the sector average across firms of the ratios of earnings to assets. When used as the only environmental covariate, sector earnings performance is indeed a statistically significant covariate for failure intensity (although not for other-exit intensity). When included as an additional covariate together with distance to default and personal income growth, however, personal income growth is statistically significant, but sector earnings performance is not.
4. *Firm-level Earnings.* Another firm-specific covariate is earnings, defined as the ratio of net income (Compustat item 69) to total assets (item 44). In contrast to distance to default, this earnings covariate comes solely from accounting data. Since there is a lag in recording a firm's accounting data, we lag earnings by one period in order to ensure that it is observable at the beginning of the quarter. (Because of this, firms enter our study only at the second quarter after their first ap-

¹²Many prior works find a correlation between macroeconomic conditions and failure. (See Allen and Saunders (2002) for a survey.) For example, Fons (1991), Blume and Keim (1991), and Jonsson and Fridson (1996) document that aggregate failure rates tend to be high in the downturn of business cycles. McDonald and Van de Gucht (1999), Hillegeist, Keating, Cram, and Lundstedt (2003), Chava and Jarrow (2002), and Lennox (1999) examine macroeconomic effects on individual-level failure or bankruptcies. Pesaran, Schuermann, Treutler, and Weiner (2003) use a comprehensive set of country-specific macro variables to estimate the effect of macroeconomic shocks in one region on the credit risk of a global loan portfolio. See, also, Vassalou and Xing (2002)

pearance in our database.) When there are occasionally missing data for net income, we substitute values from the most immediately available past quarters. Earnings is complementary to distance to default, and a traditional predictor for bankruptcy since Altman (1968). When not appearing together with distance to default, earnings is a significant failure covariate in both logit and duration models, as shown by Shumway (2001).¹³ Our definition of earnings is that used by Zmijewski (1984), and close to the profitability covariates defined by Altman (1993), the ratio of retained earnings to total assets, and the ratio of earnings before interest and taxes to total assets.

5. *Firm Size.* Firm size is defined as the logarithm of a firm’s book value of total assets (Compustat item 44). Firm size may control for unobserved heterogeneity across firms, since big firms and small firms may have different market power, management strategies, or borrowing ability, all of which may affect the risk of failure. For example, it might be easier for a big firm to re-negotiate with its creditors to postpone the repayment of debt, or to raise new funds to pay the old debt. In a “too-big-to-fail” sense, firm size may negatively influence failure intensity. The statistical significance of size as a determinant of failure risk has been documented in Shumway (2001).¹⁴

Some summary statistics of the covariates that we use are reported in Table 1. In order to maintain a parsimonious model, our currently reported estimator of the term structure of conditional failure hazard rates is based on a two-covariate model, including only personal income growth and, firm by firm, distance to default.

Regarding our selection of macroeconomic covariates, many alternatives come to mind. We plan to explore some of these; our current formulation is preliminary. McDonald and Van de Gucht (1999) use quarterly U.S. industrial production growth (defined as first difference in logarithms) as a covariate for failure risk, for a sample of 382 high-yield corporate bonds issued between 1977 and 1989, and followed through 1994. They find that the conditional likelihood of failure in a given quarter is significantly and negatively correlated with industrial production growth in the current quarter, the

¹³Another robustly significant factor according to Shumway’s paper is leverage ratio, but that has been incorporated in distance to default

¹⁴In Shumway (2001), size is defined as the logarithm of a firm’s market capitalization, relative to the total size of the NYSE and AMEX market.

Table 1: Covariate summary statistics

variable	median	mean	sdev	min	max	obs.
distance to default	3.74	4.40	3.67	-3.89	73.50	28,612
firm size	4.06	4.30	1.95	0.30	12.28	28,612
firm-level earnings	0.01	-0.01	0.12	-12.78	1.19	28,612
sector earnings	1.39	1.30	0.57	-1.49	2.65	124
personal income growth	1.74	1.88	0.95	-1.38	4.39	124

preceding three quarters, and the immediately subsequent quarter.¹⁵ A key covariate in Hillegeist, Keating, Cram, and Lundstedt (2003) is the economy-wide rate of corporate bankruptcies among publicly-traded firms over the prior 12 months. Their test of relative information content shows that this bankruptcy rate provides significant incremental information for forecasting bankruptcy rates for a sample of 10,845 public firms between 1979 and 1997, beyond that provided by the “Score” variables of Altman (1968) and Ohlson (1980), and the Black-Scholes-Merton’s modeled probability of failure, calculated in the manner indicated in Appendix A. Keenan, Sobehart, and Hamilton (1999) and Helwege and Kleiman (1997) model the forecasting of aggregate year-ahead U.S. default rates on corporate bonds, using, among other covariates, credit rating, age of bond, and various macroeconomic variables, including industrial production, interest rates, trailing default rates, aggregate corporate earnings, and indicators for recession.

3.3 Covariate Time-Series Models

In this subsection, we specify particular parametric time-series models for the covariate processes that we use to estimate term structures of conditional failure probabilities, and then provide the associated maximum-likelihood estimates.

With our current formulation, we have an extremely high-dimensional state-vector, consisting of one macroeconomic covariate, personal income growth, Y_t , and the distance to default, D_{it} , for each firm i of the n firms

¹⁵McDonald and Van de Gucht (1999) also take as covariates for bond default: the issuer’s rating, coupon rate, issue size, initial maturity, and issuance period (whether before or after 1985).

that existed during our sample period, 870 in all. Unrestricted by additional structure, this presents an unwieldy time-series model to estimate. After preliminary examination of various feasibly estimated alternatives, we have opted for a simple specification in which each of $Y_t, D_{1t}, \dots, D_{nt}$ is a univariate first-order auto-regressive Gaussian process, allowing for correlation among their innovations.

Specifically, personal income growth in quarter k , Y_k , is assumed to satisfy

$$Y_{k+1} - Y_k = \kappa_Y(\theta_Y - Y_k) + \sigma_Y \epsilon_{k+1}, \quad (19)$$

where $\epsilon_1, \epsilon_2, \dots$ is an independent sequence of standard normal variables, and $\phi = (\theta_Y, \kappa_Y, \sigma_Y)$ is a parameter vector to be estimated. Here, θ_Y is the long-run mean, κ_Y is the mean-reversion rate, and σ_Y is the standard deviation of the innovations.

Similarly, for each firm i , for the quarters in which this firm appears in our sample,

$$D_{i,k+1} - D_{ik} = \kappa_D(\theta_{Di} - D_{ik}) + v w_{i,k+1}, \quad (20)$$

where $\{w_{ik} : k \geq 1\}$ is an independent sequence of standard normals, κ_D is a mean-reversion parameter common to all firms, v is an innovations standard-deviation parameter common to all firms, and θ_{Di} is a long-run mean ¹⁶ parameter that is specific to firm i . The parameters v and κ_D characterize the degree of volatility and mean reversion in this leverage-related variable. Volatility arises from uncertainty in earnings performance and in the revaluation of assets and liabilities. Mean-reversion arises from leverage targeting, by which corporations commonly pay out dividends and other forms of distributions when they achieve a sufficiently low degree of leverage, and conversely attempt to raise capital and retain earnings to a higher degree when their leverage introduces financial distress or business inflexibility, as modeled by Leland (1998) and Collin-Dufresne and Goldstein (2001). We assume homogeneity of κ_D and v across the sector, as we do not have *a priori* reasons to assume that different firms in the same sector revert to their targeted volatility-adjusted leverages differently from one another, and also in order to maintain a parsimonious model in the face of limited time-series data on each firm. (Monte Carlo tests confirm significant small-sample bias in MLE estimators for our sample size that would arise if we estimated firm-by-firm mean reversion and volatility parameters.)

