Multiperiod Corporate Default Prediction – A Forward Intensity Approach

Jin-Chuan Duan, Jie Sun and Tao Wang*

(First Version: July 14, 2010; This Version: March 17, 2011)

Abstract

A forward intensity approach for the prediction of corporate defaults over different future periods is proposed. Maximum pseudo-likelihood analysis based on this new approach is then conducted on a large sample of the US industrial and financial firms spanning the period 1991-2009 on a monthly basis. Several frequently used factors and firm-specific attributes are shown to be useful for prediction at both short and long horizons. The prediction is very accurate for shorter horizons. The accuracy deteriorates somewhat when the horizon is increased to two or three years, but its performance still remains reasonable. The forward intensity model is also amenable to aggregation, which allows analysts to assess default behavior at the portfolio and/or economy level.

Keywords: default, bankruptcy, forward intensity, maximum pseudo-likelihood, forward default probability, cumulative default probability, accuracy ratio.

*Duan is with Risk Management Institute and Department of Finance, National University of Singapore. Email: bizdjc@nus.edu.sg. Sun is with Risk Management Institute, National University of Singapore. Email: rmisj@nus.edu.sg. Wang is with Department of Finance, National University of Singapore. Email: wangtao@nus.edu.sg.

1 Introduction

Understanding term structure of default probabilities is critical to credit risk management, macro policy making and financial regulation. Firms may have totally different short-term and long-term credit risk profiles due to their debt structures, liquidity positions and other attributes. Major credit rating agencies usually provide both short-term and long-term credit ratings for corporates. However, the academic literature has been lacking as far as the term structure of defaults is concerned. Credit risk modeling can be grouped into two large categories – structural and reduced-form approaches. We tackle the issue of default term structures using an reduced-form approach.

The first generation of reduced-form models dates back to Beaver (1966, 1968) and Altman (1968). These studies mainly relied on discriminant analysis whose output is credit scores which offer only ordinal rankings. The second generation of reduced-form models, e.g., Ohlson (1980) and Zmijewski (1984), mostly employed binary response models such as logit and probit regressions. Such methods assess a firm's likelihood of default in the next period but remain silent for default prediction beyond one period. In a recent paper, Campbell, *et al* (2008) employed a multiple logit model to predict bankruptcy for different time horizons. Recent development in reduced-form credit risk modeling is dominated by duration analysis, such as Shumway (2001), Chava and Jarrow (2004), Hillegeist, *et al* (2004). Duffie, *et al* (2007) proposed a doubly stochastic Poisson intensity approach to default modeling in which state variables governing Poisson intensities are assumed to follow a specific high-dimensional time-series dynamic.[1] Specifying the time dynamics of the state variables is simply for the purpose of multiperiod default prediction.

In this paper, we propose a new reduced-form approach based on a forward intensity construction to estimate a firm's default probabilities for different periods ahead. Our construction takes into account both defaults/bankruptcies and other types of firm exits such as mergers and acquisitions. Our method can estimate forward default probabilities and cumulative default probabilities for longer than one future period (month, quarter or year). We can estimate term structure of default probabilities solely using the known data at the time of performing prediction, and can circumvent the difficult task of specifying time dynamics for covariates. Our forward intensity approach can be implemented by maximum pseudo-likelihood estimation. Particularly, the pseudo-likelihood function is decomposable to independent components, making it less numerically intensive in estimation. Moreover, the nature of the pseudo-likelihood function makes the parameter estimations corresponding to different forward periods non-sequential so that the numerical implementation of the model for multiple period is easily parallelizable.

[1]Duffie, et al (2007) employed two state variables common to all firms and two specific to individual firms. If a sample contains 10,000 firms, the dimension of the state variables will become 20,002.

Our empirical analysis uses a large sample of the US listed companies (both industrial and financial) covering more than 12,000 firms and over 1,000,000 firm-month observations for the period from 1991 to 2009 on a monthly basis. We examine the effects of several frequently used macroeconomic factors and firm-specific attributes on companies' one-month forward default probabilities starting one month ahead to as long as 36 months ahead. We find that a firm's leverage, liquidity, profitability and volatility are four important attributes affecting its forward default probabilities for almost all the horizons considered. Interestingly, our empirical results suggest that large companies seem to be able to delay defaults, but cannot fully avoid defaults simply because of the size advantage.

We also consider the influence of state variables in terms of both level and trend. Intuitively, a firm attribute's historical average (over some period) can distinguish it cross-sectionally from other firms in a particular dimension. The current value of a firm's attribute relative to its own historical average can also reveal its current momentum and suggests a direction of its future movements. Our empirical analysis indeed reveals that firm's distance-to-default (a commonly used variable in default analysis), along with several other variables, has effect in both dimensions. Although the trending aspect of a firm's attribute contains valuable information and enhances prediction power, its effect seems to be short-term except for distance-to-default.

Our forward intensity approach actually coincides with that of Duffie, *et al* (2007) when the application is limited to the one month ahead prediction. This is not at all surprising because forward intensity is basically spot intensity for one period ahead. Our implementation, however, uses more state variables and also considers the possibility of trending effect. The likelihood ratio test and individual t statistics suggest that both the additional variables and the trending treatment have highly significant impacts.

We also conduct a prediction accuracy analysis based on the commonly employed cumulative accuracy profile. The results show that the forward intensity approach is able to generate accurate predictions for short horizons such as one and three months. Their in-sample accuracy ratios exceed 90%, and the conclusion remains robust when the sample is split into two cross-sectionally and use one set to predict the other. The same conclusion holds true when an out-of-sample analysis is performed by rolling the sample forward over time. When the prediction horizon is extended to six months and one year, the accuracy ratios drop slightly to the 80% range. If the prediction horizon is further extended to two (or three) years, the performance drops to the 70% (or 60%) range. Note that the accuracy ratio for a totally uninformative model is supposed to be 0%. Again, the findings for longer prediction horizons are robust when the sample is split cross-sectionally and rolling over time so that the analysis is out-of-sample.

Our forward intensity model can naturally employ the convolution-based default aggregation algorithm of Duan (2010) to to study portfolio behavior. We are able to show that the predicted number of defaults is quite close to the actual numbers of defaults for the US corporate sector over the intended period when the prediction period is one and three months. For longer prediction periods, the performance is not as good but is still able to reflect the overall default pattern over the past twenty years.

Following Duan's (2010) treatment of distance-to-default, we are able to include financial firms in our analysis. Particularly, we single out Lehman Brothers as a case of interest. The analysis reveals that three months prior to Lehman Brothers's filing of Chapter 11 bankruptcy, the model has already suggested a substantially raised term structure of default probabilities. For example, the estimated probability of default in one year, predicted three months prior to Lehman's bankruptcy, had already reached about 10%. Interestingly, the peak forward default probability moved to around the third to fourth month, suggesting that default month will most likely coincide with its actual bankruptcy filing month.

2 A forward intensity approach to multiperiod default prediction

The Poisson process with stochastic intensities is often used to model the occurrence of defaults/bankruptcies. By the so-called doubly stochastic process approach, the stochastic intensity is a function of some state variables, either observable or unobservable, but the dynamics of these state variables are not affected by default. Since the relationship is unidirectional from state variables to the Poisson process, such a doubly stochastic model is easy to work with both in terms of computing quantities of interest and estimating the model parameters. This approach has been widely applied in the literature, for example, Duffie, *et al* (2007).

Mergers/aquisitions happen routinely. A public company traded in a stock exchange can be delisted for a variety of reasons. Naturally, default/bankruptcy is not the sole reason that a firm leaves the sample. Considering other forms of exit is critical in the analysis of default, because a default cannot happen after a firm has already exited due to other reasons. Exit due to reasons other than default/bankruptcy is usually modeled as another doubly stochastic process independent of the default process. It is worth noting that default and other form of exit are in principle mutually exclusive events. Thus, they are competing as opposed to independent risks. When they are modeled as two independent Poisson processes, the probability of joint occurrence happens to equal zero, blurring the distinction between competing and independent risks.

