Multi-Period Corporate Default Prediction With Stochastic Covariates[*]

Darrell Duffie, Leandro Saita, and Ke Wang

First Version: August 30, 2003
Current Version: September 19, 2005

Abstract

We provide maximum likelihood estimators of term structures of conditional probabilities of corporate default, incorporating the dynamics of firm-specific and macroeconomic covariates. For U.S. Industrial firms, based on over 390,000 firm-months of data spanning 1979 to 2004, the level and shape of the estimated term structure of conditional future default probabilities depends on a firm's distance to default (a volatility-adjusted measure of leverage), on the firm's trailing stock return, on trailing S&P 500 returns, and on U.S. interest rates, among other covariates. Variation in a firm's distance to default has a substantially greater effect on the term structure of future default hazard rates than does a comparatively significant change in any of the other covariates. Default intensities are estimated to be lower with higher short-term interest rates. The out-of-sample predictive performance of the model is an improvement over that of other available models.

Keywords: default, bankruptcy, duration analysis, doubly stochastic, distance to default. JEL classification: C41, G33, E44

1 Introduction

We provide maximum likelihood estimators of term structures of conditional corporate default probabilities. Our main contribution over prior work is to exploit the time-series dynamics of the explanatory covariates in order to estimate the likelihood of default over several future periods (quarters or years). We estimate our model for U.S.-listed Industrial firms, using over 390,000 firm-months of data on over 2700 firms for the period 1979 to 2004. We find evidence of significant dependence of the level and shape of the term structure of conditional future default probabilities on a firm's distance to default (a volatility-adjusted measure of leverage) and on U.S. interest rates and stock-market returns, among other covariates. Variation in a firm's distance to default has a substantially greater effect on the term structure of future default hazard rates than does a comparatively significant change in any of the other covariates. The shape of the term structure of conditional default probabilities reflects the time-series behavior of the covariates, especially leverage targeting by firms and mean reversion in macroeconomic performance. Out-of-sample predictive performance improves on that of other available models, as detailed in Section 5.

Our model is based on a Markov state vector X_t of firm-specific and macroeconomic covariates that causes variation over time in a firm's default intensity $\lambda_t = \Lambda(X_t)$, which is the conditional mean arrival rate of default measured in events per year. The firm exits for other reasons, such as merger or acquisition, with an intensity $\alpha_t = A(X_t)$. The total exit intensity is thus $\alpha_t + \lambda_t$. We specify a doubly-stochastic formulation of the point process for default and other forms of exit under which the conditional probability of default within s years is

$$q(X_t, s) = E\left(\int_t^{t+s} e^{-\int_t^z (\lambda(u)+\alpha(u))\, du} \lambda(z)\, dz \;\middle|\; X_t \right). \tag{1}$$

This calculation reflects the fact that a firm cannot default at time t if it has already disappeared for some other reason.

While there is a significant prior literature treating the estimation of one-period-ahead default (or bankruptcy) probabilities, for example with logit models, we believe that this is the first empirical study of conditional default probabilities over multiple future time periods that incorporates the time dynamics of the covariates. The sole exception seems to be the practice of treating the credit rating of a firm as though a Markov chain, with ratings

1

transition probabilities estimated as long-term average ratings transition frequencies. 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 default rates (Nickell, Perraudin, and Varotto (2000), Kavvathas (2001), Wilson (1997a), Wilson (1997b)), nor the important effect of prior ratings history (Behar and Nagpal (1999), Lando and Skødeberg (2002)). There is, moreover, significant heterogeneity in the short-term default probabilities of different firms of the same current rating (Kealhofer (2003)). As explained in Section 5, the out-of-sample performance of ratings-based default probabilities are poorer than those of other available models.

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, *(ii)* the determination by banks and bank regulators of the appropriate level of capital to be held by a bank, in light of the credit risk represented by its loan portfolio, especially given the upcoming Basel II accord, under which borrower default probabilities play an explicit role in capital requirements, *(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 default risks of corporations.

Absent a model that incorporates the dynamics of the underlying covariates, one cannot extrapolate prior models of one-quarter-ahead or one-year-ahead default probabilities to longer time horizons. Campbell, Hilscher, and Szilagyi (2005) instead estimate separate logit models for default probabilities at each of a range of time horizons, although not taking advantage of joint consistency conditions across those horizons.[1]

The conditional default probability $q(X_t, s)$ of (1) depends on:

- a parameter vector β determining the dependence of the default 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 path of the underlying state process X determining

---

[1]Philosophov and Philiosophov (2002) estimate a model of default probabilities at various horizons, although not exploiting the time dynamics of the covariates.

default and other exit intensities, exit times are the first event times of independent Poisson processes.[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 estimator of the default probability $q(X_t, s)$, and other properties of the model, such as probabilities of joint default of more than one firm. Because of this decoupling, the resulting estimator of the intensity parameter vector β is the same as that of a conventional competing-risks duration model with time-varying covariates. The maximum likelihood estimator of the time-series parameter vector γ would depend of course on the particular specification adopted for the behavior of the state process X. For examples, we could allow the state process X to have GARCH volatility behavior, to depend on Markov chain "regimes," or to have jump-diffusive behavior. For our specific empirical application, we have adopted a simple Gaussian vector auto-regressive specification.

The doubly-stochastic assumption is overly restrictive to the extent that default or another form of exit by one firm could have an important direct influence on the default 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. The doubly-stochastic property does not fit the data well, according to tests developed by Das, Duffie, Kapadia, and Saita (2005). Our empirical results should therefore be treated with caution. In any case, we show substantially improved out-of-sample performance over prior models. For example, the average accuracy ratio (as defined in Section 5) for one-year-ahead out-of-sample default prediction by our model, over 1993-2003, is 88%. Hamilton and Cantor (2004) report that, for 1999-2003, accuracy ratios based on Moody's[3] credit ratings average 65%, while allowing ratings adjustments for placements on Watchlist and Outlook are 69%, and those based on sorting firms by bond yield spreads average 74%. More details on out-of-sample performance are provided in Section 5.

Our methods also allow estimation of the likelihood, by some future date,

---

[2]One 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.

[3]Krämer and Güttler (2003) report no statistically significant difference between the accuracy ratio of Moodys and Standard-and-Poors ratings in a sample of 1927 borrowers.

of either default *or* a given increase in conditional default probability. This and related transition-risk 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 default of groups of firms, or other properties related to default correlation. In a doubly-stochastic setting, default correlation between firms arises from correlation in their default intensities due to *(i)* common dependence of these intensities on macro-variables and *(ii)* correlation across firms of firm-specific covariates. Our model underestimates default correlation relative to average pairwise sample correlations of default reported in DeServigny and Renault (2002).

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 could include the timing of real options such as technology switch, mortgage prepayment, securities issuance, and labor mobility. We are unaware of previously available econometric methodologies for multi-period event prediction that exploit the estimated time-series behavior of the underlying stochastic covariates.

1.1   Related Literature

A standard structural model of default timing assumes that a corporation defaults when its assets drop to a sufficiently low level relative to its liabilities. For example, the 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 firm's conditional default probability is completely determined by its distance to default, which is the number of standard deviations of annual asset growth by which the asset level (or expected asset level at a given time horizon) exceeds the firm's liabilities. This default 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 default 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. Even in the context of a standard structural default model of this type, however, Duffie and Lando (2001) show that if the distance to default cannot be accurately measured, then a filtering problem arises, and the default intensity depends on the measured distance to de-

4

fault and also on other covariates that may reveal additional information about the firm's conditional default 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, and especially given the need to model the joint time-series behavior of all covariates chosen, the model of default probabilities estimated in this paper adopts a relatively small set of firm-specific and macroeconomic covariates.

Prior empirical models of corporate default probabilities, reviewed by Jones (1987) and Hillegeist et al. (2004), have relied on many types of covariates, both fixed and time-varying. Empirical corporate default analysis originated with Beaver (1966), Beaver (1968a), Beaver (1968b), and Altman (1968), who applied multivariate discriminant analysis. Among the covariates 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. Distance to default 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 default 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 default prediction, using time-independent covariates.[4] These models typically apply a Cox proportional-hazard model. Lee and Urrutia (1996) used a duration model based on a Weibull distribution of default 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 a 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 defaults and call exercises.[5] Related duration analysis by Shumway (2001), Kavvathas (2001), Chava and Jarrow (2004), and Hillegeist, Keating, Cram, and Lundstedt (2004) predict bankruptcy.[6]

Shumway (2001) uses a discrete duration model with time-dependent covariates. Computationally, this is equivalent to a multi-period logit model

---

[4]Whalen (1991) and Wheelock and Wilson (2000) also used Cox proportional-hazard models for bank default analysis.

[5]Meyer (1990) used a similar approach in a study of unemployment duration.

[6]Kavvathas (2001) also analyzes the transition of credit ratings.

with an adjusted-standard-error structure. In predicting one-year default, Hillegeist, Keating, Cram, and Lundstedt (2004) also exploit a discrete duration model. By taking as a covariate the theoretical probability of default implied by the Black-Scholes-Merton's model, based on distance to default, Hillegeist, Keating, Cram, and Lundstedt (2004) find, at least in this model setting, that distance to default does not entirely explain variation in default probabilities across firms. This is supported by Bharath and Shumway (2004) and Campbell, Hilscher, and Szilagyi (2005), who find that in the presence of market leverage and volatility information, among other covariates, distance to default adds relatively little informatiom. Further discussion of the selection of covariates for corporate default prediction may be found in Section 3.2.