¹⁶This is the long-run mean ignoring the effect of exit.

A key question is how to empirically model the targeted distance to default, θ_{Di} of firm i . Despite the arguments that swayed us to assume homogeneity across firms of the mean-reversion and volatility parameters κ_D and v , our preliminary analysis showed that applying the same assumption to the targeted distance to default parameter θ_{Di} caused estimated term structures of future failure probabilities to rise dramatically for firms that had consistently maintained low failure probabilities during our sample period. Perhaps some firms might derive reputational benefits from low distress risk, or have firm-specific costs of exposure to financial distress. In the end, we opted to estimate θ_{Di} firm by firm. As a long-run-mean parameter is challenging to pin down statistically in samples of our size, the standard errors in our estimates of θ_{Di} are responsible for a significant contribution to the standard errors of our estimated term structures of future failure probabilities.

After specifying joint normality for the innovations w_{1k}, \dots, w_{nk} and ϵ_k , we tested for, and rejected at conventional confidence levels, positive correlation between w_{ik} and ϵ_k , at least when that correlation is restricted to be common across all firms. While it is somewhat counter to our original intuition that firms' distances to default and national personal income growth do not show significantly positive correlation, the failure of this correlation to appear significantly in our sample may be due to mis-specification, for example in the manner in which correlation arises (perhaps there are substantial lag effects), or in the assumed homogeneity of correlation across different firms. In any case, we adopt a model in which ϵ is independent of $w = (w_1, \dots, w_n)$.

As for the correlation between w_i and w_j , we again adopted a simple homogeneous structure under which

$$w_{ik} = rz_k + \sqrt{1 - r^2} u_{ik}, \quad (21)$$

where u_{1k}, \dots, u_{nk} and z_k are independent standard normals, and r is a constant, so that $\text{corr}(w_{ik}, w_{jk}) = r^2$ whenever $i \neq j$, and $\text{corr}(w_{ik}, w_{jl}) = 0$.

We estimated the time-series parameter vector

$$\gamma = (\kappa_Y, \theta_Y, \sigma_Y, \kappa_D, v, r, \theta_{D1}, \theta_{D2}, \dots, \theta_{Dn})'$$

by maximum likelihood. By independence, the sub-vector $\phi = (\kappa_Y, \theta_Y, \sigma_Y)'$ can be estimated separately from the sub-vector $\xi = (\kappa_D, v, r, \theta_{D1}, \theta_{D2}, \dots, \theta_{Dn})'$, whose high dimension (873 coordinates) required special iterative numerical treatment.

With quarterly data on personal income growth from 1971 to 2001, the MLE estimate of ϕ is:

$$\hat{\phi} = \begin{pmatrix} \hat{\theta}_Y \\ \hat{\kappa}_Y \\ \hat{\sigma}_Y \end{pmatrix} = \begin{pmatrix} 1.8901 \\ 0.6524 \\ 0.8888 \end{pmatrix}, \quad (22)$$

with asymptotic variance-covariance matrix estimate

$$\hat{\Sigma}_\phi = \begin{pmatrix} 0.0076 & 0.0002 & -0.0000 \\ 0.0002 & 0.0156 & -0.0001 \\ -0.0000 & -0.0001 & 0.0033 \end{pmatrix}. \quad (23)$$

Using notation similar to that of Section 2.2, for the firms' distance-to-default processes, we estimate ξ by maximizing the joint likelihood

$$\mathcal{L}(D; \xi) = \prod_k f_{R(k+1), R(k)}(\tilde{D}_{k+1} | \tilde{D}_k; \xi). \quad (24)$$

The likelihood $f_{R(k+1), R(k)}(\cdot | \tilde{D}_k; \xi)$ is the density of a vector of $m(k)$ jointly normal random variables with mean vector \bar{w}_k and covariance matrix Ω_k , where $m(k)$ is the number of firms (indexed as firm $k(1)$ through firm $k(m)$) existing in both periods k and period $k + 1$. We have

$$\bar{w}_k = \begin{pmatrix} \theta_{Dk(1)}\kappa_D + (1 - \kappa_D)D_{k(1)k} \\ \theta_{Dk(2)}\kappa_D + (1 - \kappa_D)D_{k(2)k} \\ \dots \\ \theta_{Dk(m)}\kappa_D + (1 - \kappa_D)D_{k(m)k} \end{pmatrix} \quad (25)$$

and

$$\Omega_k = v^2 \begin{pmatrix} 1 & r^2 & r^2 & \dots & r^2 \\ r^2 & 1 & r^2 & \dots & r^2 \\ \dots & \dots & \dots & \dots & \dots \\ r^2 & r^2 & r^2 & \dots & 1 \end{pmatrix}_{m(k) \times m(k)}. \quad (26)$$

The MLE estimate of the common-parameter vector $(\kappa_D, v, r)'$ is

$$\begin{pmatrix} \hat{\kappa}_D \\ \hat{v} \\ \hat{r} \end{pmatrix} = \begin{pmatrix} 0.1185 \\ 0.9657 \\ 0.2684 \end{pmatrix}. \quad (27)$$

Because the asymptotic covariance matrix for $\hat{\xi}$ is 873×873 , we report in Appendix C only that portion of it that is relevant for the parameter vector $(\hat{\theta}_{Di}, \hat{\kappa}_D, \hat{v}, \hat{r})$ associated with a particular example firm.

3.4 Failure and Other-exit Intensity

As for the failure and other-exit intensity parameterizations, we take

$$\Lambda((Y_k, D_{ik}); \mu) = \exp(\mu_0 + \mu_1 Y_k + \mu_2 D_{ik}) \quad (28)$$

$$A((Y_k, D_{ik}); \nu) = \exp(\nu_0 + \nu_1 Y_k + \nu_2 D_{ik}), \quad (29)$$

respectively, for parameter vectors $\mu = (\mu_0, \mu_1, \mu_2)$ and $\nu = (\nu_0, \nu_1, \nu_2)$ common to all firms. The sample relationship between distance to default and failure frequency shown in Appendix A suggests that the assumed form of exponential dependence of failure intensity on distance to default is at least reasonable.

The likelihood maximization problems (16) and (17), with parameterizations (28) and (29), are solved numerically using a BFGS quasi-Newton method, based on a mixed quadratic-and-cubic line-search procedure. We have tried a range of alternative initial parameter choices to mitigate the risk of achieving only local maxima. In most cases, the search algorithm achieved near convergence within fifteen iterations. The intensity parameter-vector estimates, $\hat{\mu}$ and $\hat{\nu}$, and their estimated asymptotic standard errors,¹⁷ are reported in Table 2. The associated asymptotic covariance matrices are reported in Appendix C.

Parameter estimates for failure intensity are reported in the first column of Table 2. Consistent with the Black-Scholes-Merton model of default, the estimated failure intensity is monotonically decreasing in distance to default. (The estimated standard error implies statistical significance at conventional confidence levels.) For example, consider for illustration a firm whose current failure intensity is 100 basis points (1%) per quarter. Noting that the logarithm of the failure intensity is modeled as linear with respect to the covariates, we see from Table 2 that the estimated marginal sensitivity of this firm's failure intensity is approximately a 44 basis-point increase in quarterly failure intensity per unit reduction in distance to default, and an estimated 46 basis-point increase in quarterly failure intensity per 1% reduction in U.S. personal income growth. As we shall see in Section 3.5, while the magnitude of the impacts of these two covariates on immediate failure likelihoods are comparable, the conditional likelihood of failure more than 1 year ahead is

¹⁷Standard error estimates, shown in parentheses, are asymptotic standard errors obtained from Fisher's information matrix, associated with (14). These asymptotic estimates are within about 1% of bootstrap estimates of finite-sample standard errors obtained by independent resampling firms with replacement.