Default and other exit for the i-th firm in a group are governed by two independent doubly stochastic Poisson processes – M_{it} with stochastic intensity λ_{it} and L_{it} with stochastic intensity ϕ_{it}. λ_{it} and ϕ_{it} are instantaneous intensities and are only known at or after time t. Applying the standard argument, the probability of a standing firm surviving the period $[t, t+\tau]$ equals $E_t\left[\exp\left(-\int_t^{t+\tau}(\lambda_{is}+\phi_{is})ds\right)\right]$. The probability of default in the period $[t, t+\tau]$ is $E_t\left[\int_t^{t+\tau}\exp\left(-\int_t^s(\lambda_{iu}+\phi_{iu})du\right)\lambda_{is}ds\right]$. These quantities can only be computed with the exact knowledge of the stochastic processes: λ_{it} and ϕ_{it}. We contend that a more convenient way is to use the device of forward intensity rate.

First we define the spot combined exit intensity for default and other exits together for the period $[t, t+\tau]$, and through which we deduce the forward exit intensity. Denote by $F_{it}(\tau)$ the time-t conditional distribution function of the combined exit time evaluated at $t+\tau$. We assume that it is differentiable.

$$\psi_{it}(\tau) \equiv -\frac{\ln(1-F_{it}(\tau))}{\tau} = -\frac{\ln E_t\left[\exp\left(-\int_t^{t+\tau}(\lambda_{is}+\phi_{is})ds\right)\right]}{\tau}. \tag{1}$$

Obviously, $\exp[-\psi_{it}(\tau)\tau]$ becomes the survival probability over $[t, t+\tau]$.

Assume that $\psi_{it}(\tau)$ is differentiable. The forward exit intensity is defined as

$$g_{it}(\tau) \equiv \frac{F'_{it}(\tau)}{1-F_{it}(\tau)} = \psi_{it}(\tau) + \psi'_{it}(\tau)\tau. \tag{2}$$

Thus, $\psi_{it}(\tau)\tau = \int_0^\tau g_{it}(s)ds$. Finally, we define the forward default intensity censored by other forms of exit. Denote the default time of the i-th firm by τ_{Di} and the corresponding combined exit time by τ_{Ci}. Naturally, $\tau_{Ci} \leq \tau_{Di}$. Let $P_t(\cdot)$ denote the time-t conditional probability.

$$f_{it}(\tau) \equiv e^{\psi_{it}(\tau)\tau} \lim_{\Delta t \to 0} \frac{P_t(t+\tau < \tau_{Di} = \tau_{Ci} \leq t+\tau+\Delta t)}{\Delta t} \tag{3}$$

$$= e^{\psi_{it}(\tau)\tau} \lim_{\Delta t \to 0} \frac{E_t\left[\int_{t+\tau}^{t+\tau+\Delta t}\exp\left(-\int_t^s(\lambda_{iu}+\phi_{iu})du\right)\lambda_{is}ds\right]}{\Delta t}, \tag{4}$$

and the default probability over $[t, t+\tau]$ becomes $\int_0^\tau e^{-\psi_{it}(s)s} f_{it}(s)ds$.

Although we motivate the forward intensity model using a reduced-form approach involving doubly stochastic Poisson processes, the method conceptually encompasses the structural approach or a combination. For example, a combination can be (1) default is driven by a structural argument of asset value falling below promised debt payment, and (2) other forms of exit occur due to a Poisson event. Duffie and Lando (2001) argued that instantaneous default intensity does not exist unless the default time is totally inaccessible. Unfortunately, the structure model with the asset value driven by a diffusion process (or jump-diffusion process) is accessible (or neither accessible nor totally inaccessible). Therefore, such structural models cannot be given an intensity interpretation. Chen (2007), however, showed that forward intensity exists even for those structural models. Therefore, the forward intensity approach is not only more natural for multiperiod default prediction as will be demonstrated later, but also conceptually more widely applicable.

Instead of modeling λ_{it} and ϕ_{it} as some functions of state variables available at time t, we will later deal with $f_{it}(\tau)$ and $g_{it}(\tau)$ directly as functions of state variables available at time t and the horizon of interest, τ. Moreover, we need to ensure that $f_{it}(\tau) \leq g_{it}(\tau)$ to reflect the fact that default intensity must be no greater than combined exit intensity.

Let $X_{it} = (x_{it,1}, x_{it,2}, \cdots, x_{it,k})$ be the set of the state variables (stochastic and/or deterministic) that affect the forward intensities for the i-th firm. These variables may include two types of variables: macroeconomic factors and firm-specific attributes. Therefore, X_{it} and X_{jt} may share some common elements. $f_{it}(\tau)$ and $g_{it}(\tau)$ can be all kinds of functions of X_{it} as long as they are non-negative and $g_{it}(\tau) \geq f_{it}(\tau)$. For convenience, we let

$$\begin{aligned} f_{it}(\tau) &= \exp\left(\alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \alpha_2(\tau)x_{it,2} + \cdots \alpha_k(\tau)x_{it,k}\right) && (5) \\ g_{it}(\tau) &= f_{it}(\tau) + \exp\left(\beta_0(\tau) + \beta_1(\tau)x_{it,1} + \beta_2(\tau)x_{it,2} + \cdots \beta_k(\tau)x_{it,k}\right) && (6) \end{aligned}$$

Note that $f_{it}(\tau)$ and $g_{it}(\tau)$ do not need to share the same set of state variables. This can be achieved in the above specification by setting some coefficients to zero.

We need to discretize the model for empirical implementation, and for that we set one month as the basic time interval, i.e., $\Delta t = 1/12$. To simplify notation, from this point onwards, we view $t = 0, \Delta t, 2\Delta t, \cdots$ and $\tau = 0, \Delta t, 2\Delta t, \cdots$ as time sequences with an increment of Δt. The forward intensities in the discretized version, i.e., $f_{it}(\tau)$ and $g_{it}(\tau)$, should be understood as at time t for the period $[t + \tau, t + \tau + \Delta t]$.

When $\tau = 0$, our forward intensity set-up is the same as the spot intensity formulation of Duffie, *et al* (2007). The reason for using the forward intensity formulation is to deal with multiperiod default predictions without having to specify the dynamics for state variables, which in turn avoid estimating the state variable models and simulating these variables in computing predicted default probabilities.

We are interested in the following quantities in the discretized model for the firms that have not yet exited at time t. They can all be computed from $f_{it}(\tau)$ and $g_{it}(\tau)$.

1. Forward default probability at time t for the period $[t + \tau, t + \tau + \Delta t]$:

$$P_t(t + \tau < \tau_{Di} = \tau_{Ci} \leq t + \tau + \Delta t) = e^{-\psi_{it}(\tau)\tau}\left(1 - e^{-f_{it}(\tau)\Delta t}\right) \quad (7)$$

2. Forward combined exit probability at time t for the period $[t+\tau, t+\tau+\Delta t]$:

$$P_t(t+\tau < \tau_{Ci} \leq t+\tau+\Delta t) = e^{-\psi_{it}(\tau)\tau}\left(1 - e^{-g_{it}(\tau)\Delta t}\right) \tag{8}$$

3. Cumulative default probability at time t for the period $[t, t+\tau]$:

$$P_t(t < \tau_{Di} = \tau_{Ci} \leq t+\tau) = \sum_{j=1}^{\tau/\Delta t} e^{-\psi_{it}((j-1)\Delta t)(j-1)\Delta t}\left(1 - e^{-f_{it}((j-1)\Delta t)\Delta t}\right) \tag{9}$$

4. Spot combined exit intensity at time t for the period $[t, t+\tau]$:

$$\psi_{it}(\tau) = \frac{1}{\tau}\left[\psi_{it}(\tau-\Delta t)(\tau-\Delta t) + g_{it}(\tau-\Delta t)\Delta t\right] \tag{10}$$

Note that $\psi_{it}(0)$ need not be specified because it is irrelevant.