Moving from the empirical literature on corporate default prediction to statistical methods available for this task, typical econometric treatments of stochastic intensity models include those of Lancaster (1990) and Kalbfleisch and Prentice (2002).[7] 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 large-sample 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.[8]

---

[7]For 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).

[8]For 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 occured during the previous period, which again does not treat a firm that disappears once it defaults.

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 U.S.-listed Industrial firms, and Section 5 discusses out-of-sample performance.

2.1   Conditional Survival and Default Probabilities

Fixing a probability space $(\Omega, \mathcal{F}, P)$ and an information filtration $\{\mathcal{G}_t : t \geq 0\}$ satisfying the usual conditions,[9] let $X = \{X_t : t \geq 0\}$ be a time-homogeneous 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 default, and the first event time of M, corresponding to exit for some other reason. In our application to U.S.-listed Industrial firms, the portion of exits for reasons other than default is far too substantial to be ignored.

   The main idea is that, so long as the firm has not exited for some reason, its default 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 default 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

---

[9]See Protter (1990) for technical definitions.

σ-algebra generated by

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

The firm's default time is the stopping time $T = \inf\{t : N_t > 0, M_t = 0\}$. We now verify the formula (1) for default probabilities.

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

$$P(\tau > t + s \,|\, \mathcal{F}_t) = p(X_t, s) \equiv E\left(e^{-\int_t^{t+s}(\lambda(u)+\alpha(u))\,du} \,\Big|\, X_t\right), \qquad (2)$$

and the $\mathcal{F}_t$-conditional probability of default by $t + s$ is

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

where $q(X_t, s)$ is given by (1).

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

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

$$P(\tau > t + s \,|\, \mathcal{G}_t) = E\left(e^{-\int_t^{t+s}(\lambda(u)+\alpha(u))\,du} \,\Big|\, X_t\right), \qquad (3)$$

and

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

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 default time T, exploiting the X-conditional independence of M and N is, with the standard abuse of notation,

$$
\begin{aligned}
P(T \in dz \,|\, X) &= P(\inf\{u : N_u \neq N_t\} \in dz, M_z = M_t \,|\, X) \\
&= P(\inf\{u : N_u \neq N_t\} \in dz \,|\, X) P(M_z = M_t \,|\, X) \\
&= e^{-\int_t^z \lambda(u)\,du} \lambda(z)\,dz \; e^{-\int_t^z \alpha(u)\,du} \\
&= e^{-\int_t^z (\alpha(u)+\lambda(u))\,du} \lambda(z)\,dz.
\end{aligned}
$$

8

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

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

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. In our eventual application, however, the intensities depend non-linearly on an affine process X_t, which calls for numerical computation of $p(X_t, s)$ and $q(X_t, s)$. Fortunately, this numerical calculation is done *after* obtaining maximum-likelihood estimates of the parameters of the model.

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 default 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. 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 monthly or less frequently, 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_{it} is constant between periodic observations, a form of time-inhomogeneity that involves only a slight extension of our basic theory of Section 2.1. We continue to measure time continuously, however, in order to take advantage of information on intra-period timing of exits.

9

Extending our notation from Section 2.1, for all i, we let $\Lambda(X_{it}, \beta)$ and $A(X_{it}, \beta)$ denote the default 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 default 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 $Z_1, \ldots, Z_n$ of the n firms. For example, some firms may have different target leverage ratios than others.

We assume that the exit-counting process $(M_1, N_1, \ldots, M_n, N_n)$ of the n firms is doubly-stochastic driven by $X = (X_1, \ldots, X_n)$, in the sense of Section 2.1, so that the exit times $\tau_1, \ldots, \tau_n$ of the n firms are X-conditionally independent, as discussed in Section 1. There is some important loss of generality here, for this implies that the exit of one firm has no direct impact on the default intensity of another firm. Their default times are correlated only insofar as their exit intensities are correlated.

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

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

where

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

is the information set for firm i, 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.

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

Conditional on the current combined covariate vector $Z_k = (Z_{1k}, \ldots, Z_{nk})$, we suppose that Z_{k+1} has a joint density $f(\cdot \mid Z_k; \gamma)$, 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 $K \subset L \subset \{1, \ldots, n\}$ 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 $\tilde{U}_k = \{U_{ik} : i \in R(k)\}$, $S_i(t) = \min(t, S_i)$, $S(t) = (S_1(t), \ldots, S_n(t))$, and likewise define $T_i(t)$ and $T(t)$. Under our doubly-stochastic assumption, the likelihood for the information set $\mathcal{I}_t$ is

$$\mathcal{L}(\mathcal{I}_t; \gamma, \beta) = \mathcal{L}(\tilde{U}, Y, t; \gamma) \times \mathcal{L}(S(t), T(t); Y, \tilde{U}, \beta), \tag{5}$$

where

$$\mathcal{L}(\tilde{U}, Y, t; \gamma) = \prod_{k=0}^{k(t)} f_{R(k+1), R(k)}(Y_{k+1}, \tilde{U}_{k+1} \,|\, Y_k, \tilde{U}_k; \gamma), \tag{6}$$

and

$$\mathcal{L}(S(t), T(t); Y, \tilde{U}; \beta) = \prod_{i=1}^{n} G_{it}(\beta), \tag{7}$$

for

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

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

Because the logarithm of the joint likelihood (5) is the sum of separate terms involving γ and β respectively, we can decompose the overall maximum likelihood estimation problem into the separate problems

$$\sup_{\gamma} \; \mathcal{L}(\tilde{U}, Y, t; \gamma) \tag{8}$$

and

$$\sup_{\beta} \; \mathcal{L}(S(t), T(t); Y, \tilde{U}, \beta). \tag{9}$$

11

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

$$\lambda_{it} = \Lambda(X_{it}; \mu); \qquad \alpha_{it} = A(X_{it}; \nu). \tag{10}$$

(This involves a slight abuse of notation.) This means that the form of dependence of the default 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 (9) reveals that this decoupling assumption allows problem (9) to be further decomposed into the pair of problems

$$\sup_{\mu} \prod_{i=1}^{n} e^{-\int_{t_i^0}^{H_i} \Lambda(X_i(u); \mu) \, du} \left(1_{H_i \neq T_i} + \Lambda(X_i(T_i); \mu) 1_{H_i = T_i} \right) \tag{11}$$

and

$$\sup_{\nu} \prod_{i=1}^{n} e^{-\int_{t_i^0}^{H_i} A(X_i(u); \nu) \, du} \left(1_{H_i \neq S_i} + A(X_i(S_i); \nu) 1_{H_i = S_i} \right). \tag{12}$$

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

Proposition 2. *Solutions γ^* and β^* of the respective maximum-likelihood problems (8) and (9) collectively form a solution to the overall maximum-likelihood problem*

$$\sup_{\gamma, \beta} \mathcal{L}(\mathcal{I}_t; \gamma, \beta). \tag{13}$$

Under the parameter-decoupling assumption (10), solutions μ^* and ν^* to the maximum-likelihood problems (11) and (12), respectively, form a solution $\beta^* = (\mu^*, \nu^*)$ to problem (9).

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 maximium-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 default

probability $q(X_t, s)$ and the survival probability $p(X_t, s)$ are obtained by (1) and (2), respectively, after replacing $\beta = (\mu, \nu)$ and γ with their maximum likelihood estimators, $\hat{\beta}$ and $\hat{\gamma}$.

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 a well-known covariance matrix (Amemiya 1985). In our case, it is apparent that a consistency result would require that both the number n of firms and the number $k(t)$ of periods of data become large in this sense. We defer precise consistency conditions to future research.

3 Empirical Analysis

This section describes our data set, the parameterization of our covariate processes and intensity models, a summary of the properties of our parameter estimates, and some of the substantive conclusions regarding the behavior of conditional term structures of default hazard rates. We are particularly interested in the sensitivity of these term structures of default hazard rates to firm-specific and macroeconomic variables.

3.1 Data

Our sample period is 1979 to 2004. For each firm, short-term and long-term debt data are from quarterly and yearly Compustat files. Short-term debt is estimated as the larger of Compustat items DATA45 and DATA49. Long-term debt is taken from DATA51, while DATA61 provides the number of common shares outstanding, quarterly.[10] The number of common shares

---

[10]For annual data, the corresponding records are DATA34, DATA5, DATA9, and DATA25, respectively. For cases with missing debt data, if the missing value corresponds to year Y, quarter Q, then we complete the data whenever possible as follows. First, we only consider there to be missing data for firm F at year Y, quarter Q if there is debt data for firm F before and after Y, Q. Given that, if there is non-missing quarterly debt data for F in year Y, we take the closest observation to quarter Q, after quarter Q, to complete the dataset. If all quarterly debt observations for F in year Y are missing, we look in the yearly debt file for firm F, year Y. If this observation is not missing we take it to be the debt level for year Y, quarter Q, firm F. In all other cases of missing observations we take

outstanding is combined with Compustat's stock price data (data item PRC) to compute the value of total equity.