Table 2: Parameter Estimates for Exit Intensities

Covariates	Failure intensity		Other-exit intensity	
Constant	$\hat{\mu}_0$	-4.2017 (0.2465)	$\hat{\nu}_0$	-3.9855 (0.1254)
Personal income growth (percent)	$\hat{\mu}_1$	-0.4597 (0.1382)	$\hat{\nu}_1$	-0.1711 (0.0564)
Distance to default	$\hat{\mu}_2$	-0.4411 (0.0592)	$\hat{\nu}_2$	0.0137 (0.0129)

estimated to have greater sensitivity with respect to normalized shocks to distance to default than it does to normalized shocks to personal income growth, due to the relatively greater time-series persistence of shocks to distance to default.

With regard to the important roles of both firm-specific leverage and macroeconomic performance for short-term failure probabilities, our results are generally consistent with the prior literature, although we use somewhat different covariates and methods. In particular, distance to default is apparently not a sufficient statistic for conditional failure probabilities, as it would have been in the Black-Scholes-Merton model. Of course, it is conceivable that covariate measurement error, small-sample noise, and mis-specification could have masked the true role of distance to default as a potentially more powerful covariate. Given the limits of empirical modeling of corporate failure prediction, however, it seems prudent to incorporate additional covariates beyond distance to default.

On average across the firms in our study, we attribute approximately 21% of the variation (sample variance) of a firm’s failure intensity to variation of personal income growth, 74% to variation in the firm’s distance to default, and another 5% to covariation between these two covariates.

The second column of Table 2 reports parameter estimates, and estimates of their asymptotic standard errors, for the dependence of other-exit intensity on the covariates. Distance to default does not pass a conventional test of significance as a determinant of the arrival intensity of other exits (most of which, about 80% of those in our sample, were mergers or acquisition). Personal income growth, however, does appear to play a significant role in

the intensity of other exits.¹⁸

As a rough diagnostic of the reasonableness of the overall fit of the model, we compared the actual failure rate in our sample, 0.24% (70 firms out of 28,612 firm-quarters) with the average model-implied expected failure rate during the study period,

$$\frac{\sum_{k=0}^{T-1} \sum_{i \in R(k)} (1 - e^{-\hat{\lambda}_i(k)})}{\sum_{t=0}^T \sum_{i \in R(k)} 1} = 0.23\%, \quad (30)$$

where $R(k)$ is the risk set at quarter k , that is the set of firms operating at quarter k , and $\hat{\lambda}_i(k)$ is the estimated failure intensity of firm i at quarter k . The denominator is, as for the actual failure rate, merely the sample size.

Appendix B reports an estimated model of failure and other-exit intensities that is augmented with two additional covariates: firm-level accounting earnings and firm size. Adding these covariates does not lead to significant changes, from the basic two-covariate model reported in Table 2, in the coefficients representing the dependence of failure intensity on distance to default and personal income growth. (It is perhaps noteworthy, however, that with the addition of earnings and size covariates, the dependence of other-exit intensity on distance to default becomes statistically significant at conventional confidence levels.)

In the end, we have opted to use the basic two-covariate model for estimation of term structures of conditional hazard rates. The four-covariate model does not offer a better average fit (for example, the associated average expected failure probability of the four-covariate model, 20 basis points, is no closer than the 23-basis-point estimate of (30) of the basic model to the sample failure rate of 24 basis points). Firm size and earnings, moreover, are intricately structurally linked in their time-series behavior with distance to default, and incorporating the essence of these structural links into a time-series model for the four covariates seems fraught with mis-specification risk, not to mention loss of parsimony.¹⁹

¹⁸We are not aware of other studies using the similar econometric setting as ours, in particular, duration models, of merger prediction.

¹⁹See Chava and Jarrow (2002) for additional discussion of the relative importance of firm-specific covariates.

3.5 Term Structures of Failure Hazards

We are now in a position to obtain maximum-likelihood estimates, by firm, of the conditional survival and failure probabilities, (1) and (2), for each future time horizon. For the i -th firm in our sample surviving to a given time t , these conditional probabilities, denoted $p(Y_t, D_{it}, s; \psi_i)$ and $q(Y_t, D_{it}, s; \psi_i)$ respectively, depend on the parameter vector $\psi_i = (\mu, \nu, \phi, \kappa_D, v, \theta_{D_i})$ associated with firm i . Under standard technical conditions, the maximum-likelihood estimators of these conditional probabilities at time horizon s are $p(Y_t, D_{it}, s; \hat{\psi}_i)$ and $q(Y_t, D_{it}, s; \hat{\psi}_i)$ respectively, where $\hat{\psi}_i$ is the maximum-likelihood estimator of ψ_i .

In order to illustrate the results more meaningfully, we will report the estimated probability density $q_s(X_t, s; \hat{\psi}_i)$ (partial of $q(\cdot)$ with respect to time horizon s) of the failure time,²⁰ and the estimated failure hazard rate

$$H(Y_t, D_{it}, s; \hat{\psi}_i) = \frac{q_s(Y_t, D_{it}, s; \hat{\psi}_i)}{p(Y_t, D_{it}, s; \hat{\psi}_i)}. \quad (31)$$

We emphasize that this failure hazard rate at time horizon s conditions on survival to time s from both failure and from other forms of exit. (The total hazard rate is, notationally suppressing all arguments of the survival function $p(\cdot)$ except for the time horizon s , given as usual by $-p_s(s)/p(s)$.)

As an illustration, we fix a particular time t , the end of our sample period at the fourth quarter of 2001, fix a particular firm, General Binding Corporation (GBC), calculate GBC's estimated conditional term structure of failure hazard rates, and show how that term structure responds to changes to the business-cycle variable Y_t and to changes in GBC's distance to default, D_{it} . GBC is a natural choice for illustration, given that it has a non-trivial level of credit risk at time t (with a current Moodys rating of B3, set on December 16, 1999), and is a reasonably closely followed firm that existed for 113 quarters, most of our sample period. GBC, based in Illinois, is engaged in the design, manufacture, and distribution of office equipment, related supplies, and laminating equipment and films. Founded in 1947, GBC first appears in the Compustat database at the fourth quarter of 1973. At the end of

²⁰This density is most easily calculated by differentiation through the expectation, as $E\left(\exp\left(-\int_t^{t+s}[\lambda(u) + \alpha(u)] du\right)\lambda(t+s) \mid X_t\right)$, which we compute by Monte-Carlo simulation. We emphasize that this density is "improper" (integrates over all s to less than one) because of other exit events.

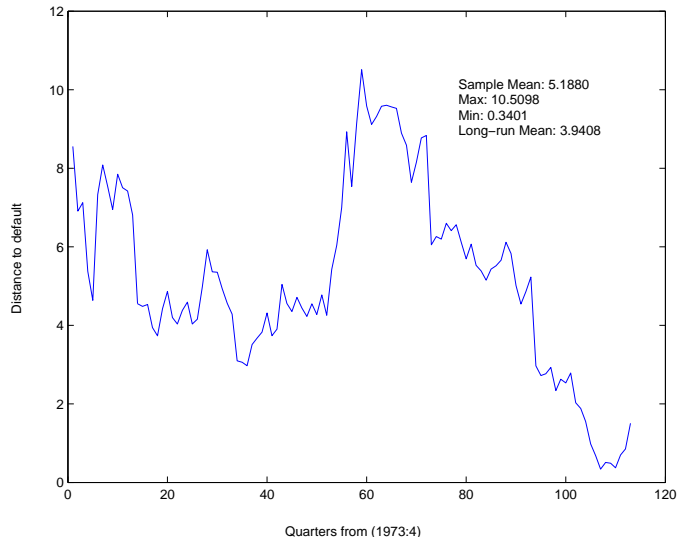


Figure 1: Distance to default of General Binding Corporation, quarters from 1973:4 to 2001:4.