3 Estimating the forward intensity model

3.1 Overlapped pseudo-likelihood function

First, we extend our notations used in the preceding section. Suppose that our sample period is $[0, T]$ and is divided into $T/\Delta t$ periods. Let N be the total number of companies. For firm i, we let t_{0i} be the first month that it appeared in the sample. τ_{Di} is the default time and τ_{Ci} is the combined exit time. If a firm exits due to default, then $\tau_{Di} = \tau_{Ci}$, and otherwise, $\tau_{Ci} < \tau_{Di}$. The covariates X_{it} consist of two parts $X_{it} = (W_t, U_{it})$. W_t are the factors common to all firms, and U_{it} are the firm-specific variables which cease to be observable after a company exits the sample. Suppose τ is the intended prediction horizon measured in terms of the number of basic periods with each equal to Δt.

We denote the model's parameter set by $\alpha = \{\alpha(0), \alpha(\Delta t), \cdots, \alpha(\tau-\Delta t)\}$ and $\beta = \{\beta(0), \beta(\Delta t), \cdots, \beta(\tau-\Delta t)\}$. The pseudo-likelihood function for prediction horizon τ is

$$\begin{aligned}&\mathscr{L}_\tau(\alpha, \beta; \tau_C, \tau_D, X)\\ =& \prod_{j=0}^{T/\Delta t-1} P(\tau_{Ci} \wedge ((j+1)\Delta t+\tau), \tau_{Di} \wedge ((j+1)\Delta t+\tau), i = 1, 2, \cdots, N | X_{j\Delta t}; \alpha, \beta)\end{aligned} \tag{11}$$

Note that when $\tau > 1$, the above pseudo-likelihood is constructed with observations from overlapped periods. As an example, when $\tau = 2$, default over the next two periods corresponding to one data point is correlated with default used in the next data point (one period after) due to two observations sharing a common one period.

We assume

$$
\begin{aligned}
&P(\tau_{Ci} \wedge ((j+1)\Delta t + \tau), \tau_{Di} \wedge ((j+1)\Delta t + \tau), i = 1, 2, \cdots, N | X_{j\Delta t}; \alpha, \beta) \\
=&\prod_{i=1}^{N} P(\tau_{Ci} \wedge ((j+1)\Delta t + \tau), \tau_{Di} \wedge ((j+1)\Delta t + \tau) | X_{i,j\Delta t}; \alpha, \beta)
\end{aligned} \tag{12}
$$

This assumption lets firms' survival and default probabilities depend only upon the common factors and firm-specific attributes. Hence, different firms are conditionally independent among themselves. If there is any dependency, it must arise from their sharing of the common factors and/or any correlation among the firm-specific attributes. This assumption is in essence similar to the doubly stochastic assumption (also known as conditional independence assumption) used in the traditional intensity model. One firm's exit does not feed back to the state variables. Neither does it influence the exit probabilities of other firms. Under this assumption, the pseudo-likelihood function can be expressed as

$$
\mathscr{L}_\tau(\alpha, \beta; \tau_C, \tau_D, X) = \prod_{i=1}^{N} \prod_{j=0}^{T/\Delta t - 1} P_{\tau,i,j}(\alpha, \beta) \tag{13}
$$

where

$$
\begin{aligned}
P_{\tau,i,j}(\alpha, \beta) \equiv& P(\tau_{Ci} \wedge ((j+1)\Delta t + \tau), \tau_{Di} \wedge ((j+1)\Delta t + \tau) | X_{i,j\Delta t}; \alpha, \beta) \\
=& 1_{\{t_{0i} \leq j\Delta t, \tau_{Ci} > j\Delta t + \tau\}} \exp\left\{ -\sum_{k=0}^{\tau/\Delta t - 1} g_{i,j\Delta t}(k\Delta t)\Delta t \right\} \\
& + 1_{\{t_{0i} \leq j\Delta t, \tau_{Di} = \tau_{Ci} \leq j\Delta t + \tau\}} \exp\left\{ -\sum_{k=0}^{\tau_{Di}/\Delta t - j - 2} g_{i,j\Delta t}(k\Delta t)\Delta t \right\} \\
& \quad \times (1 - \exp\{-f_{i,j\Delta t}(\tau_{Di} - (j+1)\Delta t)\Delta t\}) \\
& + 1_{\{t_{0i} \leq j\Delta t, \tau_{Di} > \tau_{Ci}, \tau_{Ci} \leq j\Delta t + \tau\}} \exp\left\{ -\sum_{k=0}^{\tau_{Ci}/\Delta t - j - 2} g_{i,j\Delta t}(k\Delta t)\Delta t \right\} \\
& \quad \times (\exp\{-f_{i,j\Delta t}(\tau_{Ci} - (j+1)\Delta t)\Delta t\} - \exp\{-g_{i,j\Delta t}(\tau_{Ci} - (j+1)\Delta t)\Delta t\}) \\
& + 1_{\{t_{0i} > j\Delta t\}} + 1_{\{\tau_{Ci} \leq j\Delta t\}}
\end{aligned}
$$

The first term on the right-hand side of the above expression is the probability of surviving both forms of exit. The second term is the probability that firm defaults. The third term is the probability that firm exits due to other reasons. If the firm does not appear in the sample in month t, then we set $P_{\tau,i,j}$ to 1, which is transformed to 0 in the pseudo-log-likelihood function.

The pseudo-likelihood function $\mathscr{L}_\tau$ can be numerically maximized to obtain estimates $\hat{\alpha}$ and $\hat{\beta}$. Due to the overlapping nature of the pseudo-likelihood function, the associated inference is not immediately clear, however. This "overlapped" pseudo-likelihood function, for example, violates the standard assumption. We thus derive the large sample properties in Appendix A.

3.2 Decomposable pseudo-likelihood function

Because the pseudo-likelihood function (13) is the product of separate terms involving α and β, we can maximize its two components separately to obtain $\hat{\alpha}$ and $\hat{\beta}$. The two components are

$$\mathscr{L}_{\tau}^{\alpha} = \mathscr{L}_{\tau}(\alpha; \tau_C, \tau_D, X) = \prod_{i=1}^{N} \prod_{j=0}^{T/\Delta t-1} \mathscr{L}_{\tau,i,j}^{\alpha} \tag{14}$$

$$\mathscr{L}_{\tau}^{\beta} = \mathscr{L}_{\tau}(\beta; \tau_C, \tau_D, X) = \prod_{i=1}^{N} \prod_{j=0}^{T/\Delta t-1} \mathscr{L}_{\tau,i,j}^{\beta} \tag{15}$$

where

$$\begin{aligned}
\mathscr{L}_{\tau,i,j}^{\alpha} =& 1_{\{t_{0i}\le j\Delta t, \tau_{Ci}>j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau/\Delta t-1} f_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&+ 1_{\{t_{0i}\le j\Delta t, \tau_{Di}=\tau_{Ci}\le j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau_{Di}/\Delta t-j-2} f_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&\quad \times (1-\exp\{-f_{i,j\Delta t}(\tau_{Di}-(j+1)\Delta t)\Delta t\}) \\
&+ 1_{\{t_{0i}\le j\Delta t, \tau_{Di}>\tau_{Ci}, \tau_{Ci}\le j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau_{Di}/\Delta t-j-2} f_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&\quad \times \exp\{-f_{i,j\Delta t}(\tau_{Di}-(j+1)\Delta t)\Delta t\} \\
&+ 1_{\{t_{0i}>j\Delta t\}} + 1_{\{\tau_{Ci}\le j\Delta t\}}
\end{aligned}$$

$$\begin{aligned}
\mathscr{L}_{\tau,i,j}^{\beta} =& 1_{\{t_{0i}\le j\Delta t, \tau_{Ci}>j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau/\Delta t-1} h_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&+ 1_{\{t_{0i}\le j\Delta t, \tau_{Di}=\tau_{Ci}\le j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau_{Di}/\Delta t-j-2} h_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&+ 1_{\{t_{0i}\le j\Delta t, \tau_{Di}>\tau_{Ci}, \tau_{Ci}\le j\Delta t+\tau\}} \exp\left\{-\sum_{k=0}^{\tau_{Di}/\Delta t-j-2} h_{i,j\Delta t}(k\Delta t)\Delta t\right\} \\
&\quad \times (1-\exp\{-h_{i,j\Delta t}(\tau_{Di}-(j+1)\Delta t)\Delta t\}) \\
&+ 1_{\{t_{0i}>j\Delta t\}} + 1_{\{\tau_{Ci}\le j\Delta t\}}
\end{aligned}$$

and

$$h_{it}(\tau) = g_{it}(\tau) - f_{it}(\tau) = \exp\left(\beta_0(\tau) + \beta_1(\tau)x_{it,1} + \beta_2(\tau)x_{it,2} + \cdots \beta_k(\tau)x_{it,k}\right)$$