Data on the timing of default, merger, other-exit, and bankruptcy are mainly from Moodys Default Risk Service[11] and the CRSP/Compustat database. For cases in which a firm exits our database and no exit reason appears in either of these sources, we refer to Bloomberg's CACS function, SDC, CRSP, and, when necessary, other sources. CRSP/Compustat provides reasons for deletion of firms, and the year and month of deletion (data items AFTNT35, AFTNT34, and AFTNT33, respectively). The reasons for deletion are coded 1-10. (Code 2 is bankruptcy under Chapter 11; Code 3 is bankruptcy under Chapter 7.)

A firm is included in our dataset provided it has a common firm identifier for the Moodys and Compustat databases, and is of the Moodys "Industrial" category. We also restrict attention to firms for which we have at least 6 months of monthly Compustat data. We are left with 2,770 firms, covering 392,404 firm-months of data.

Table 1 shows the number of firms in each of the following exit categories:

- Bankruptcy. An exit is treated for our purposes as a bankruptcy if coded in Moodys database under any of the following categories of events: Bankruptcy, Bankruptcy Section 77, Chapter 10,[12] Chapter 11, Chapter 7, and Prepackaged Chapter 11. A bankrutcy is also recorded if data item AFTNT35 of Compustat is 2 or 3 (for Chapter 11 and Chapter 7, respectively). In some cases, our data reflect bankruptcy exits based on information from Bloomberg and other data sources. Our dataset has 175 bankruptcy exits, although many defaults that eventually led to bankruptcies may not appear as bankruptcy exits if the default was triggered earlier than a bankruptcy, for example by a missed debt payment.

- Default. A default is defined as a bankruptcy, as above, or as any of the following additional default types in the Moodys database: distressed exchange, dividend omission, grace-period default, indenture

---

the data to be missing.

[11]Moodys Default Risk Service provides detailed issue and issuer information on rating, default or bankruptcy, date and type of default (such as bankruptcy, distressed exchange, or missed interest payment), tracking 34,984 firms starting in 1938.

[12]Chapter 10 is limited to businesses engaged in commercial or business activities, not including real estate, whose aggregate debts do not exceed $2,500,000.

modified, missed interest payment, missed principal and interest payments, missed principal payment, payment moratorium, suspension of payments. We also include any defaults recorded in Bloomberg or other data sources.

- Failure. A failure includes any default, as above, and any failures to meet exchange listing requirements, as documented in data from Bloomberg, CRSP, or Compustat.

- Acquisition. Exits due to acquisitions and mergers are as recorded by Moodys, CRSP/Compustat, and Bloomberg.

- Other exits. Some firms are dropped from the CRSP/Compustat database or the Moodys database for other specified reasons, such as reverse acquisition, "no longer fits original file format," leveraged buyout, "now a private company," or "Other" (CRSP/Compustat AFTNT35 codes 4, 5, 6, 9, 10 respectively). We have also temporarily included in this category, from the Moodys database: Cross-default, Conservatorship, Placed under administration, Seized by regulators, or Receivership, although these will fall under "failure exits" in the next revision of the paper. We also include firms that are dropped from CRSP/Compustat for no stated reason (under item AFTNT35). When such a failure to include a firm continues for more than 180 days, we take the last observation date to be the exit date from our dataset. Most of the other exits are due to data gaps of various types.

Exit type	Number
bankruptcy	175
default	495
failure	497
merger-acquisition	872
other	877

Table 1: Number of firm exits of each type.

3.2 Covariates

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

1. The firm's *distance to default,* which, roughly speaking, is the number of standard deviations of quarterly asset growth by which assets exceed a standardized measure of liabilities. As explained in 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 (2004), Crosbie and Bohn (2002), and Hillegeist, Keating, Cram, and Lundstedt (2004), although Bharath and Shumway (2004) point out that default prediction performance is robust to the method by which distance to default is estimated. Details are given in Appendix A.

2. The firm's trailing 1-year stock return.

3. The 3-month Treasury bill rate (in percent).

4. The trailing 1-year return on the S&P 500 index.

We also considered, and rejected for lack of significance in the presence of the above covariates, a number of additional covariates: the U.S. 10-year treasury yield, U.S. personal income growth, U.S. GDP growth rate, average Aaa-to-Baa bond yield spread, the firm's size (in terms of the logarithm of the model-implied assets), and the industry-average distance to default. Prior studies find correlation between macroeconomic conditions and default, using a variety of macroeconomic variables. (See Allen and Saunders (2002) for a survey.) For example, McDonald and Van de Gucht (1999) used quarterly industrial production[13] growth in the U.S. as a covariate for high-yield bond default. Hillegeist, Keating, Cram, and Lundstedt (2004) exploit the national rate of corporate bankruptcies in a baseline-hazard-rate model of

---

[13]If included as an additional covariate, industrial production is marginally significant as a determinant of default intensities, with a coefficient that is approximately twice its standard error. We did not include it as a covariate because of its marginal role and because of the loss in parsimony, particularly with respect to the time-series model.

default. Fons (1991), Blume and Keim (1991), and Jonsson and Fridson (1996) document that aggregate default rates tend to be high in the downturn of business cycles. 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. 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.

Firm-level earnings is a traditional predictor for bankruptcy since Altman (1968). When not appearing together with distance to default, earnings is a significant default covariate in both logit and duration models, as shown by Shumway (2001). Chava and Jarrow (2004), Bharath and Shumway (2004), Beaver, McNichols, and Rhie (2004), and Campbell, Hilscher, and Szilagyi (2005) provide additional discussion of the importance of earnings and other firm-specific accounting covariates.

The lack of significance of firm size as a default covariate is somewhat surprising. For example, large firms are thought to have more financial flexibility than small firms. The statistical significance of size as a determinant of default risk was documented in Shumway (2001).[14]

4 Covariate Time-Series Model

We now specify a particular parameterization of the time-series model for the covariates. Because of the extremely high-dimensional state-vector, which includes the macroeconomic covariates as well as the distance to default and size of each of almost 3000 firms, we have opted for a Gaussian first-order vector auto-regressive time series model, with the following simple structure.

The 3-month and 10-year treasury rates, r_{1t} and r_{2t}, respectively, are

---

[14]Shumway (2001), takes the size covariate to be the logarithm of the firm's stock-market capitalization, relative to the total size of the NYSE and AMEX stock markets. We did not find statistical significance, in the presence of our other covariates, of the logarithm of stock-market capitalization. (The associated standard error is approximately equal to the coefficient estimate.)

modeled by

$$r_{t+1} = r_t + k_r(\theta_r - r_t) + C_r\epsilon_{t+1},$$

where $\epsilon_1, \epsilon_2, \ldots$ are independent standard-normal vectors, C_r is a 2×2 lower-triangular matrix, and the time step is one month. Maximum-likelihood parameter estimates and standard errors are reported in Appendix B.

For the distance to default D_{it} and log-assets V_{it} of firm i, and the trailing one-year S&P500 return, S_t, we assume that

$$\begin{bmatrix} D_{i,t+1} \\ V_{i,t+1} \end{bmatrix} = \begin{bmatrix} D_{it} \\ V_{it} \end{bmatrix} + \begin{bmatrix} k_D & 0 \\ 0 & k_V \end{bmatrix} \left(\begin{bmatrix} \theta_{iD} \\ \theta_{iV} \end{bmatrix} - \begin{bmatrix} D_{it} \\ V_{it} \end{bmatrix} \right) +$$
$$+ \begin{bmatrix} b \cdot (\theta_r - r_t) \\ 0 \end{bmatrix} + \begin{bmatrix} \sigma_D & 0 \\ 0 & \sigma_V \end{bmatrix} \eta_{i,t+1}, \tag{14}$$

$$S_{t+1} = S_t + k_S(\theta_S - S_t) + \xi_{t+1}, \tag{15}$$

where

$$\eta_{it} = A z_{it} + B w_t, \tag{16}$$
$$\xi_t = \alpha_S u_t + \gamma_S w_t,$$

for $\{z_{1t}, z_{2t}, \ldots, z_{nt}, w_t : t \geq 1\}$ that are *iid* 2-dimensional standard-normal, all independent of $\{u_1, u_2, \ldots\}$, which are independent standard normals. The 2×2 matrices A and B have $A_{12} = B_{12} = 0$, and are normalized so that the diagonal elements of $AA' + BB'$ are 1. For estimation, some such standardization is necessary because the joint distribution of η_{it} (over all i) is determined by the 6 (non-unit) entries in $AA' + BB'$ and BB'. Our standardization makes A and B equal to the Cholesky decompositions of AA' and BB', respectively. For simplicity, although this is unrealistic, we assume that ϵ is independent of (η, ξ). Maximum-likelihood parameter estimates, with standard errors, are provided in Appendix B.

A positive mean-reversion parameter k_D for distance to default might be characterized as leverage targeting, by which corporations 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

18

(2001). Empirical evidence of leverage targeting, allowing for delays due to frictional costs in adjusting capital structure, is provided by Leary and Roberts (2004). We assume homogeneity of k_D 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. (Our Monte Carlo tests confirm substantial small-sample bias of MLE estimators for firm-by-firm mean-reversion parameters.) Because the distance to default is constructed by normalizaing for asset volatility, as explained in Appendix A, the Merton theory would imply that σ_D does not vary across firms. In any case, we did not allow σ_D to vary by firm, particularly in light of the advantages of parsimony.