2002, GBC had approximately 4,250 employees, a market capitalization of \$208 million. The sample path of the distance to default of GBC during our sample period is illustrated in Figure 1. The maximum likelihood estimate of θ_{D_i} for GBC is 4.72.

As of the end of 2001, GBC’s estimated term structure of failure hazard rates is shown in Figure 2. The asymptotic 95% confidence intervals of these estimated hazard rates are shown with dashed lines, and obtained by the usual “delta” chain-rule calculation. That is, the variance of $H(Y_t, D_{it}, s; \hat{\psi}_i)$ due to parameter uncertainty is estimated as $H_\psi \Sigma H_\psi^\top$, where H_ψ denotes the partial of H with respect to the parameter vector ψ_i , and where Σ is the asymptotic estimate of the covariance matrix of $\hat{\psi}_i$ reported in Appendix C, which contains additional details on this.

The estimated term structure of failure hazard rates of GBC is downward-sloping because, on the conditioning date t at the end of 2001, the growth rate $Y(t)$ of personal income, at 0.2197%, was well below its estimated long-run mean, 1.8901%, and because GBC’s distance to default, $D_{it} = 1.51$, was also well below its estimated “target,” $\hat{\theta}_{D_i} = 4.72$. The estimated reversion of these covariates toward their respective long-run means causes a reduction in the estimated mean failure arrival rate s quarters ahead, $H(Y_t, D_{it}, s; \hat{\psi}_i)$,

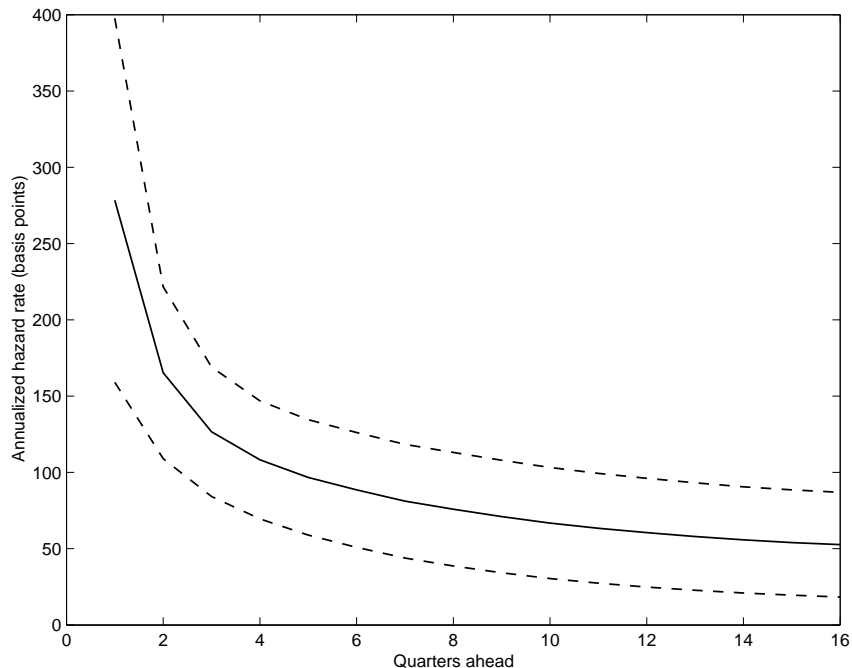


Figure 2: Annualized GBC failure hazard rates, with personal income growth and distance to default measured at the fourth quarter of 2001: when $Y_t = 0.2197\%$ and $D_{it} = 1.51$, shown with dashed lines representing asymptotic estimates of 95% confidence intervals.

conditional on survival to that time, as the time horizon s increases. There is also a small downward effect on this curve due to covariate uncertainty, based on the effect of Jensen's Inequality.

Figures 3 and 4 show the hypothetical effects on GBC's hazard rates of varying the initial conditions $Y(t)$ and D_{it} , respectively, from their estimated long-run means to one-standard deviation above and below their respective long-run means. Specifically, these shifts are the estimated standard deviations²¹ of the stationary distributions of Y and D_i , which are $d_Y = 0.95\%$ and $d_{D_i} = 2.05$, respectively. Comparing Figure 3 and 4 shows that, mainly due to its greater time-series persistence, varying the distance to default has a relatively greater effect than varying U.S. personal income growth on GBC's

²¹For example, with a mean-reversion parameter of κ_Y and an innovation standard deviation of σ_Y , the stationary distribution of Y has a standard deviation whose maximum likelihood estimate is $d_Y = \hat{\sigma}_Y^2 / (1 - (1 - \hat{\kappa}_Y)^2)$.

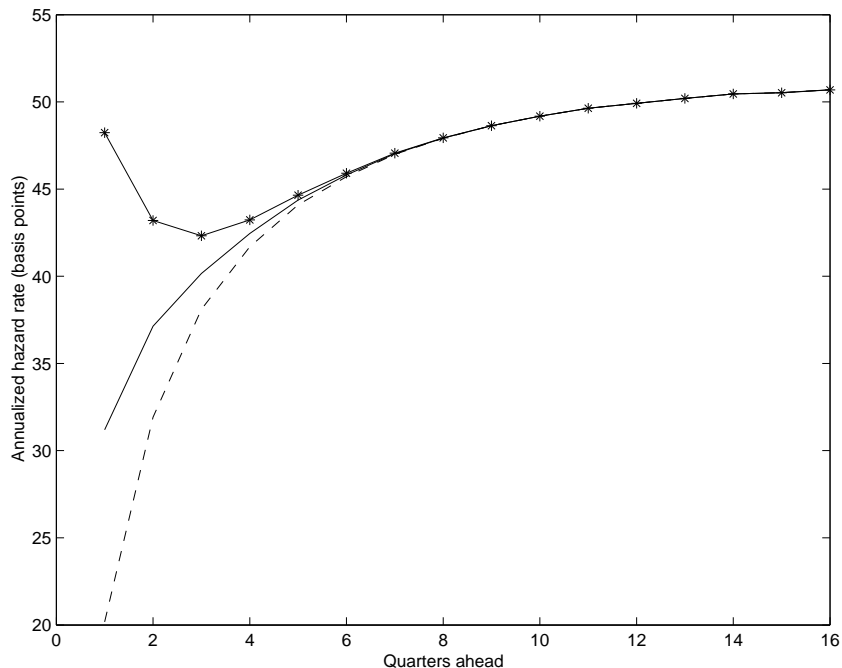


Figure 3: Annualized GBC failure hazard rates, with distance to default set hypothetically at the estimated long-run mean distance to default ($\hat{\theta}_{Di} = 4.72$), and with U.S. personal income growth at three hypothetical levels, the long run mean ($\hat{\theta}_Y = 1.89\%$, solid line); one standard deviation d_Y of its stationary distribution above its long-run mean ($\hat{\theta}_Y + d_Y = 2.84\%$, dashed line) and one standard deviation below its long-run mean ($\hat{\theta}_Y - d_Y = 0.94\%$, solid line with asterisks).

hazard rate for failure more than one year into the future. (We recall that $\hat{\kappa}_Y = 0.6524$, while $\hat{\kappa}_D = 0.1192$.) Figure 5 shows the effects of moving both covariates above, and both covariates below, their respective long-run means.