Note that equations (14)-(15) can be further decomposed to separate terms involving $\alpha(s)$ and $\beta(s)$ for different s. Therefore, we can obtain the maximum pseudo-likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$ without having to perform estimation sequentially from shorter to longer prediction horizons. The horizon-specific pseudo-likelihood functions are

$$\mathscr{L}_{\alpha(s)} = \prod_{i=1}^{N} \prod_{j=0}^{(T-s)/\Delta t-1} \mathscr{L}_{\alpha(s),i,j}, \quad s = 0, \Delta t, \cdots, \tau - \Delta t \tag{16}$$

$$\mathscr{L}_{\beta(s)} = \prod_{i=1}^{N} \prod_{j=0}^{(T-s)/\Delta t-1} \mathscr{L}_{\beta(s),i,j}, \quad s = 0, \Delta t, \cdots, \tau - \Delta t \tag{17}$$

where

$$\begin{aligned}
\mathscr{L}_{\alpha(s),i,j} =& 1_{\{t_{0i} \le j\Delta t, \tau_{Ci} > (j+1)\Delta t + s\}} \exp\left\{-f_{i,j\Delta t}(s)\Delta t\right\} \\
&+ 1_{\{t_{0i} \le j\Delta t, \tau_{Di} = \tau_{Ci} = (j+1)\Delta t + s\}} (1 - \exp\left\{-f_{i,j\Delta t}(s)\Delta t\right\}) \\
&+ 1_{\{t_{0i} \le j\Delta t, \tau_{Di} \ne \tau_{Ci}, \tau_{Ci} = (j+1)\Delta t + s\}} \exp\left\{-f_{i,j\Delta t}(s)\Delta t\right\} \\
&+ 1_{\{t_{0i} > j\Delta t\}} + 1_{\{\tau_{Ci} < (j+1)\Delta t + s\}}
\end{aligned}$$

$$\begin{aligned}
\mathscr{L}_{\beta(s),i,j} =& 1_{\{t_{0i} \le j\Delta t, \tau_{Ci} > (j+1)\Delta t + s\}} \exp\left\{-h_{i,j\Delta t}(s)\Delta t\right\} \\
&+ 1_{\{t_{0i} \le j\Delta t, \tau_{Di} = \tau_{Ci} = (j+1)\Delta t + s\}} \\
&+ 1_{\{t_{0i} \le j\Delta t, \tau_{Di} \ne \tau_{Ci}, \tau_{Ci} = (j+1)\Delta t + s\}} (1 - \exp\left\{-h_{i,j\Delta t}(s)\Delta t\right\}) \\
&+ 1_{\{t_{0i} > j\Delta t\}} + 1_{\{\tau_{Ci} < (j+1)\Delta t + s\}}
\end{aligned}$$

4 Data and the choice of covariates

4.1 Data

Our data set is a large sample of U.S. public firms over the period from 1991 to 2009. The stock market data are from the CRSP monthly and daily files. We only include companies traded on NYSE, AMEX and Nasdaq (exchange code 1 to 3) with share code 10 and 11 (common stocks). The accounting data are taken from the Compustat quarterly file. Since the accounting statements are usually released several months after the reporting period, we lag all the accounting items by three months. If the accounting variable is missing, we substitute it with the closest observation prior to the relevant date. Our default and bankruptcy data are obtained from three different sources. We use the CRSP delisting code "574" for bankruptcy. We also identify a delisting as bankruptcy if the delisting reason is "02" in Compustat.[2] A default or bankruptcy is also recorded if the CACS function of Bloomberg indicates so. Similar to Shumway (2001), firms that defaulted or filed for any type of bankruptcy within 1 year of delisting are considered to be in default status by the time of delisting. There are altogether 12,196 companies (including financial firms) giving rise to 1,030,305 firm-month observations in our sample. Table 1 summarizes the number of active companies, defaults/bankruptcies and other exits each year. The summary statistics show, as expected, that the overall default rate is low ranging between 0.18% and 3.2% of the firms in each sample year. Other forms of exit are significantly higher, ranging from 5.16% to 13.63%.

4.2 Covariates

We use the following set of common factors and firm-specific attributes to characterize the forward intensity functions:

1. SP500: trailing 1-year return on the S&P500 index.
2. Treasury rate: 3-month US Treasury bill rate.
3. DTD: firm's distance-to-default, which is a volatility adjusted leverage measure based on Merton (1974). The DTD is estimated once a month using the preceding one year of daily equity values. To include financial firms in our analysis, we follow Duan's (2010) adjustment method to include firm's liabilities beyond short- and long-term debts. The model parameters are estimated by the transformed-data maximum likelihood method in Duan (1994, 2000). The parameter estimates are then used to compute DTDs and the last valid DTD is the one used as the covariate. The methodological details are provided in Appendix B.
4. CASH/TA: ratio of the sum of cash and short-term investments to the total assets.
5. NI/TA: ratio of net income to the total assets.
6. SIZE: log of the ratio of firm's market equity value to the average market equity value of the S&P500 firm.
7. M/B: market-to-book asset ratio.
8. SIGMA: 1-year idiosyncratic volatility, calculated by regressing individual monthly stock return on the value-weighted CRSP monthly return over the preceding 12 months. SIGMA is the standard deviation of the residuals from the regression.

[2]Duffie, *et al* (2007) regarded both "02" and "03" as bankruptcy. However, we have confirmed with Standard & Poor's that code "03" stands for liquidation for any reasons.

Following Shumway (2001), we treat SIGMA as missing if there are less than 12 monthly returns.

Our DTD differs from that of Duffie, *et al* (2007) in two aspects. First, they estimated the parameters of the Merton (1974) model for each firm once and for all using the entire sample (monthly data) instead of using a moving window approach, which in a sense has inappropriately peeked into the future. Second, we have adopted a different debt specification by incorporating other liabilities, which in turn allows us to include financial firms.

The first three variables were used in Duffie, *et al* (2007). They also used firm's own one-year trailing return as a covariate, but our analysis shows that it is insignificant after incorporating other variables. We also considered several other covariates frequently used in the previous literature, but didn't include them due to either lack of significance or creating a serious missing value problem.

Interestingly, we discover that both trend and level of some firm-specific attributes play an important role. It is not at all surprising to find that momentum plays a role in predicting defaults. For example, other things being equal, two firms with same DTD are likely to face different default likelihoods if one firm's DTD has been deteriorating in the past few months whereas the other firm has experienced improvement in its DTD. We compute the average of a variable over the preceding 12 months, and denote it by the subscript "AVG" to reflect the level of such variable. We also calculate the difference between its current value and the 12-month moving average, and denote it by the subscript "DIF". The "DIF" measure proxies for the trending aspect of a variable. We found both trend and level measures for DTD, CASH/TA, NI/TA and SIZE to be significant. To dampen the effect of outliers, we winsorize each of the above firm-specific attributes. We cap all the observations at the 999-th permille value. Similarly, all values are subject to the floor at the first permille value. The summary statistics and correlation matrix for the firm-specific attributes are reported in Tables 2-3.

5 Empirical results

5.1 Parameter estimates

We present in Tables 4-5 the maximum pseudo-likelihood estimates for $\alpha(\tau)$ and $\beta(\tau)$ with different τ ranging from 1 month to 36 months. To show the impact of various factors/attributes on firms' default probabilities, we plot in Figure 1 the estimated coefficients corresponding to different prediction horizons. Also plotted is the 90% confidence interval for each variable used in the forward default intensity function.