A key question is how to empirically model the targeted distance to default, θ_{iD} of firm i. Despite the arguments that swayed us to assume homogeneity across firms of the mean-reversion and volatility parameters k_D and σ_D, our preliminary analysis showed that assuming a common targeted distance to default θ_{iD} leads to estimated term structures of future default probabilities that rise dramatically for firms that had consistently maintained low leverage during our sample period. Perhaps some firms 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_{iD} firm by firm. The median estimate of θ_{iD} across the 2770 firms in the sample is approximately 3.1, with an inter-quartile range of approximately 1.4 to 4.8. The full cross-sectional distribution of $\hat{\theta}_{iD}$ and their standard errors is illustrated in Appendix B. As a long-run-mean parameter is challenging to pin down statistically in samples of our size, however, the standard errors of our estimates of θ_{iD} are responsible for a significant contribution to the standard errors of our estimated term structures of default probabilities.

The high-dimensional parameter search required an iterative numerical treatment. Appendix B provides portions of the parameter estimates and standard errors that are relevant to other calculations appearing in the paper.

4.1   Default and Other-Exit Intensities

We take the default intensities to be of the proportional-hazards form

$$\Lambda(x; \mu) = e^{\mu_0 + \mu_1 x_1 + \cdots + \mu_n x_n}, \tag{17}$$

for a covariate vector x of firm-specific and macroeconomic variables, and for a parameter vector $\mu = (\mu_0, \mu_1, \ldots, \mu_n)$ common to all firms. The other exit intensities have the same proportional-hazards form. The default intensity parameter estimates and their estimated asymptotic standard errors[15] are also reported in Table 2. The associated asymptotic covariance matrix is reported in Appendix C. For all forms of exit, the estimated standard errors imply statistical significance of all covariates at conventional confidence levels, with the exception of the dependence of the merger-acquisition exit intensity on distance to default.

Consistent with the Black-Scholes-Merton model of default, estimated default intensities are strongly monotonically decreasing in distance to default. For example, the parameter estimate in Table 2 reveals that a 10% reduction in distance to default causes an estimated 11.3% proportional increase in default intensity. As we shall see in Section 4.2, distance to default dominates the other covariates in economic importance when viewed in terms of the impact of a typical (one-standard-deviation) variation of the covariate on the term structure of default probabilities. Figure 1 shows the empirical frequency of default within one year as a function of distance to default (with kernel smoothing), indicating that the exponential dependence in (17) is at least reasonable for this crucial covariate.

Controlling for other covariates, the default intensity is estimated to be sigificantly declining in short-term interest rates. While this runs counter to the role of interest rates in determining the interest expense of corporations (by which higher rates place firms under more financial distress, not less), the sign of the coefficient for the short rate is consistent with the fact that short rates are often increased by the U.S. Federal Reserve in order to "cool down" business expansions. Default intensities are estimated to *increase* in the trailing one-year return of the S& P 500, controlling for other covariates. This could be due to correlation between individual stock returns and S& P 500 stock returns, perhaps the trailing nature of the returns and business-cycle dynamics.

As a rough diagnostic of the reasonableness of the overall fit of the model, one can compare the total predicted number of defaults implied by the estimated default intensity paths, about 471 (which is the integral of the total

---

[15]Standard error estimates, shown in parentheses, are asymptotic standard errors obtained from Fisher's information matrix, associated with (9). These asymptotic estimates are within about 1% of bootstrap estimates of finite-sample standard errors obtained by independent resampling firms with replacement.

Exit type	constant	DTD	return	3-mo. r	SPX
bankruptcy	-3.099	-1.089	-0.930	-0.153	1.074
	(0.198)	(0.062)	(0.141)	(0.037)	(0.489)
default	-2.156	-1.129	-0.694	-0.105	1.203
	(0.113)	(0.036)	(0.075)	(0.021)	(0.289)
failure	-2.148	-1.129	-0.692	-0.106	1.185
	(0.113)	(0.036)	(0.074)	(0.021)	(0.289)
merger	-3.220	0.021	0.310	-0.137	1.442
	(0.098)	(0.013)	(0.050)	(0.014)	(0.241)
other	-2.773	-0.072	0.677	-0.167	0.674
	(0.095)	(0.014)	(0.040)	(0.015)	(0.231)

Table 2: Maximum likelihood estimates of intensity parameters, with parenthetic standard errors. "DTD" is distance to default, "return" is the trailing one-year stock return of the firm, "3-mo. r" is the current 3-month treasury rate, and "SPX" is the trailing one-year return of the S& P 500.

default intensity path shown in Figure 2), with the actual number of defaults during the same period, 495. The out-of-sample predictive power of the estimated model is reviewed in Section 5.

4.2 Term Structures of Default Hazards

We are now in a position to obtain maximum-likelihood estimates, by firm and conditioning date, of the term structure of conditional default probabilities. These are obtained from their theoretical counterparts by substituting parameter estimates into (1). In order to illustrate the results more meaningfully, we will report examples of the estimated probability density $q_s(X_{it}, s)$ (the partial derivative of $q(\,\cdot\,)$ with respect to time horizon s) of the default time,[16] and the estimated default hazard rate

$$H(X_{it}, s) = \frac{q_s(X_{it}, s)}{p(X_{it}, s)},\qquad(18)$$

---

[16]This density is most easily calculated by differentiation through the expectation, as $E\left(e^{-\int_t^{t+s}[\lambda(u)+\alpha(u)]\,du}\lambda(t + s)\,|\,X_t\right)$, which we compute by Monte-Carlo simulation. We emphasize that this density is "improper" (integrates over all s to less than one) because of nondefault exit events.
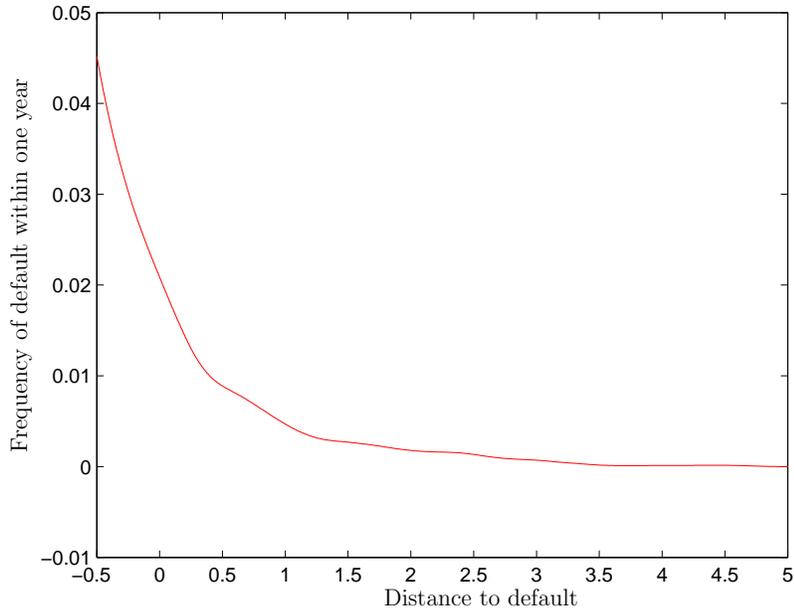
Figure 1: The dependence of empirical default frequency on distance to default.

where $p(X_{it}, s)$ is the estimated survival probability, from (2). The hazard rate $H(X_{it}, s)$ is mean rate of arrival of default at time $t + s$, conditioning on the covariate vector X_{it} at time t, and conditioning as well on the event of survival up until time $t + s$. We emphasize that this default hazard rate at time horizon s conditions on survival to time s from both default and from other forms of exit.[17] If the intensity of default and the intensity of other forms of exit are independent processes, then controlling for survivorship from other forms of exit has no effect on the default hazard rate. In our case, the default intensity and other-exit intensity are correlated since they depend on the same covariates, however the effect of this correlation on the default hazard rates is small. In our illustrative calculations, we account for the other-exit effects associated with merger and acquisition, viewing the other forms of exit as less relevant in practical terms for avoiding default. Even a merger or acquisition need not prevent the future default of a particular

---

[17]The total-exit hazard rate is, notationally suppressing all arguments of the survival function $p(\cdot)$ except for the time horizon s, given as usual by $-p_s(s)/p(s)$.
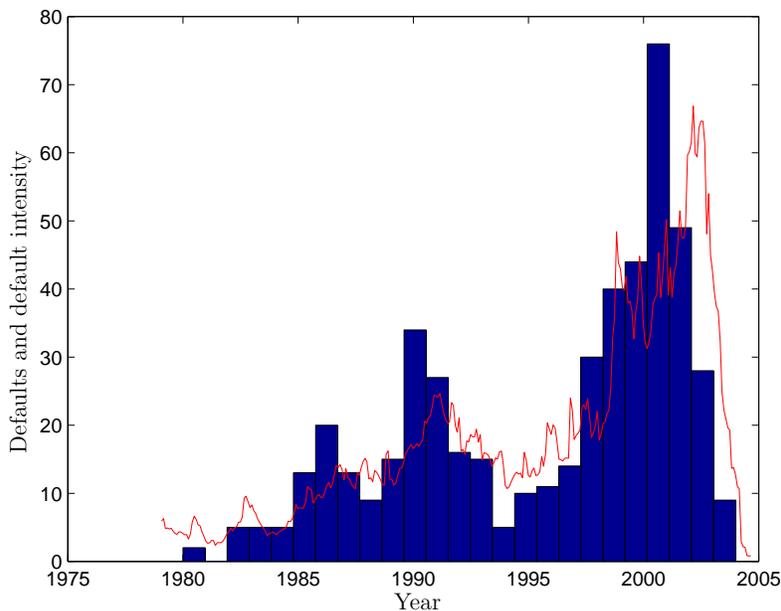
22

Figure 2: The total across firms of default intensities (line), and the number of defaults in each year (bars), 1979-2004.

debt instrument (depending, for example, on whether that debt instrument is paid down immediately, assumed by the new corporation, or exchanged for a new form of debt issued by the new corporation), although of course an acquisition rules out a future bankruptcy by the acquired firm itself.