Figure 6 shows the estimated probability density function of GBC's failure time (solid line) setting the initial conditions for personal income growth and GBC's distance to default at their respective long run means. Figure 6 also shows what this failure-time density would be if one were to ignore the effect of other exits (that is, if one assumes that the other-exit intensity parameter vector ν is zero). For example, GBC obviously cannot itself fail more than one year into the future in the event that it is merged with another firm in less than one year.

In summary, the shape of the term structure of GBC's failure hazard rates

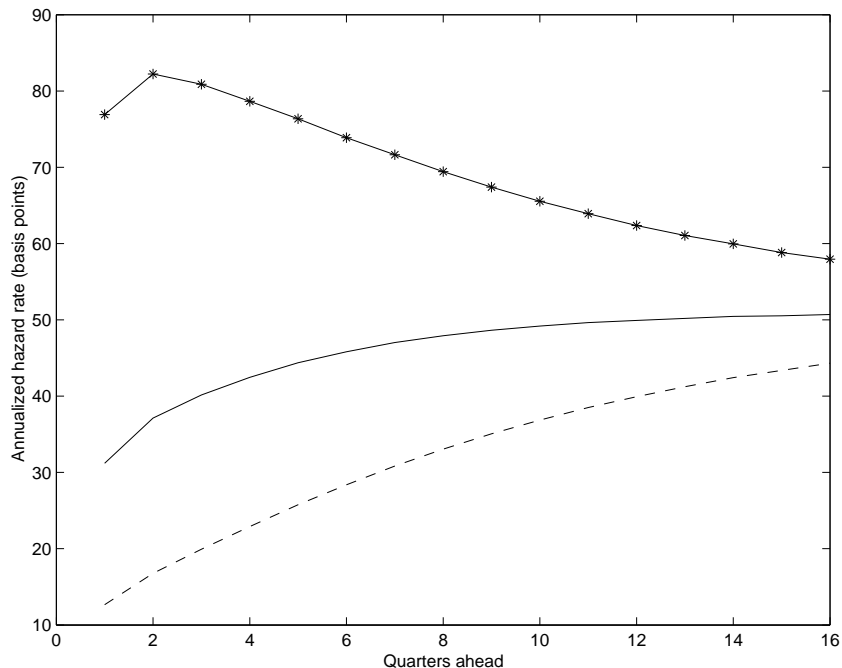


Figure 4: Annualized GBC failure hazard rates, at a hypothetical level of U.S. personal income growth equal to its estimated long-run mean ($\hat{\theta}_Y = 1.89\%$), and with GBC's distance to default at three hypothetical levels: the estimated long-run mean, ($\hat{\theta}_{Di} = 4.72$, solid line); at one standard deviation d_{Di} of its stationary distribution above its long-run mean, ($\hat{\theta}_{Di} + d_{Di} = 7.24$, solid lines with asterisks); and at an estimate of one standard deviation below its long run mean ($\hat{\theta}_{Di} - d_{Di} = 3.13$, dashed line).

for future quarters, conditioning on GBC's current distance to default and on current U.S. personal income growth, clearly reflects the time-series dynamics of these covariates over the coming quarters. The counter-cyclical behavior of failure probabilities is already well documented in such prior studies as Fons (1991), Blume and Keim (1991), Jonsson and Fridson (1996), McDonald and Van de Gucht (1999), Hillegeist, Keating, Cram, and Lundstedt (2003), Chava and Jarrow (2002), and Vassalou and Xing (2003). The main marginal contribution of this paper is the ability to estimate the influence of firm-specific and macro-covariates on the likelihood of corporate failure, not just during the subsequent time period, but also for subsequent quarters into the future.

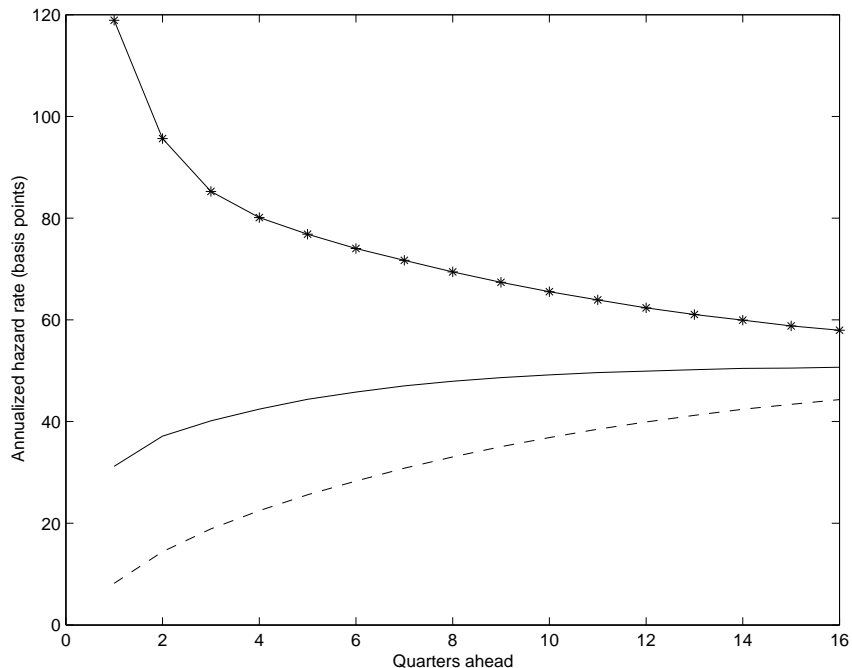


Figure 5: Annualized GBC failure hazard rates. Solid line: covariates initialized at their respective long-run means, $\hat{\theta}_{Di} = 4.72$ and $\hat{\theta}_Y = 1.89$. Dashed line: covariates each initialized one standard deviation (of the respective stationary distributions) above long-run means. Solid line with asterisks: covariates initialized one standard deviation below long-run means.

4 Discussion and Additional Applications

This paper offers an econometric model, and an empirical implementation of this method for the U.S. industrial machinery and instruments sector, for estimating the term structure of corporate failure probabilities over multiple future periods, conditional on firm-specific and macroeconomic covariates. The method, under its probabilistic assumptions, allows one to combine traditional duration analysis of the dependence of event intensities on time-varying covariates with conventional time-series analysis of covariates, in order to obtain maximum-likelihood estimation of multi-period failure probabilities.