In terms of the trailing 1-year S&P500 index return, the forward default intensity coefficients for most of the prediction horizons are positive but their magnitudes first decrease with the prediction horizon and then rise later. What it suggests is that when the equity market performs well, firms are more likely to default, a result seems counter-intuitive. This could be caused by the correlation between the S&P500 index return and other firm-specific attributes. For example, as suggested by Duffie, *et al* (2009), that "after boom years in the stock market, a firm's distance to default overstates its financial health". Hence, the S&P500 index return simply serves as a correction.

The forward default intensities are estimated to decrease with the 3-month Treasury bill rate in the short run but to increase in the long run. The signs of the coefficients at short horizons are consistent with the fact that the short-term interest rate is typically lowered by the US Federal Reserve to stimulate the economic growth during recessions and increased to fight inflation during expansions. The opposite signs of the coefficients may simply reflect the business cycle effect.

The estimated forward default intensities decrease with firm's moving average of distances-to-default for all prediction horizons. Although our distance-to-default measure is somewhat different, this finding is consistent with those reported in the literature such as Hillegeist, *et al* (2004), Duffie, et al (2007), and Bharath and Shumway (2008), showing that distance-to-default is a highly useful attribute for differentiating a firm's credit risk from other firms. Moreover, we find that forward default intensity also decreases in a significant manner with the distance-to-default trend for all prediction horizons analyzed. To our knowledge, this is the first study that the distance-to-default trend measure is used to characterize default likelihood.

The CASH/TA variable captures the liquidity position of a company. Other things being equal, a firm with more liquid assets available to meet interest and principal payments is more likely to avoid default. The forward default intensities are estimated to decrease with both the trend and level of CASH/TA, but the trend measure loses its significance when the prediction horizon becomes longer. This suggests that the liquidity trend measure is more indicative of short-run default likelihood.

We measure a firm's profitability by the NI/TA ratio. A firm's ultimate existence is based on the profitability of its business. This measure is expected to play a role in the default/bankruptcy analysis. Bharath and Shumway (2008) found that this measure provides significant predicting power in addition to distance-to- default. We also find that estimated forward default intensities are strongly decreasing in the level of profitability for all prediction horizons considered. The trend measure for profitability turns out to be significant for shorter prediction horizons.

Firm size has long been regarded as an important predictor for default/bankruptcy ever since the early days of reduced-form modeling. Large firms are usually thought to have more diversified business lines and financial flexibility than smaller firms, which may help them better weather financial distress. Large firms are also more likely to be bailed out by governments because they may be "too big to fail". Our results show that forward default intensities do decrease with size in the short run but however increase in the long run. This means that other things being equal, large companies can postpone defaults rather than fully avoid them. The trend measure of size can be viewed as a proxy for a firm's growth pattern. The forward intensities are found to be decreasing in this trend measure only for short prediction horizons, indicating that fast growth may lower default likelihood in the short run.

Market-to-book asset ratio is a mixed measure for the market mis-valuation and future growth opportunities. If the market mis-valuation effect dominates, then the forward default intensities should be increasing in market-to-book asset ratio. Otherwise, the signs of the coefficients should be negative. Our results show that after controlling for other covariates, estimated forward intensities are increasing in market-to-book asset ratio for up to the 2-year prediction horizon, which is consistent with Campbell, *et al* (2008). Such finding supports the mis-valuation explanation. The effect of market-to-book asset ratio on default probability can be further studied by decomposing this measure into mis-valuation and growth option components using the methodology developed in Rhodes-Kropf, *et al* (2005). However, our interest here is not on finding why market-to-book asset ratio is significant and will leave this matter to future research.

The idiosyncratic standard deviation measure is first employed by Shumway (2001), who argued that "If a firm has more variable cash flows (and hence more variable stock return), then the firm ought to have a higher probability of bankruptcy." Our finding is consistent with Shumway's (2001) argument. The forward default intensities are strongly increasing in this idiosyncratic risk measure for almost all the horizons being considered.

Our estimates of the forward intensity function for exits due to reasons other than default are presented in Table 5. All common factors and firm-specific attributes used in the forward default intensity functions continue to be relevant. The results show that all variables are significant even though they may not be so for all prediction horizons. We skip the detailed discussions here to conserve space.

5.2 Aggregate number of defaults

At each month-end, we compute the predicted number of defaults among the active firms in the sample for a prediction horizon. We then compare it with the observed number of defaults in the intended prediction period. Repeating this for the entire sample and for different prediction horizons. Figure 2 plots the comparisons for the following horizons: 1 month, 3 months, 6 months, 12 months, 24 months and 36 months. The line depicts the predicted values whereas the bars are the observed numbers of defaults. For shorter horizons, our model fits the reality quite well. However, as the horizon increases, the line deviates from the bars, implying a deteriorating performance in the long run. Generally speaking, our model overstates the overall credit risk in the beginning of the sample period and understates the overall credit risk towards the end of the sample period.

There are many possibilities for the model's deteriorating performance for longer prediction horizons. One natural speculation is that our model has missed out some variables that are capable of reflecting long-term credit risk. A potential quick fix is to introduce the frailty effect as suggested in the previous literature such as Koopman, *et al* (2008) and Duffie, *et al* (2009) or to employ the regime-switching approach as in Chuang and Kuan (2010). Koopman, *et al* (2009a&b) studied the relation between macroeconomic fundamentals and cycles in defaults and rating activities. They found that portfolio credit risk models which are solely based on observable common risk factors omit one of the strongest determinants of credit risk. By accounting for the latent frailty factor or hidden regimes, one may be able to improve our forward intensity model. Another possibility is to experiment with different functional forms in relating the forward intensity to the covariates.

5.3 Prediction accuracy

In this section, we employ the cumulative accuracy profile and its associated accuracy ratio to evaluate our model's prediction accuracy. The cumulative accuracy profile, also known as power curve, examines a model's performance based on risk rankings. A detailed description can be found in Crosbie and Bohn (2002) and Vassalou and Xing (2004). To check our model's in-sample performance, we estimate the cumulative default probabilities for each firm-month observation employing the parameter estimates reported in Tables 4-5 where all the firm-month observations are included in the estimation. Figure 3A plots the cumulative accuracy profiles for the prediction horizons: 1 month, 3 months, 6 months, 12 months, 24 months and 36 months. Table 6 (Panel A) reports the accuracy ratios. The predictions for short horizons are very accurate with the accuracy ratios for 1 month and 3 months prediction exceeding 90%. The accuracy ratios for 6 months and 12 months are also very good with their values staying above 80%. As the horizon increases to 24 months and 36 months, the accuracy ratios reduce to 72.74% and 65.80%, respectively.

We also implement out-of-sample analysis to ascertain the model's performance. First, we randomly and equally divide all companies into two groups: the estimation group and the evaluation group. Then we estimate the parameters using the estimation group and apply the estimated coefficients to the evaluation group to generate the cumulative accuracy profiles and to compute the associated accuracy ratios for different prediction horizons. Figure 3B plots the cumulative accuracy profiles for this out-of-sample analysis, and Table 6 (Panel B) reports the accuracy ratios. The results show that the model is very stable in the sense that the accuracy ratios in the cross-sectional out-of-sample analysis are very close to those obtained from the in-sample analysis.

An out-of-sample analysis in the time dimension is also conducted. We use a moving-window approach. At each month-end starting from January 2001, we re-estimate the model using all the data available up to that time and compute predicted default probabilities for different prediction horizons. This analysis is more indicative of the performance of the model in line with the situation in real applications. Figure 3C plots this out-of-sample performance result based on the cumulative accuracy profile. Their out-of-sample accuracy ratios are reported in Table 6 (Panel C). Again, the accuracy ratios are very close to the in-sample results.

5.4 A case study of Lehman Brothers

We use Lehman Brothers as an illustrative example to see whether the term structure of predicted default probabilities is informative. Our analysis is conducted in the out-of-sample sense employing only data that were available at the time of computing the term structure. Lehman Brothers filed for the Chapter 11 bankruptcy on September 15th, 2008. We plot in Figure 4 the estimated term structure of forward and cumulative default probabilities at several time points prior to its bankruptcy filing. On the same graph, we also plot the forward and cumulative default probabilities for Merrill Lynch, Bank of America as well as the average values of the financial sector. Our results reveal that the term structure is very informative, particularly in light of other financial firms over the same time period.