We consider Xerox as an illustrative firm, and take January 1, 2000 as the conditioning date t. The estimated term structure of Xerox's default hazard rates as of that date is shown in Figure 3. The asymptotic one-standard-error bands of the estimated hazard rates associated with parameter uncertainty are shown with dashed lines, and are obtained by the usual "Delta method," as explained in Appendix C.

The estimated term structure of default hazard rates shown in Figure 3 is downward-sloping mainly because Xerox's distance to default, 0.95, was well below its estimated target, $\hat{\theta}_{iD} = 4.4$ (which has an estimated standard error of 1.4). Other indications that Xerox was in signficant financial distress
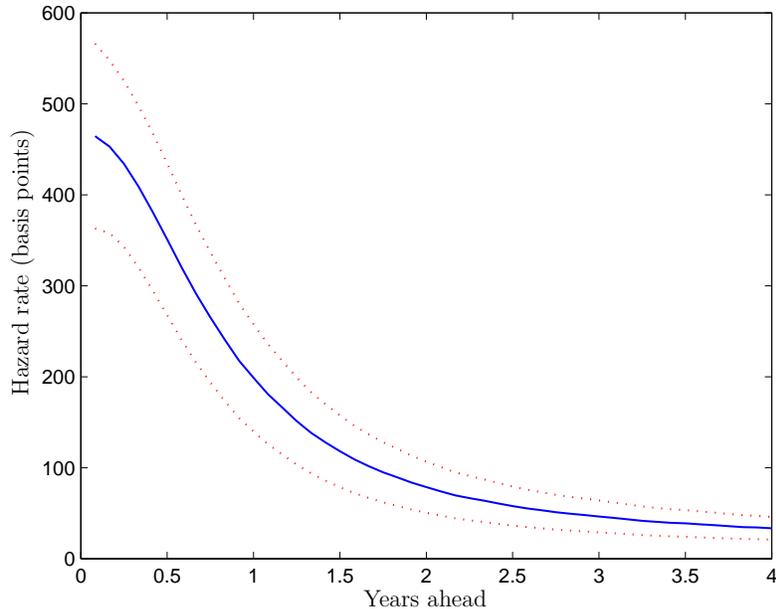
Figure 3: Annualized Xerox default hazard rates as of January 1, 2001 (solid curve), with one-standard-error bands associated with parameter uncertainty (dotted curves).

at this point were its 5-year default swap rate of 980 basis points[18] and its trailing 1-year stock return[19] of -71%.

Figure 4 shows the hypothetical effects on the term structure of Xerox's default hazard rates of one-standard-deviation shifts (from its stationary distribution) of its distance to default, above or below its current level.[20] In terms of both the impact of normalized shocks to default intensity as well time-series presistence, shocks to distance to default have a relatively greater effect on the term structure of Xerox's default hazard rates than do similarly significant shocks to any of the other covariates. Figure 5, for example, shows the analogous effects of a one-standard-deviation shock (from its stationary

---

[18]This CDS rate is an average of quotes provided from GFI and Lombard Risk.

[19]As an intensity covariate, the trailing-stock-return covariate is measured on a continuously compounding basis, and was -124%.

[20]For example, with a mean-reversion parameter of κ_Y and an innovation standard deviation of σ_Y, the stationary distribution of a first-order autoregressive process Y has a standard deviation whose maximum likelihood estimate is $d_Y = \hat{\sigma}_Y^2 / (1 - (1 - \hat{\kappa}_Y)^2)$.
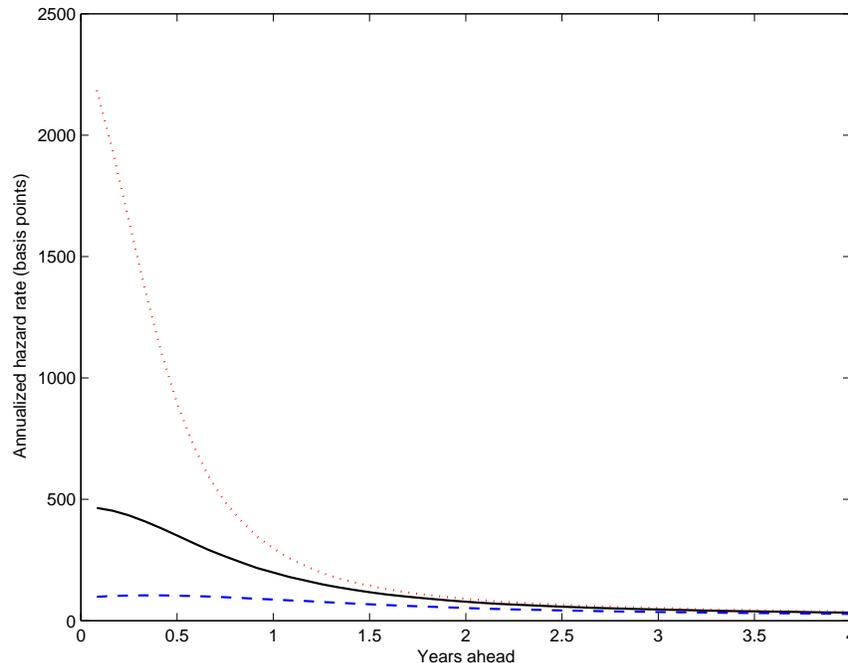
Figure 4: Annualized Xerox default hazard rates as of January 1, 2000 (solid curve), and with distance to default at one standard deviation (1.33) below its current level of 0.95 (dotted curve), and with distance to default at one standard deviation above current level (dashed curve). The trailing S& P 500 return was −8.6%, the trailing one-year stock return of Xerox was −71%, the 3-month treasury rate was 5.8%, and the 10-year treasury yield was 5.2%.

distribution) to the current short-term interest rate. The effect of such a shift in interest rates has a smaller effect than the effect of the analogous shift to Xerox's distance to default, both because of the relative sizes of these shocks, as scaled by the corresponding intensity coefficients, and also because interest rates are less persistent (have a higher mean-reversion rate) than distance to default.

By the beginning of 2004, Xerox's term structure of hazard rates had shrunk dramatically to that shown in Figure 6, mainly because its distance to default had grown to 3.7. The short-maturity hazard rates were also reduced by the Xerox's high trailing one-year stock return of[21] +95%. The

---

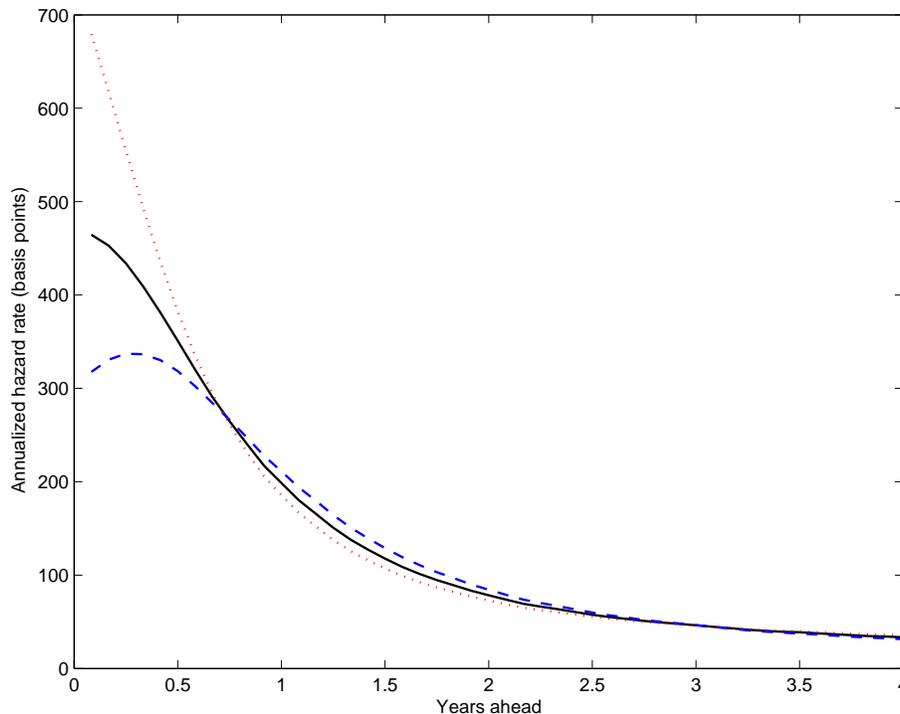[21]On a continuously-compounding basis, the trailing one-year return was 67%.