Applying this model to data on U.S. firms in the industrial machinery and instrument sector over the years 1971 to 2001, we find that the estimated term

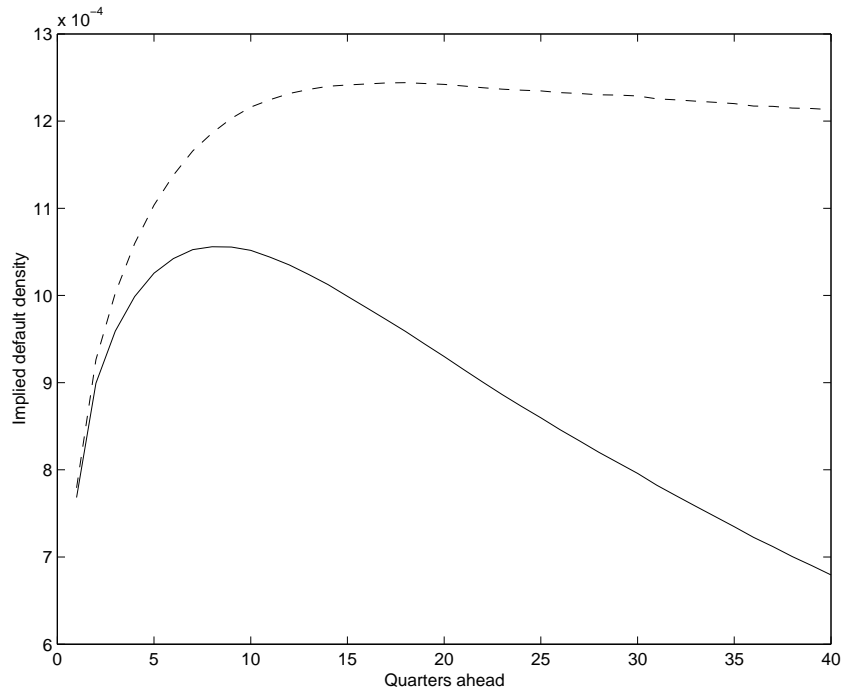


Figure 6: Estimated density of GBC failure time. Solid line: the estimated failure time density $p_s(\hat{\theta}_Y, \hat{\theta}_{Di}, s; \hat{\psi}_i)$ of GBC. Dashed line: the estimated failure-time density obtained by ignoring (setting to zero) the intensity of other exits. Both cases take the covariates at their respective estimated long-run means.

structures of failure hazard rates of individual firms in this sector depend significantly, in level and shape, on the current state of the economy, and on the current leverage of the firm, as captured by the popular volatility-adjusted leverage measure, distance to default. For some firms, variation in distance to default has a greater relative effect on the term structure of future failure hazard rates than does a comparatively sized change in the business-cycle covariate, U.S. personal income growth, especially for the conditional likelihood of failure over long future time periods.

Our methodology could be applied to other settings involving the forecasting of discrete events over multiple future periods, in which the time-series behavior of covariates could play a significant role, for example: mortgage prepayment and default, consumer default, initial and seasoned equity offerings, merger, acquisition, and the exercise of real timing options, such as the

option to change or abandon a technology.

Appendices

A Construction of distance to default

In constructing the distance to default, we follow a recipe similar to those of Vassalou and Xing (2003), Crosbie and Bohn (2002), and Hillegeist, Keating, Cram, and Lundstedt (2003). For a given firm, the distance to default is the number of standard deviations of asset growth by which a firm’s market value of assets exceeds a liability measure. Formally, for a given firm at time t , the distance to default is

$$D_t = \frac{\ln\left(\frac{V_t}{L_t}\right) + (\mu_A - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}, \quad (\text{A.1})$$

where V_t is the market value of the firm’s assets at time t and L_t is a liability measure, defined below, that is often known in industry practice as the “default point”. Here, μ_A and σ_A measure the firm’s mean rate of asset growth and asset volatility, respectively, and T is a chosen time horizon, typically taken to be 4 quarters.

The default point L_t , following the standard established by Moodys KMV (see Crosbie and Bohn (2002), as followed by Vassalou and Xing (2003)), is measured as the firm’s book measure of short-term debt (“Debt in current liabilities”, Compustat item 45), plus one half of its long-term debt (item 51), based on its quarterly accounting balance sheet. If these accounting measures of debt are missing in the Compustat quarterly file, but available in the annual file, we replace the missing data with the associated annual debt data (Compustat items 34 and item 9 for short-term and long term debt, respectively). Of 28,612 firm-quarters in our sample, there are 3,086 firm-quarters in which we use annual debt data to approximate quarterly debt data in this way.

We estimate the assets V_t and volatility σ_A according to a call-option pricing formula, following the theory of Merton (1974), under which equity may be viewed as a call option on the value of a firm’s assets, V_t . In this setting, the market value of equity, as W_t , is the option price at strike L_t and time T to expiration.

We take the initial asset value V_t to be the sum of W_t (end-of-quarter stock price times number of shares outstanding, from CRSP database) and

the book value of total debt (the sum of short-term debt and long-term debt from Compustat). We take the risk-free return r to be the one-year T-bill rate. We solve for the asset value V_t and asset volatility σ_A by iteratively applying the equations:

$$W_t = V_t \Phi(d_1) - L_t e^{-rT} \Phi(d_2) \quad (\text{A.2})$$

$$\sigma_A = \text{sdev}(\ln(V_t) - \ln(V_{t-1})), \quad (\text{A.3})$$

where

$$d_1 = \frac{\ln\left(\frac{V_t}{L_t}\right) + (r + \frac{1}{2}\sigma_A^2)T}{\sigma_A \sqrt{T}}, \quad (\text{A.4})$$

$d_2 = d_1 - \sigma_A \sqrt{T}$, and $\Phi(\cdot)$ is the standard-normal cumulative distribution function, and $\text{sdev}(\cdot)$ denotes sample standard deviation. Equation (A.2) is the call-option pricing formula of Black and Scholes (1973), allowing, through (A.3), an estimate of the asset volatility σ_A . For simplicity, by using (A.3), we avoided the calculation of the volatility implied by the option pricing model (See Crosbie and Bohn (2002) and Hillegeist, Keating, Cram, and Lundstedt (2003) for this alternative approach), but instead estimated σ_A as the sample standard deviation of the time series of asset-value growth, $\ln(V_t) - \ln(V_{t-1})$. A histogram²² of our sample of distances to default is provided in Figure 7.

Our construction of distance to default has the property that, in the theoretical setting of Merton (1974), a firm whose current distance to default is D has a conditional probability (given all available information) of failure in one year of $\Phi(-D)$, where Φ is the cumulative standard-normal distribution function. Figure 8 shows the average relationship in our sample between distance to default and failure rate. For the purpose of this figure, distance to default is “bucketed” into intervals of length 0.25. The denominator for a given bucket is the number of firm-quarters with distance to default in the associated interval; the numerator is the number of failures from that bucket within the subsequent quarter. We did not include in Figure 8 those buckets with distances to default of less than -2 , which constitutes 0.8% of the firm-quarters in our sample. Figure 8 illustrates an average relationship between distance to default and failure frequency that is roughly consistent with our assumption that failure intensity depends exponentially on distance to default, fixing other covariates.

²²Of all 28,612 firm-quarters, four had distances to default larger than 40. These are not shown in Figure 7.

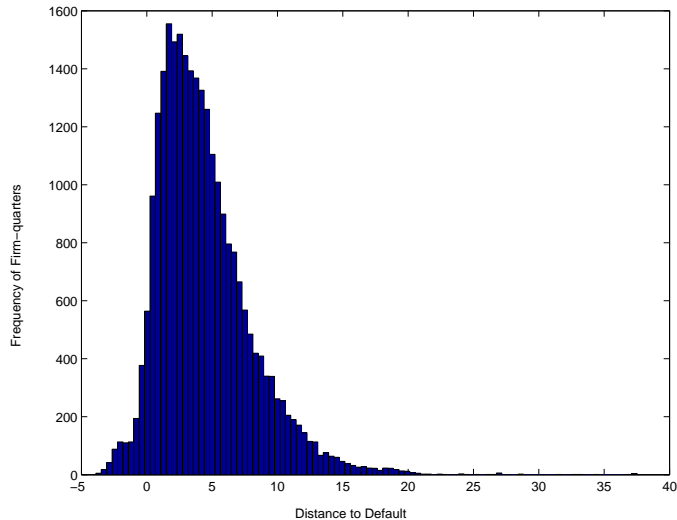


Figure 7: Histogram of distance to default for firm-quarters in the sample.