The first set of two plots shows the estimated term structure of forward default probabilities and that of cumulative default probabilities in September 2005, which was 36 months before Lehman Brothers' bankruptcy filing. The term structure for the forward default probabilities was upward sloping, making the cumulative rising faster when the prediction horizon increases. The predicted cumulative default probabilities were quite low in value, however, with the 1-year cumulative default probability being 0.2% and 3-years cumulative default probability being around 1.5%. This result suggests that the market did not foresee any noticeable problem with Lehman Brothers three years prior to its bankruptcy filing. Lehman Brothers had its distance-to-default at 2.7 and was trending up by comparing with its preceding 12-month average of 1.5. The company also had enough liquid assets with CASH/TA ratio higher than 25%. Its profitability was, however, less than 1%, which possibly led to the upward sloping forward default probability term structure. The same pattern applied to other financial firms as well.

The second set of plots is the term structures in September 2006, which was 24 months before its bankruptcy. The term structure of forward default probabilities was hump-shaped and peaked at around 24 months. The 1-year cumulative default probability rose to 0.4% while the 3-years cumulative default probability rose to 2.7%. The stock market was bullish then with the S&P500 index increasing by over 8% in the previous year. Lehman Brothers remained highly liquid then. Its distance-to-default reduced to 1.2, and net income remained less than 1% of its book asset value.

The third set of plots presents Lehman Brothers' term structures of forward and cumulative default probabilities in September 2007, which was 12 months before its bankruptcy filing. The forward curve remained hump-shaped with the peak moving to 16 months. The 1-year cumulative default probability further rose to 1.6% and the 3-year cumulative default probability rose significantly to 6.1%. The S&P500 index increased by over 14% in the previous year. But Lehman Brothers' distance-to-default dropped to 0.1 and its stock had lost by more than 10% over the previous twelve months.

The last set of plots is the term structures for Lehman Brothers in June 2008, just 3 months before its bankruptcy filing. The company's short-term credit risk reached its historical high. The peak of the forward default probability curve moved to 3 months. The 1-year cumulative default probability increased sharply to 9.9% which is about 50 times of the value 3 years earlier. The 3-year cumulative default probability climbed to 16.4%. The stock market turned bearish with the S&P500 index dropping by almost 15% in the previous year. Lehman Brothers' distance-to-default further decreased to -1.7. And the company's stock price also reached the lowest level in 5 years then. Interestingly, other US financial firms did not follow Lehman Brothers' pattern. This case analysis seems to suggest that our forward intensity model is highly informative about the dynamic evolution of Lehman Brothers' default prospect.

6 Conclusion

We have developed a reduced-form model for predicting corporate defaults/bankruptcies over different prediction horizons. Our approach relies on constructing forward intensities. The forward intensity model is implemented on a large sample of the US public firms listed on three major stock exchanges. We use two common factors and six firm-specific attributes to characterize the two forward intensity functions: default and other forms of exit. We found that some firm-specific attributes influence the forward intensity both in terms of level and trend. The forward intensity model is shown to perform very well for shorter prediction horizons. For longer prediction horizons (two to three years), the model's performance deteriorates somewhat, but still seems to track the general default pattern over time. We believe that improvement in performance should be possible with further research.

We have demonstrated that the forward intensity approach can be operationally implemented for default prediction for different horizons. Needless to say, it can be used for credit risk analysis of individual firms such as credit ratings. The forward intensity model also lends itself naturally to portfolio aggregation. By applying the aggregation algorithm of Duan (2010) for the standard intensity model, one can generate the default distribution (in terms of the number of defaults or the size of exposure) for any credit portfolio. In short, it is also a practical bottom-up approach to credit portfolio analysis.

7 Appendix

A Large sample properties of the estimator

We characterize the large sample properties of the estimator based on maximizing the pseudo-likelihood function in (13). The parameter set is denoted by θ and its true value is θ_0. $\log \mathscr{L}_N(y, \theta)$ is the pseudo-log-likelihood function when there are N companies and y denotes the companies' status indicators. To prove the consistency of the maximum pseudo-likelihood estimator, We make the following assumptions:

Assumption 1. *The parameter space Θ is an open bounded subset of the Euclidean K-space.*

Assumption 2. *The covariate vectors $\{x_{it}\}$ are uniformly bounded and the nonsingularity condition holds such that*

$$\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sum_{j=0}^{T/\Delta t-1} \left(\frac{\exp(-f_{i,j\Delta t}(\Delta t)\Delta t)}{1-\exp(-f_{i,j\Delta t}(\Delta t)\Delta t)} + \frac{\exp(-h_{i,j\Delta t}(\Delta t)\Delta t)}{1-\exp(-h_{i,j\Delta t}(\Delta t)\Delta t)} \right) x_{it}x_{it}'$$

is a finite nonsingular matrix.

The form of nonsingularity assumption is due to our forward intensity specification. It should be noted that although only the total number of firms N is required to be large for the consistency result, the total number of periods $T/\Delta t$ need to be larger than the dimension of the common attributes in order to allow the nonsingularity condition to hold. We first state the lemmas used in the proof below. These lemmas are corresponding to Theorem 4.1.2 and 4.2.2 in Amemiya (1986).

Lemma 1. *Under the conditions:*
(A) The parameter space Θ is an open subset of the Euclidean K-space.
(B) $\log \mathscr{L}_N(y, \theta)$ is a measurable function of y for all $\theta \in \Theta$, and $\partial \log \mathscr{L}_N / \partial\theta$ exists and is continuous in an open neighborhood $N_1(\theta_0)$ of θ_0.
(C) There exists an open neighborhood $N_2(\theta_0)$ of θ_0 such that $N^{-1} \log \mathscr{L}_N(\theta)$ converges to a nonstochastic function $l(\theta)$ in probability uniformly in θ in $N_2(\theta_0)$, and $l(\theta)$ attains a strict local maximum at θ_0.
Let Θ_N be the set of roots of the equation

$$\frac{\partial \log \mathscr{L}_N}{\partial \theta} = 0$$

corresponding to the local maxima. Then for any $\epsilon > 0$,

$$\lim_{N\to\infty} P[\inf_{\theta\in\Theta_N} (\theta-\theta_0)'(\theta-\theta_0) > \epsilon] = 0$$

Lemma 2. *Let $g_i(y,\theta)$ be a measurable function of y in Euclidean space for each i and for each $\theta \in \Theta$, a compact subset of Euclidean K-space, and a continuous function of θ for each y uniformly in i. Assume $Eg_i(y,\theta) = 0$. Let $\{y_i\}$ be a sequence of independent and not necessarily identically distributed random vectors such that $E\sup_{\theta\in\Theta}|g_i(y_i,\theta)|^{1+\delta} \le M < \infty$ for some $\delta > 0$. Then $N^{-1}\sum_{i=1}^{N} g_i(y_i,\theta)$ converges to 0 in probability uniformly in $\theta \in \Theta$.*

To prove the consistency, we verify the conditions of Lemma 1. Conditions (A) and (B) are obviously satisfied. To verify (C), we make use of Lemma 2 and define $g_i(y,\theta) = \sum_{j=0}^{T/\Delta t-1}\log P_{\tau,i,j}(\theta) - E_{\theta_0}\sum_{j=0}^{T/\Delta t-1}\log P_{\tau,i,j}(\theta)$. $g_i(y,\theta)$ in a compact neighborhood of θ_0 satisfies all the conditions in Lemma 2 because of the assumptions. Therefore,

$$N^{-1}\sum_{i=1}^{N}\sum_{j=0}^{T/\Delta t-1}\log P_{\tau,i,j}(\theta) \to l(\theta) = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E_{\theta_0}\sum_{j=0}^{T/\Delta t-1}\log P_{\tau,i,j}(\theta)$$

uniformly in θ as $N \to \infty$. By making use of Assumption 2 as well as the exact function form of the pseudo-log-likelihood function, we can also prove that $l(\theta)$ attains a strict local maximum at $\theta = \theta_0$. Thus we complete the proof of consistency.