Figure 5: Annualized Xerox default hazard rates as of January 1, 2000 (solid curve), and with the 3-month treasury rate at one standard deviation (3.6%) below the current level of 5.8% (dotted curve), and at one standard deviation above the current level (dashed curve).

shape of the term structure of hazard rates is also influenced by the non-linear dependence of the covariate-conditional default probabilities on future covariates combined with uncertainty over these future covariates. For example, the term structure of hazard rates can be upward sloping even when the initial covariates are at their long-run means. There are two opposing "convexity" effects here, both due to Jensen's Inequality. First, the intensities are convex with respect to the covariates so the expected future intensities are higher than the intensities evaluated at the expected covariates. Second, the conditional survival probabilities are convex with respect to the path of intensities, so the survival probablities are higher (and default probabilities are lower) than they would be when evaluated at the expected path of the intensities. These competing effects are not canceling.
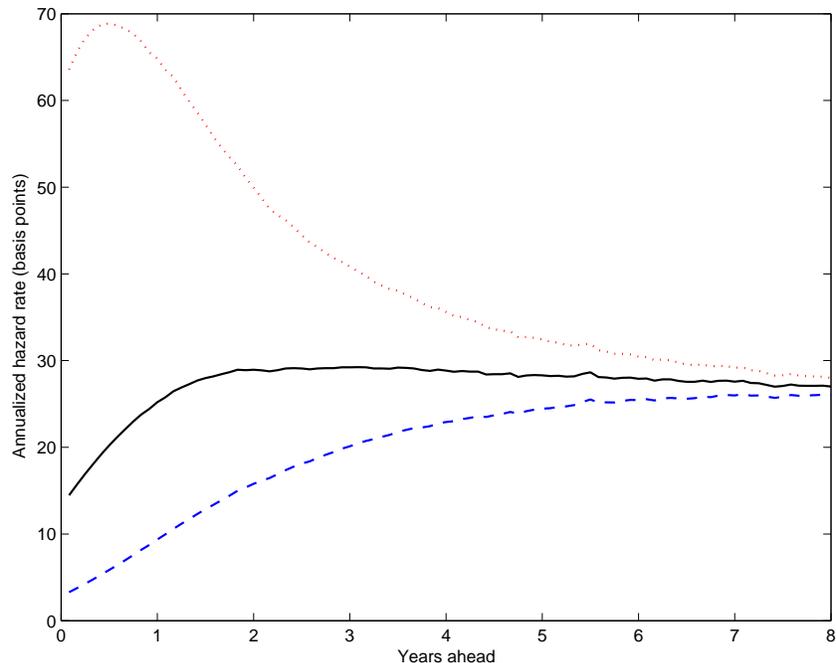
26

Figure 6: Annualized Xerox default hazard rates as of January 1, 2004 (solid curve), and with distance to default at one standard deviation (1.33) below its current level of 3.7 (dotted curve), and with distance to default at one standard deviation above current level (dashed curve).

Figure 7 shows the estimated conditional probability density function of Xerox's default time as of January 1, 2004, and how much larger this default-time density would be if one were to ignore the effect of merger and acquisition (that is, if one assumes that the merger-acquisition intensity parameter vector ν is zero). For example, Xerox 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.

As illustrated above, the shapes of the term structure of Xerox's conditional default hazard rates for future years reflects the time-series dynamics of the covariates. The counter-cyclical behavior of default 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 (2004), Chava and Jarrow (2004), and Vassalou and Xing (2004). A marginal contribution of this paper is the
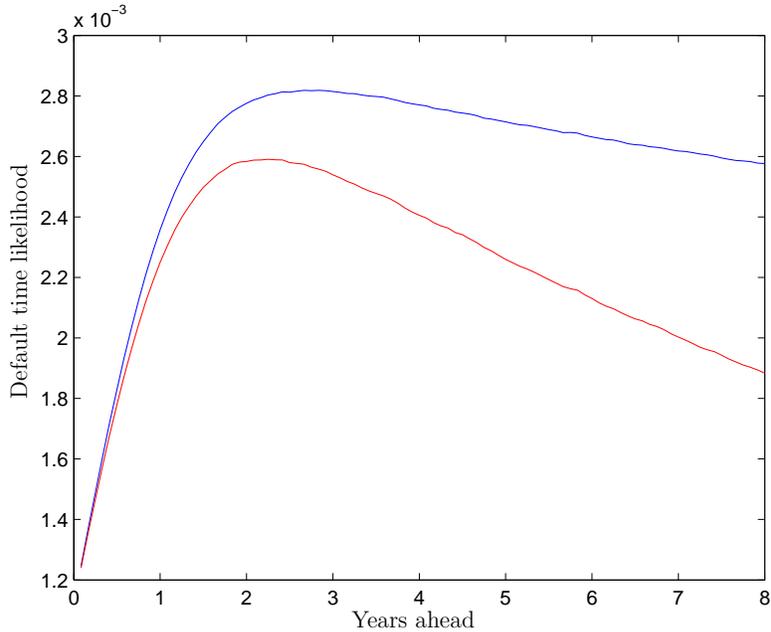
Figure 7: Estimated conditional density of Xerox's default time as of January 1, 2001. Bottom plot: the estimated default time density, incorporating the impact of survival from merger and acquisition. Top plot: the estimated default-time density obtained by ignoring (setting to zero) the intensity of other exits.

ability to estimate the influence of firm-specific and macro-covariates on the likelihood of corporate default, not just during the immediately subsequent time period, but also for a number additional time periods into the future, incorporating the effects of mean reversion, volatilities, and correlation.

5   Out-of-Sample Performance

We now review the out-of-sample ability of our model to sort firms according to estimated default likelihoods at various time horizons. Traditional tools for this purpose are the "power curve" and the associated "accuracy ratio." The one-year power curve illustrated in Figure 8, for example, shows that the "worst" 20% of the firms in the sample, according to estimated one-year default probability, accounted for approximately 92% of the firms that

actually defaulted in the subsequent one-year period, on average over the period 1993 to 2004. These results are out of sample, in that the model used to produce the estimated default probabilities was that estimated from data for 1979 to the end of 1992. Figure 9 provides the analogous results for 5-year-ahead prediction, for 1993 to 2000.
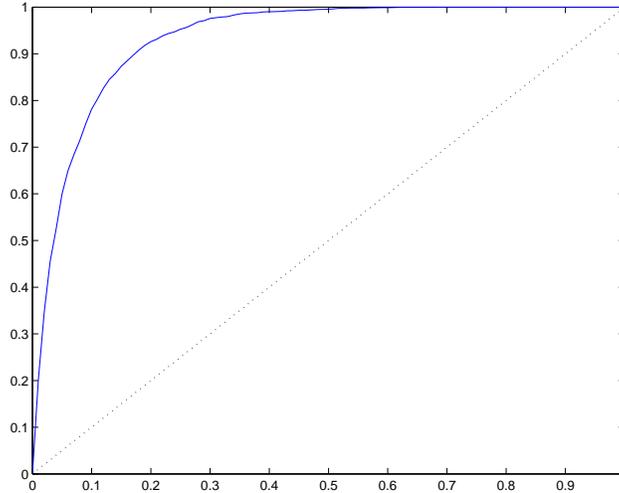

Figure 8: Average out-of-sample power curve for 1-year default prediction, January 1993 to December, 2003.

The accuracy ratio associated with a given power curve is defined to be twice the area between the power curve and the 45-degree line. (The maximum possible accuracy ratio is therefore below 100% by the sample average of the annual default rate.) Table 3 shows average 1-year and average 5-year accuracy ratios for each of the exit types considered, over the post-1993 sample periods. Notably, the accuracy ratios are essentially unchanged when replacing the model as estimated in 1993 with the sequence of models estimated at the beginnings of each of the respective forecast periods. The one-year-ahead out-of-sample accuracy ratio for default prediction from our model, over 1993-2003, is 88%. Hamilton and Cantor (2004) report that, for 1999-2003, the average accuracy ratio for default prediction based on Moody's credit ratings[22] is 65%, while those based on ratings adjustments

---

[22]Krämer and Güttler (2003) report no statistically significant difference between the accuracy ratio of Moodys and Standard-and-Poors ratings in a sample of 1927 borrowers.
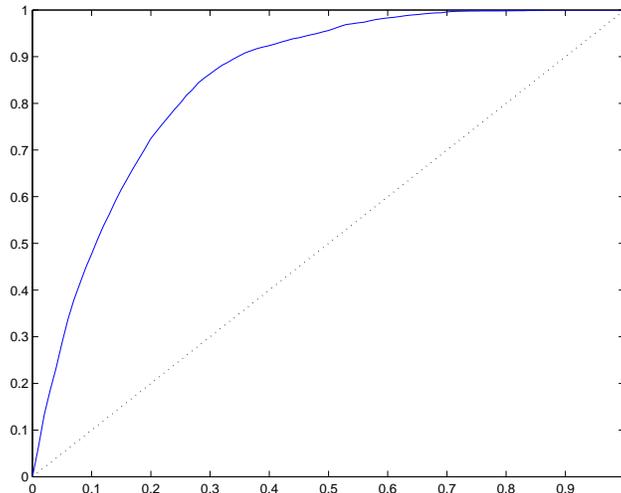
Figure 9: Average out-of-sample power curve for 5-year default prediction, January 1993 to December, 1999.

for placements on Watchlist and Outlook are 69%, and those based on sorting firms by bond yield spreads average 74%.

The out-of-sample accuracy for prediction of merger and acquisition is negative, indicating no out-of-sample power to discriminate among firms regarding their likelihood of being merged or acquired. (Randomly sorting the firms would produce an accuracy ratio, in a large sample, of approximately zero.)

	AR1	AR5
bankruptcy	0.88	0.69
default	0.87	0.70
failure	0.86	0.70
merger	-0.03	-0.01
other exit	0.19	0.11

Table 3: Out-of-sample average accuracy ratios, 1993 to 2004 for 1-year prediction, and 1993 to 2000 for 5-year prediction.