B Four-covariate intensity model

As a check of the robustness of our model of failure and other-exit intensities, in this appendix, we examine the implications of adding two covariates, earnings and firm size.

Instead of assuming (28) and (29), we specify failure intensity and other-exit intensity to be the form:

$$\Lambda((Y_k, D_{ik}, R_{ik}, S_{ik}); \mu) = e^{\mu_0 + \mu_1 Y_k + \mu_2 D_{ik} + \mu_3 R_{ik} + \mu_4 S_{ik}} \quad (\text{B.1})$$

$$A((Y_k, D_{ik}, R_{ik}, S_{ik}); \nu) = e^{\nu_0 + \nu_1 Y_k + \nu_2 D_{ik} + \nu_3 R_{ik} + \nu_4 S_{ik}}, \quad (\text{B.2})$$

where R_{ik} is the earnings (relative to assets) of firm i in quarter k , and S_{ik} is the size of firm i in quarter k , as explained in Section 3.2. Maximizing a likelihood function analogous to that of the two-covariate model, we obtain the coefficient estimates reported in Table 3. Compared with the results in Table 2, the coefficient estimates for the dependence of failure intensity on distance to default and personal income growth have not changed dramatically. As for the other-exit intensity, although the coefficient of distance to default is not statistically significant different from zero in the two-covariate model, this coefficient is significantly positive in the four-covariate model, the most notable change that we observed.

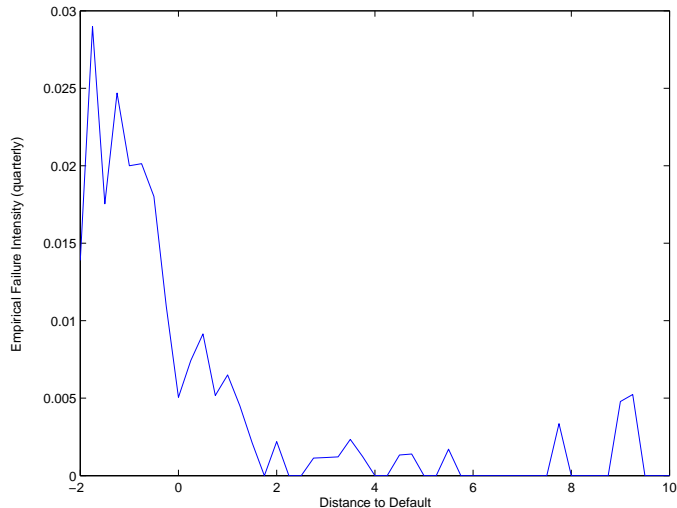


Figure 8: Empirical mapping from distance to default to failure intensity.

The two new covariates, earnings and firm size, show a statistically significant effect on failure intensity and other-exit intensity. Consistent with our expectation, firms with a higher earnings ratio are less likely to go bankrupt, and large firms have a larger chance to survive. It is also notable that high-profit and big firms are less likely to be merged or acquired. We plan to explore the time series of these and other covariates, and further address their effects on failure intensity.

C Asymptotic covariances of MLEs

The confidence intervals plotted in Figure 2 are, as explained in the main text, based on the asymptotic standard errors obtained from the Delta method. For this, we require an estimate of the covariance matrix Σ of the MLE estimator $\hat{\psi}_i$ of the parameters affecting the hazard rates of firm i , which in this case is General Binding Corporation. We let $\psi_i = (\gamma_i, \mu, \nu)$, where γ_i is the vector of parameters of the time-series model for (Y_t, D_{it}) , and where μ and ν parameterize the failure and other-exit intensities, respectively, as in Section 2.

Table 3: Parameter Estimates for Exit Intensities

Covariate	Failure intensity		Other-exit intensity	
Constant	$\hat{\mu}_0$	-3.1718 (0.3566)	$\hat{\nu}_0$	-3.5550 (0.1626)
Personal income growth	$\hat{\mu}_1$	-0.4979 (0.1479)	$\hat{\nu}_1$	-0.1933 (0.0560)
Distance to default	$\hat{\mu}_2$	-0.4473 (0.0699)	$\hat{\nu}_2$	0.0289 (0.0111)
Firm-level earnings	$\hat{\mu}_3$	-0.3752 (0.0976)	$\hat{\nu}_3$	-0.3613 (0.0570)
Firm size	$\hat{\mu}_4$	-0.2910 (0.0785)	$\hat{\nu}_4$	-0.1136 (0.0280)

Asymptotic standard errors are in parentheses.

Fixing Y_t, D_{it} , and s , we write

$$H(Y_t, D_{it}, s; \psi_i) = G(\psi_i). \quad (\text{C.1})$$

The failure probability $p(Y_t, D_{it}, s; \psi_i)$, failure-time density $p_s(Y_t, D_{it}, s; \psi_i)$, survival probability $q(Y_t, D_{it}, s; \psi_i)$, and failure hazard rate $H(Y_t, D_{it}, s; \psi_i)$ are all continuous with respect to the parameter vector ψ_i , by the dominated convergence theorem, using the fact that $e^{-\int_t^{t+s} [\lambda(u)+\alpha(u)] du}$ is strictly positive and bounded by 1, using the continuity of the probability distribution of the covariate process with respect to the parameters, using the monotonicity of the failure and other-exit time intensities with respect to the parameters, and finally using the fact that $\lambda(t+s)$, being the double-exponential of a normal variable, is integrable. Thus, under the consistency assumption that $\hat{\psi}_i$ converges in distribution with sample size to ψ_i , the continuity of $G(\cdot)$ implies that the maximum-likelihood estimator $G(\hat{\psi}_i)$ of $G(\psi_i)$ is also consistent. Moreover, with the addition of differentiability and other technical conditions, $G(\hat{\psi}_i)$ has the asymptotic variance estimate $\nabla G(\psi_i)\Sigma\nabla G(\psi_i)^\top$, where $\nabla G(\cdot)$ is the gradient of G and where

$$\Sigma = \begin{pmatrix} \Sigma_{\gamma_i} & & 0 \\ & \Sigma_{\mu} & \\ 0 & & \Sigma_{\nu} \end{pmatrix} \quad (\text{C.2})$$

is determined by the asymptotic covariance matrices Σ_{γ_i} , Σ_{μ} , and Σ_{ν} of γ_i , μ , and ν , respectively. These asymptotic covariance matrices are obtained by the usual method of inverting the Hessian matrix of the likelihood functions, evaluated at the parameter estimates.