To show the asymptotic normality of the estimator, denote the maximum pseudo-likelihood estimates as $\hat{\theta}$ and use the Taylor expansion to obtain

$$\begin{aligned}\frac{\partial \log \mathscr{L}_N(\hat{\theta})}{\partial\theta} &= \frac{\partial \log \mathscr{L}_N(\theta_0)}{\partial\theta} + \frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'}(\hat{\theta}-\theta_0), \text{ where } \tilde{\theta} \text{ lies between } \theta_0 \text{ and } \hat{\theta}\\ \Rightarrow \hat{\theta}-\theta_0 &= -\left(\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'}\right)^{-1}\left(\frac{1}{N}\frac{\partial \log \mathscr{L}_N(\theta_0)}{\partial\theta}\right)\end{aligned}$$

Consider

$$\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=0}^{T/\Delta t-1}\frac{\partial^2 \log P_{\tau,i,j}(\tilde{\theta})}{\partial\theta\partial\theta'} \to_p H(\tilde{\theta}) \ \ as\ N \to \infty,$$

where $H(\theta) = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E_{\theta_0}\sum_{j=0}^{T/\Delta t-1}\frac{\partial^2 \log P_{\tau,i,j}(\theta)}{\partial\theta\partial\theta'}$. Since $\hat{\theta}$ converges to θ_0 and $\tilde{\theta}$ lies between $\hat{\theta}$ and θ_0, $H(\tilde{\theta})$ converges to $H(\theta_0)$. So

$$\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=0}^{T/\Delta t-1}\frac{\partial^2 \log P_{\tau,i,j}(\tilde{\theta})}{\partial\theta\partial\theta'} \to_p H(\theta_0) \ \ as\ N \to \infty.$$

Therefore,

$$\begin{aligned}&\sqrt{N}(\hat{\theta}-\theta_0)\\ =& -\left(\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'}\right)^{-1}\left(\frac{1}{\sqrt{N}}\frac{\partial \log \mathscr{L}_N(\theta_0)}{\partial\theta}\right)\\ =& -\left(\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial\theta\partial\theta'}\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\sum_{j=0}^{T/\Delta t-1}\frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial\theta}\right)\end{aligned}$$

where $\left\{\sum_{j=0}^{T/\Delta t-1} \frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial \theta}, i = 1, 2, \cdots, N\right\}$ are independent. Then, according to Lindeberg's central limit theorem, $\sqrt{N}(\hat{\theta} - \theta_0)$ is asymptotically normally distributed with the mean vector equal to 0 and the variance-covariance matrix being

$$H(\theta_0)^{-1} \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left[\left(\sum_{j=0}^{T/\Delta t-1} \frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial \theta}\right)\left(\sum_{j=0}^{T/\Delta t-1} \frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial \theta}\right)'\right] H(\theta_0)^{-1}.$$

The asymptotic variance can thus be approximated by

$$\begin{aligned}
&\mathrm{Var}(\hat{\theta} - \theta_0) \\
= &\left(\frac{\partial^2 \log \mathscr{L}_\tau(\hat{\theta})}{\partial\theta\partial\theta'}\right)^{-1} \sum_{i=1}^{N} \left[\left(\sum_{j=0}^{T/\Delta t-1} \frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial \theta}\right)\left(\sum_{j=0}^{T/\Delta t-1} \frac{\partial \log P_{\tau,i,j}(\theta_0)}{\partial \theta}\right)'\right] \\
&\times \left(\frac{\partial^2 \log \mathscr{L}_\tau(\hat{\theta})}{\partial\theta\partial\theta'}\right)^{-1}.
\end{aligned}$$

B Estimating distance-to-default (DTD)

This appendix briefly reviews the Merton (1974) model and explains the numerical scheme employed to calculate distance-to-default. Merton's model assumes that firms are financed by equity and one single pure discount bond with maturity date T and principal L. The asset value V_t follows geometric Brownian motion:

$$dV_t = \mu V_t dt + \sigma V_t dB_t.$$

Due to limited liability, the equity value at maturity is $E_T = \max(V_T - L, 0)$. Therefore, the equity value at time $t \leq T$ by the Black-Scholes option pricing formula becomes

$$E_t = V_t N(d_t) - e^{-r(T-t)} L N(d_t - \sigma\sqrt{T-t}) \tag{18}$$

where r is the instantaneous risk-free rate, $N(\cdot)$ is the cumulative distribution function for standard normal random variable, and

$$d_t = \frac{\ln(V_t/L) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}. \tag{19}$$

According to Merton's model, the company's bankruptcy probability at time t is $N(-\mathrm{DTD}_t)$ where DTD_t denotes distance-to-default and it is

$$\mathrm{DTD}_t = \frac{\ln(V_t/L) + (\mu - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

To implement Merton's model, the so-called KMV assumption is typically adopted which sets $T-t$ to one year and L to the firm's book measure of short-term debt plus one half of its long-term debt. The KMV implementation assumption becomes problematic for financial firms. Financial firms typically have large amount of liabilities that are neither classified as short-term nor long-term debt, and thus the KMV assumption will grossly understate the amount of debt.

In order to deal with for financial firms, we follow Duan (2010) to include a firm's other liabilities which is adjusted by a fraction. Denote this unknown fraction by δ and note the resulting debt level used in estimation is a function of δ, i.e., $L(\delta)$. This unknown fraction can be estimated along with μ and σ. The KMV assumption can therefore be viewed a special case by setting $\delta = 0$. Our estimation method does not preclude the estimated fraction to become zero.

Following Duffie, Saita, and Wang (2007), we measure the short-term debt as the maximum of "Debt in current liabilities" and "Total current liabilities". A firm's other liabilities are defined as total liabilities minus short-term debt and then minus long-term debt. Hence, the liability measure $L(\delta)$ equals short-term debt plus one half of the long-term debt and plus a fraction of the other liabilities.

We then apply the maximum likelihood estimation method developed in Duan (1994, 2000) to estimate this unknown fraction parameter together with the asset return's mean and standard deviation. Since a firm's asset value could significantly change with a major investment and financing action, it makes more sense to standardize the firm's market value of assets by its book value so that the pure scaling effect will not distort the parameter values in the time series estimation. We thus divide the model's implied asset value by its book asset value in constructing the log-likelihood function. Obviously, if the book asset value stays unchanged throughout the sample period, such standardization will not have any effect. The log-likelihood function is

$$\begin{aligned}\mathscr{L}(\mu,\sigma,\delta) \quad = \quad & -\frac{n-1}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\ln(\sigma^2 h_t) - \sum_{t=2}^{n}\ln\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\right) - \sum_{t=2}^{n}\ln(N(\hat{d}_t(\sigma,\delta))) \\ & -\sum_{t=2}^{n}\frac{1}{2\sigma^2 h_t}\left[\ln\left(\frac{\hat{V}_t(\sigma,\delta)}{\hat{V}_{t-1}(\sigma,\delta)}\times\frac{A_{t-1}}{A_t}\right) - \left(\mu - \frac{\sigma^2}{2}\right)h_t\right]^2\end{aligned}$$

where n is the total number of equity values in the time series sample, $\hat{V}_t$ is the model's implied asset value solved using equation (18), $\hat{d}_t$ is computed using equation (19) with $\hat{V}_t$, A_t is the book asset value, and h_t is the length of time between two consecutive equity values (measured in trading days as a fraction of a year). Introducing h_t is mainly to take care of missing equity values in the sample. Note that δ becomes part of the log-likelihood function through $L(\delta)$.

To avoid the "look-ahead bias", we employ a rolling window method to estimate DTD. More specifically, at the end of each month, we estimate DTD for each firm using its daily market values of equity capitalization in the preceding year. We set the DTD to a missing value if there are less than 50 valid equity values in the preceding year. Whenever there are three or more consecutive equity values being identical, we will only consider the first and the last equity values in the sequence to be valid. The last valid DTD is used as the final DTD of each estimation.