Figures 10 and 11 show the time series of the accuracy ratios for 1-year-

30

out and 5-year-out default prediction throughout the out-of-sample period.
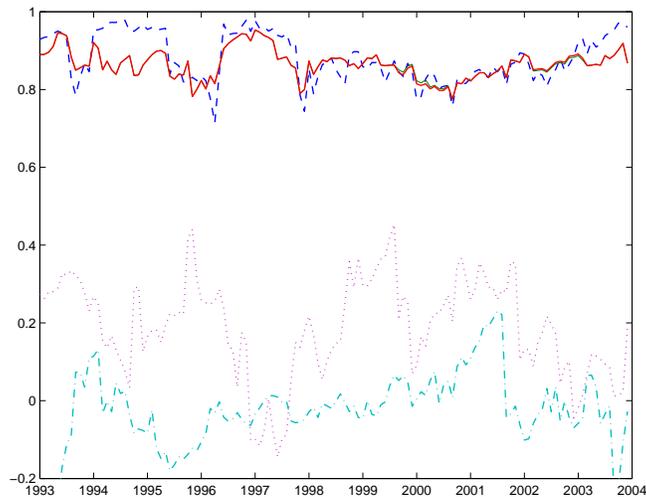

Figure 10: One-year accuracy ratios. Dashed line: bankruptcy. Solid line: default. Dotted line: other exit. Broken dash-dot line: merger.
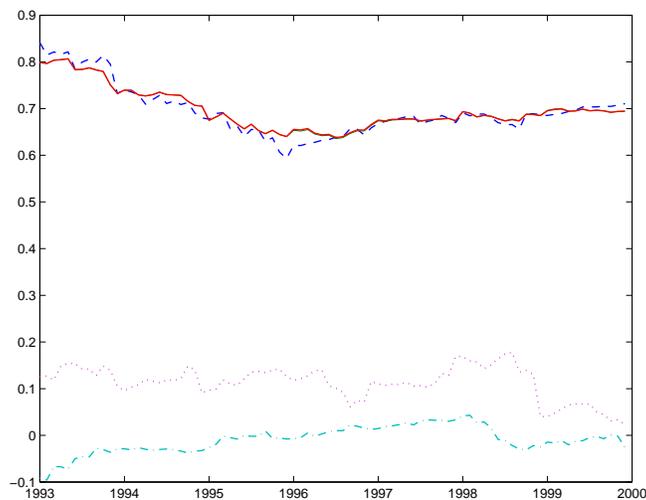

Figure 11: Five-year accuracy ratios. Dashed line: bankruptcy. Solid line: default. Dotted line: other exit. Broken dash-dot line: merger.

Bharath and Shumway (2004) analyze out-of-sample predictive perfor-

mance in terms of the average (over quarters) of the fraction of firms in their sample that default in the subsequent quarter that were sorted by the model into the lowest-quality decile of firms at the beginning of the quarter. For example, for a particular sample, during the period 1991-2003, they report that sorting based on KMV EDFs places approximately 69% of the quarter-ahead defaulting firms in the lowest decile. Sorting based on the more elaborate models developed by Bharath and Shumway (2004) places approximately 77% in the lowest decile. For our own sample and the period 1993-2004, this accuracy measure rises to 94%.

The model of Beaver, McNichols, and Rhie (2004) based on accounting ratios places 80.3% of the year-ahead defaulters in the lowest two deciles, out of sample, for the period 1994-2002. Consistent with the important role of market variables discovered by Shumway (2001), when Beaver, McNichols, and Rhie (2004) add stock-market variables in combination with accounting ratios, this measure goes up to 88.1%, and when they further allow their model coefficients to adjust over time, this measure rises to 92%, which is roughly the measure obtained for our model and data.

6   Concluding Remarks

This paper offers an econometric method for estimating term structures of corporate default probabilities over multiple future periods, conditional on firm-specific and macroeconomic covariates. We also provide an empirical implementation of this method for the U.S.-listed Industrial firms. The method, under its 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 survival probabilities.

Applying this method to data on U.S.-listed Industrial firms over 1979 to 2004, we find that the estimated term structures of default hazard rates of individual firms in this sector depend significantly, in level and shape, on the current state of the economy, and especially on the current leverage of the firm, as captured by distance to default, a volatility-adjusted leverage measure that is popular in the banking industry.

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.

Our model also allows estimates of portfolio credit risk, as it provides maximum-likelihood estimates of joint probabilities of default. For a given maturity T, the default-event correlation between firms i and j is the correlation between the random variables $1_{\{\tau(i)<T\}}$ and $1_{\{\tau(j)<T\}}$. These correlations can be calculated by using the fact that, in a doubly-stochastic framework, for stopping times $\tau(A)$ and $\tau(B)$ that are the first jump times of counting processes with respective intensities λ_A and λ_B, the probability of joint survival to time T is

$$E(1_{\{\tau(A)>T\}}1_{\{\tau(B)>T\}}) = P(\tau(A) > T, \tau(B) > T) = E\left(e^{-\int_0^T (\lambda_A(t)+\lambda_B(t))\, dt}\right).$$

Unfortunately, our estimated model implies unrealistically low estimates of default correlation, compared to the sample correlations reported by DeServigny and Renault (2002). This is a topic of separate ongoing research.

We conclude by comparing our model with that of structural models of default, for example Merton (1974). The main distinctions between the two modeling approaches are the nature of the event that triggers default, and of course the empirical fit. Our assumptions about the asset process and distance-to-default process include those of the Merton model as a special case, and would be identical to those of the Merton model if we take the particular case of no leverage targeting. That is, removing mean reversion, our distance to default process is a Brownian motion, just as in Merton (1974), and our asset process is a geometric Brownian motion, just as in Merton. With regard to what triggers default, our model is quite different from structural models, including Merton (1974) as well as first-passage structural models, such as those of Fisher, Heinkel, and Zechner (1989) and Leland (1994). The structural models apply a solvency test, regarding whether the distance to default falls below some barrier, that is in some cases determined endogenously. Our model assumes instead that, at each "small" time period, default occurs (or not) at random, with a probability that depends on the current distance to default and other explanatory variables. It is known (for example, Duffie and Lando (2001)) that these structural models produce highly unrealistically shaped term structures of default probabilities, given the continuity properties of Brownian motion and the undue precision with which distance to default is assumed to be measured. In particular, the associated default

probabilities are extremely small for maturities of roughly two years or less (at typical parameters), even for low-quality firms. Extensions of these structural models with imperfectly observed leverage (or with significant jumps in leverage) and with leverage targeting (as suggested by Collin-Dufresne and Goldstein (2001)) would have more realistic term structures of default probabilities. Theory as well as our empirical results suggest that enriching structural models with additional state variables (beyond the distance to default), such as macro-economic variables, could lead to an improved for such extended structural models. The evidence that we have presented here may provide some clues for future research in this direction.

A Construction of Distance to Default

This appendix explains how we construct the distance to default covariate, following a recipe similar to those of Vassalou and Xing (2004), Crosbie and Bohn (2002), Hillegeist, Keating, Cram, and Lundstedt (2004), and Bharath and Shumway (2004). For a given firm, the distance to default is, roughly speaking, 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

$$D_t = \frac{\ln\left(\frac{V_t}{L_t}\right) + \left(\mu_A - \frac{1}{2}\sigma_A^2\right)T}{\sigma_A\sqrt{T}}, \tag{19}$$

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 (2004)), is measured as the firm's book measure of short-term debt, plus one half of its long-term debt (Compustat item 51), based on its quarterly accounting balance sheet. We have measured short term debt as the larger of Compustat items 45 ("Debt in current liabilities"), and 49 ("Total Current Liabilities"). 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.

We estimate the assets V_t and volatility σ_A according to a call-option pricing formula, following the theory of Merton (1974) and Black and Scholes (1973), 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, 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 the 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:

$$W_t = V_t\Phi(d_1) - L_t e^{-rT}\Phi(d_2) \tag{20}$$
$$\sigma_A = \text{sdev}\left(\ln(V_t) - \ln(V_{t-1})\right), \tag{21}$$

where

$$d_1 = \frac{\ln\left(\frac{V_t}{L_t}\right) + (r + \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}, \tag{22}$$

$d_2 = d_1 - \sigma_A\sqrt{T}$, and $\Phi(\cdot)$ is the standard-normal cumulative distribution function, and $\text{sdev}(\cdot)$ denotes sample standard deviation. Equation (20) is a variant of the call-option pricing formula of Black and Scholes (1973), allowing, through (21), an estimate of the asset volatility σ_A. For simplicity, by using (21), we avoided the calculation of the volatility implied by the option pricing model (as in Crosbie and Bohn (2002) and Hillegeist, Keating, Cram, and Lundstedt (2004)), but instead estimated σ_A as the sample standard deviation of the time series of asset-value growth, $\ln(V_t) - \ln(V_{t-1})$.