With regard to Σ_{γ_i} , we have already presented the estimated asymptotic covariances of $(\hat{\kappa}_Y, \hat{\theta}_Y, \hat{\sigma}_Y)$ in Section 3.3. The asymptotic variance-covariance matrix of $(\hat{\theta}_{Di}, \hat{\kappa}_D, \hat{v}, \hat{r})$, Σ_{Di} , is part of the full variance-covariance matrix for all firms' θ_{Di} and parameter estimates $(\hat{\kappa}_D, \hat{v}, \hat{r})$. For firm GBC, the estimate of Σ_{Di} is:

$$\hat{\Sigma}_{Di} = \begin{pmatrix} 5931 & 0.3340 & 0.7416 & 2.6375 \\ 0.3340 & 0.0962 & 0.0003 & 0.0012 \\ 0.7416 & 0.0003 & 0.3948 & 0.7936 \\ 2.6375 & 0.0012 & 0.7936 & 3.0292 \end{pmatrix} \times 10^{-4}. \quad (\text{C.3})$$

Combine this and the result for $\hat{\Sigma}_{\phi}$ as in equation (23), we obtain the asymptotic covariance for γ_i as:

$$\hat{\Sigma}_{\gamma_i} = \begin{pmatrix} \hat{\Sigma}_{\phi} & 0 \\ 0 & \hat{\Sigma}_{Di} \end{pmatrix}. \quad (\text{C.4})$$

For the two-covariate model of failure intensity, the asymptotic estimation of covariance matrix corresponding to the parameters (μ_0, μ_2, μ_1) , the coefficients for constant, distance to default and personal income growth, is

$$\hat{\Sigma}_{\mu} = \begin{pmatrix} 0.0607 & -0.0039 & -0.0274 \\ -0.0039 & 0.0035 & -0.0003 \\ -0.0274 & -0.0003 & 0.0191 \end{pmatrix}. \quad (\text{C.5})$$

For the other-exit intensity parameters, the asymptotic covariance matrix of the MLE estimators of the parameters (ν_0, ν_2, ν_1) is

$$\hat{\Sigma}_{\nu} = \begin{pmatrix} 0.0157 & -0.0008 & -0.0055 \\ -0.0008 & 0.0002 & 0.0000 \\ -0.0055 & 0.0000 & 0.0032 \end{pmatrix}. \quad (\text{C.6})$$

References

- Allen, L. and A. Saunders (2002). A Survey of Cyclical Effects in Credit Risk Measurement Models. mimeo, Stern School of Business, New York University.

- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy. *Journal of Finance* 23, 589–609.
- Amemiya, T. (1985). *Advanced Econometrics*. Cambridge, Massachusetts: Harvard University Press.
- Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (1992). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag.
- Behar, R. and K. Nagpal (1999). Dynamics of Rating Transition. Working Paper, Standard and Poor's.
- Berman, S. M. and H. Frydman (1999). Parametric Estimation of Hazard Functions with Stochastic Covariate Processes. *Sankhya, Series A* 61, 174–188.
- Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637–654.
- Blume, M. and D. Keim (1991). Realized Returns and Volatility of Low-Grade Bonds: 1977-1989. *Journal of Finance* 46, 49–74.
- Chava, S. and R. Jarrow (2002). Bankruptcy Prediction with Industry Effects. Working Paper, Cornell University.
- Collin-Dufresne, P. and R. Goldstein (2001). Do Credit Spreads Reflect Stationary Leverage Ratios? Reconciling Structural and Reduced-Form Frameworks. *Journal of Finance* 56, 1929–1958.
- Cox, D. R. and V. Isham (1980). *Point Processes*. New York: Chapman and Hall.
- Cox, D. R. and D. Oakes (1984). *Analysis of Survival Data*. New York: Chapman and Hall.
- Crosbie, P. J. and J. R. Bohn (2002). Modeling Default Risk. Technical Report, KMV, LLC.
- Daley, D. J. and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes*. New York: Springer.
- Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third edition). Princeton, New Jersey: Princeton University Press.
- Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine Processes and Applications in Finance. *Annals of Applied Probability* 13, 984–1053.

- Duffie, D. and D. Lando (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica* 69, 633–664.
- Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68, 1343–1376.
- Duffie, D. and K. Singleton (2003). *Credit Risk*. Princeton, New Jersey: Princeton University Press.
- Efron, B. and R. Tibshirani (1993). *An Introduction To The Bootstrap*. New York: Chapman and Hall.
- Engle, R. F. and J. R. Russell (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* 66, 1127–1162.
- Engle, R. F. and J. R. Russell (2002). Econometric Analysis of Discrete-Valued Irregularly-Spaced Financial Transactions Data. Working Paper, University of California, San Diego.
- Fisher, E., R. Heinkel, and J. Zechner (1989). Dynamic Capital Structure Choice: Theory and Tests. *Journal of Finance* 44, 19–40.
- Fons, J. (1991). An Approach to Forecasting Default Rates. Working Paper, Moody’s Investors Services.
- Helwege, J. and P. Kleiman (1997). Understanding Aggregate Default Rates of High Yield Bonds. *Journal of Fixed Income* 7, 55–62.
- Hillegeist, S. A., E. K. Keating, D. P. Cram, and K. G. Lundstedt (2003). Assessing the Probability of Bankruptcy. Working Paper, Northwestern University.
- Jones, F. (1987). Current Techniques in Bankruptcy Prediction. *Journal of Accounting Literature* 6, 131–164.
- Jonsson, J. and M. Fridson (1996, June). Forecasting Default Rates on High-Yield Bonds. *The Journal of Fixed Income*, 69–77.
- Kalbfleisch, J. and R. Prentice (2002). *The Statistical Analysis of Failure Time Data* (Second edition). Hoboken, New Jersey: John Wiley & Sons.
- Karr, A. (1991). *Point Processes and Statistical Inference* (Second edition). New York: Marcel Dekker.

- Kavvathas, D. (2001). Estimating Credit Rating Transition Probabilities for Corporate Bonds. Working paper, University of Chicago.
- Kealhofer, S. (2003). Quantifying Credit Risk I: Default Prediction. *Financial Analysts Journal*, January–February, 30–44.
- Keenan, S. C., J. Sobehart, and D. T. Hamilton (1999). Predicting Default Rates: A Forecasting Model For Moody’s Issuer-Based Default Rates.
- Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. New York: Cambridge University Press.
- Lando, D. and T. Skødeberg (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations. *Journal of Banking and Finance* 26, 423–444.
- Lane, W. R., S. W. Looney, and J. W. Wansley (1986). An Application of the Cox Proportional Hazards Model to Bank Failure. *Journal of Banking and Finance* 10, 511–531.
- Lee, S. H. and J. L. Urrutia (1996). Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models. *The Journal of Risk and Insurance* 63, 121–130.
- Leland, H. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance* 49, 1213–1252.
- Leland, H. (1998). Agency Costs, Risk Manangement, and Capital Structure. *Journal of Finance* 53, 1213–1242.
- Lennox, C. (1999). Identifying Failing Companies: A Re-Evaluation of the Logit, Probit and DA Approaches. *Journal of Economics and Business* 51, 347–364.
- McDonald, C. G. and L. M. Van de Gucht (1999). High-Yield Bond Default and Call Risks. *Review of Economics and Statistics* 81, 409–419.
- Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29, 449–470.
- Meyer, B. D. (1990). Unemployment Insurance and Unemployment Spells. *Econometrica* 58, 757–782.
- Miller, R. (1981). *Survival Analysis*. New York: John Wiley & Sons.
- Nickell, P., W. Perraudin, and S. Varotto (2000). Stability of Rating Transitions. *Journal of Banking and Finance* 24, 203–227.

- Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research* 19, 109–131.
- Pesaran, M. H., T. Schuermann, B.-J. Treutler, and S. M. Weiner (2003). Macroeconomic Dynamics and Credit Risk: A Global Perspective. Unpublished working paper. Wharton School, Federal Reserve Bank, New York.
- Protter, P. (1990). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business* 74, 101–124.
- Therneau, T. M. and P. M. Grambsch (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag.
- Vassalou, M. and Y. Xing (2003). Default Risk in Equity Returns. Working Paper, Columbia University, forthcoming, *Journal of Finance*.
- Whalen, G. (1991). A Proportional Hazards Model of Bank Failure: An Examination of its Usefulness as an Early Warning Model Tool. In *Federal Reserve Bank of Cleveland Economic Review*, pp. 21–31. 1991.
- Wheelock, D. C. and P. W. Wilson (2000). Why Do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions. *Review of Economics and Statistics* 82, 127–138.