B   Time-Series Parameter Estimates

For the 2-factor interest rate model parameters, our maximum likelihood parameter estimates, with standard errors in subscripted parentheses, are

$$k_r = \begin{pmatrix} 0.03_{(0.026)} & -0.021_{(0.030)} \\ -0.027_{(0.012)} & 0.034_{(0.014)} \end{pmatrix},$$

$$\theta_r = \begin{pmatrix} 3.59_{(4.08)} \\ 5.47_{(3.59)} \end{pmatrix},$$

and

$$C_r = \begin{pmatrix} 0.5639_{(0.035)} & 0 \\ 0.2247_{(0.026)} & 0.2821_{(0.008)} \end{pmatrix},$$

where θ_r is measured in percentage points.

Joint maximum likelihood estimation of equations (14), (15), and (16), simultaneously across all firms i in $\{1, \ldots, n\}$ gives the parameter estimates (with standard errors in subscripted parentheses):

$$b = \begin{pmatrix} 0.0090_{(0.0021)} & -0.0121_{(0.0024)} \end{pmatrix}'$$

$$k_D = 0.0355_{(0.0003)}$$

$$\sigma_D = 0.346_{(0.0008)}$$

$$k_v = 0.015_{(0.0002)}$$

$$\sigma_v = 0.1169_{(0.0002)}$$

$$AA' + BB' = \begin{pmatrix} 1 & 0.448_{(0.0023)} \\ 0.448_{(0.0023)} & 1 \end{pmatrix}$$

$$BB' = \begin{pmatrix} 0.0488_{(0.0038)} & 0.0338_{(0.0032)} \\ 0.0338_{(0.0032)} & 0.0417_{(0.0033)} \end{pmatrix}$$

$$k_S = 0.1137_{(0.018)}$$

$$\alpha_S = 0.047_{(0.0019)}$$

$$\theta_S = 0.1076_{(0.0085)}$$

$$\gamma_S = \begin{pmatrix} 0.0366_{(0.0032)} & 0.0134_{(0.0028)} \end{pmatrix}.'$$

Figure 12 shows the cross-sectional distribution of estimated targeted distance to default, $\hat{\theta}_{iD}$, with standard errors.
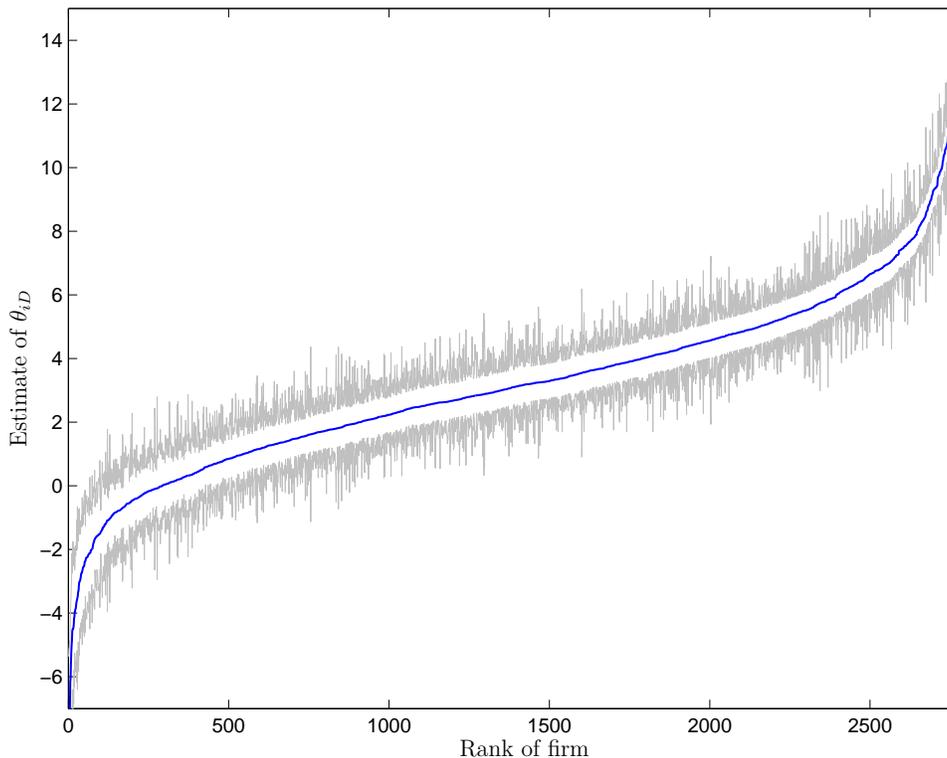

Figure 12: Cross-sectional distribution of estimated targeted distance to default, $\hat{\theta}_{iD}$, sorted, with (vertically) one-standard-error bands for each.

C   Delta-based Standard Errors

The confidence intervals plotted in Figure 3 are 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 portion of the parameter vector ψ_i affecting the hazard rates of firm i, which in this case is Xerox. We let $\psi_i = (\gamma_i, \mu, \nu)$, where γ_i is the vector of parameters of the time-series model for X_{it}, and where μ and ν parameterize the default and other-exit intensities, respectively, as in Section 2.

Fixing $X_{it} = (D_{it}, V_{it}, S_t, r_t)$ and the time horizon s, we write

$$H(X_{it}, s; \psi_i) = G(\psi_i). \tag{23}$$

The default probability $q(X_{it}, s; \psi_i)$, default-time density $q_s(X_{it}, s; \psi_i)$, survival probability $p(X_{it}, s; \psi_i)$, and default hazard rate $H(X_{it}, s; \psi_i)$ 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 default and other-exit time intensities with respect to the parameters, and finally using the fact that $\lambda(t + s)$ is the double-exponential of a normal variable. 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 $\nabla G(\psi_i) \Sigma \nabla G(\psi_i)'$, where $\nabla G(\,\cdot\,)$ is the gradient of G and where

$$\Sigma = \begin{pmatrix} \Sigma_{\gamma_i} & & 0 \\ & \Sigma_{\mu} & \\ 0 & & \Sigma_{\nu} \end{pmatrix} \tag{24}$$

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.

For example, the asymptotic estimate of the covariance matrix of the MLE estimators of the default intensity parameters $(\mu_0, \mu_2, \mu_1, \mu_2, \mu_3, \mu_4)$, corresponding to the constant, distance to default, same-firm trailing stock return, 3-month Treasury rate, and trailing S& P 500 return, respectively, is

$$\hat{\Sigma}_{\mu} = \begin{pmatrix} 0.0128 & 0.0004 & 0.0029 & -0.0019 & 0.0034 \\ 0.0004 & 0.0013 & -0.0011 & -0.0001 & -0.0015 \\ 0.0029 & -0.0011 & 0.0056 & -0.0000 & -0.0025 \\ -0.0019 & -0.0001 & -0.0000 & 0.0004 & -0.0022 \\ 0.0034 & -0.0015 & -0.0025 & -0.0022 & 0.0836 \end{pmatrix}. \tag{25}$$

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.

Beaver, B. (1966). Financial Ratios as Predictors of Failure. *Empirical Research in Accounting: Selected Studies, Supplment to Journal of Accounting Research Autumn*, 91–101.

Beaver, B. (1968a). Alternative Accounting Measures as Predictors of Failure. *Accounting Review January*, 113–122.

Beaver, B. (1968b). Market Prices, Financial Ratios, and the Prediction of Failure. *Journal of Accounting Research Autumn*, 170–192.

Beaver, W., M. McNichols, and J.-W. Rhie (2004). Have Financial Statements Become Less Informative? — Evidence from the Ability of Financial Ratios to Predict Bankruptcy. Working Paper, Stanford University.

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.

Bharath, S. and T. Shumway (2004). Forecasting Default with the KMV-Merton Model. Working Paper, University of Michigan.

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.

Campbell, J., J. Hilscher, and J. Szilagyi (2005). In Search of Distress Risk. Working Paper, Harvard University.

Chava, S. and R. Jarrow (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance 8*, 537–569.

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.

Das, S. R., D. Duffie, N. Kapadia, and L. Saita (2005). Common Failings: How Corporate Defaults are Correlated. Working Paper, Stanford University.

DeServigny, A. and O. Renault (2002). Default Correlation: Empirical Evidence. Working Paper, Standard and Poor's.

Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third edition). Princeton, New Jersey: Princeton University Press.

Duffie, D. and D. Lando (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica 69*, 633–664.

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.

Hamilton, D. and R. Cantor (2004). Ratings Transitions and Default Rates Conditionaed on Outlooks. *The Journal of Fixed Income*, September, 54–70.

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 (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies 9*, 5–34.

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. Working Paper, Moody's Investors Services.

Krämer, W. and A. Güttler (2003). Comparing the Accuracy of Default Predictions in the Ratings Industry: The Case of Moody's vs. S& P. Working Paper, Statistics Department, University of Dortmund.

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.

Leary, M. and M. Roberts (2004). Do Firms Rebalance their Capital Structures? Working Paper, Duke University.

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 Manamgement, and Capital Structure. *Journal of Finance 53*, 1213–1242.

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. Working paper. Wharton School, Federal Reserve Bank, New York, forthcoming in *Journal of Money, Credit, and Banking.*

Philosophov, L. and V. Philiosophov (2002). Corporate Bankruptcy Prognosis: An Attempt at a Combined Prediction of the Bankruptcy Event and Time Interval of its Occurrence. *International Review of Financial Analysis 11*, 375–406.

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 (2004). Default Risk in Equity Returns. *Journal of Finance 59*, 831–868.

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.

Wilson, T. (1997a). Portfolio Credit Risk, I. *RISK 10*, September, 111–117.

Wilson, T. (1997b). Portfolio Credit Risk, I. *RISK 10*, October, 56–61.