References

[1] Altman, Edward I., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589-609.

[2] Amemiya, Takeshi, 1986, Advanced Econometrics, Basil Blackwell Ltd, UK.

[3] Azizpour, S., and K. Giesecke, 2008, Self-exciting corporate defaults: contagion vs. frailty, Stanford University working Paper.

[4] Beaver, William H., 1966, Financial ratios as predictors of failure, *Journal of Accounting Research* 4, 71-111.

[5] Beaver, William H., 1968, Market prices, financial ratios, and the prediction of failure, *Journal of Accounting Research* 6, 179-192.

[6] Beaver, W. H., M. F. McNichols, and J. W. Rhie, 2005, Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy, *Review of Accounting Studies* 10, 93-122.

[7] Bharath, S. T., and T. Shumway, 2008, Forecasting default with the Merton distance to default model, *Review of Financial Studies* 21, 1339-1369.

[8] Campbell, J. Y., J. Hilscher, and J. Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899-2939.

[9] Chava, Sudheer, and Robert A. Jarrow, 2004, Bankruptcy prediction with industry effects, *Review of Finance* 8, 537-569.

[10] Chen, C. J., 2007, The instantaneous and forward default intensity of structural models, University of Alberta working paper.

[11] Chuang, H.C., and C.M. Kuan, 2010, Predicting defaults with regime switching intensity: model and empirical evidence, National Taiwan University working Paper.

[12] Crosbie, P., and J. Bohn, 2002, Modeling default risk, technical report, KMV LLC.

[13] Das, S. R., D. Duffie, N. Kapadia, and L. Saita, 2007, Common failings: How corporate defaults are correlated, *Journal of Finance* 62, 93-117.

[14] Duan, J. C., 1994, Maximum likelihood estimation using price data of the derivative contract, *Mathematical Finance* 4, 155-167.

[15] Duan, J. C., 2000, Correction: "Maximum likelihood estimation using price data of the derivative contract", *Mathematical Finance* 10, 461-462.

[16] Duan, J. C., 2010, Clustered defaults, National University of Singapore working paper.

[17] Duffie, D., and D. Lando, 2001, Term structures of credit spreads with incomplete accounting information, *Econometrica* 69, 633-664.

[18] Duffie, D., A. Eckner, G. Horel, and L. Saita, 2009, Frailty correlated default, *Journal of Finance* 64, 2089-2123.

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

[20] 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.

[21] Koopman, S. J., A. Lucas, and A. Monteiro, 2008, The multi-state latent factor intensity model for credit rating transitions, *Journal of Econometrics* 142, 399-424.

[22] Koopman, S. J., R. Kraussl, A. Lucas, and A. B. Monteiro, 2009, Credit cycles and macro fundamentals, *Journal of Empirical Finance* 16, 42-54.

[23] Koopman, S.J., A. Lucas, and B. Shwaab, 2009, Macro, industry and frailty effects in defaults: The 2008 credit crisis in perspective, VU University Amsterdam working paper.

[24] Lando, D., and M. S. Nielsen, 2009, Correlation in corporate defaults: contagion or conditional independence, Copenhagen Business School working paper.

[25] Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449-470.

[26] Ohlson, J. A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109-131.

[27] Rhodes-Kropf, M., D.T. Robinson, and S. Viswanathan, 2005, Valuation waves and merger activity: the empirical evidence, *Journal of Financial Economics* 77, 561-603.

[28] Shumway, T., 2001, Forecasting bankruptcy more accurately: a simple hazard model, *Journal of Business* 74, 101-124.

[29] Vassalou, M., and Y. H. Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831-868.

[30] Zmijewski, Mark E., 1984, Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research* 22, 59-82.

Figure 1. Parameter estimates for the forward default intensity function

This figure shows the parameter estimates for the forward default intensity function corresponding to different prediction horizons. S&P500 is the trailing 1-year S&P500 index return, Treasury rate is the 3-month US Treasury rate, DTD is the distance to default, CASH/TA is the sum of cash and short-term investments over the total assets, NI/TA is the net income over the total assets, SIZE is log of firm's market equity over the average market equity value of the S&P500 company, M/B is the market to book equity value ratio, SIGMA is the 1-year idiosyncratic volatility. The subscript AVG denotes the average in the preceding 12 months, DIF denotes the difference between its current value and the preceding 12-month average. The solid line is for the parameter estimates and the dotted lines depict the 90% confidence interval.

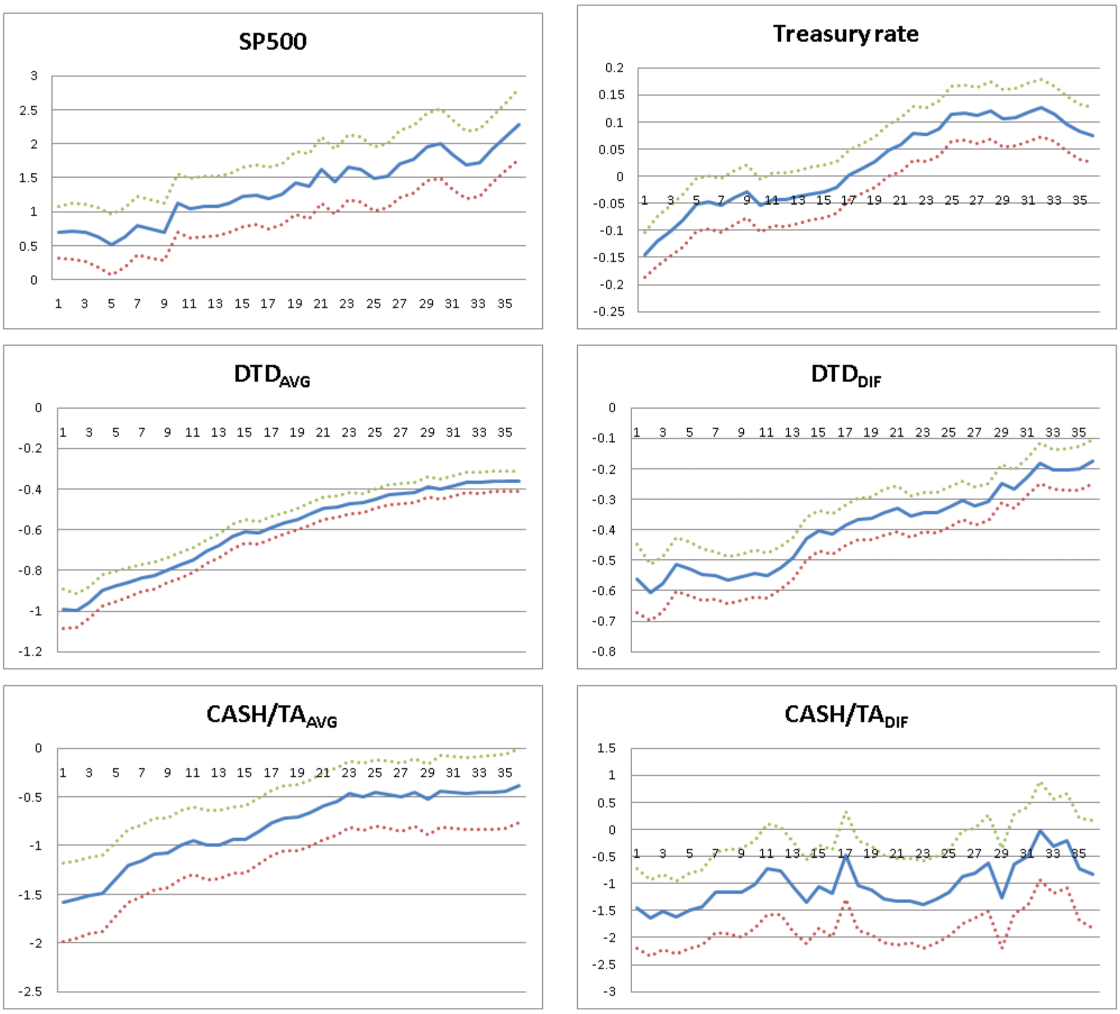

Figure 1. Parameter estimates for the forward default intensity function (Cont'd)

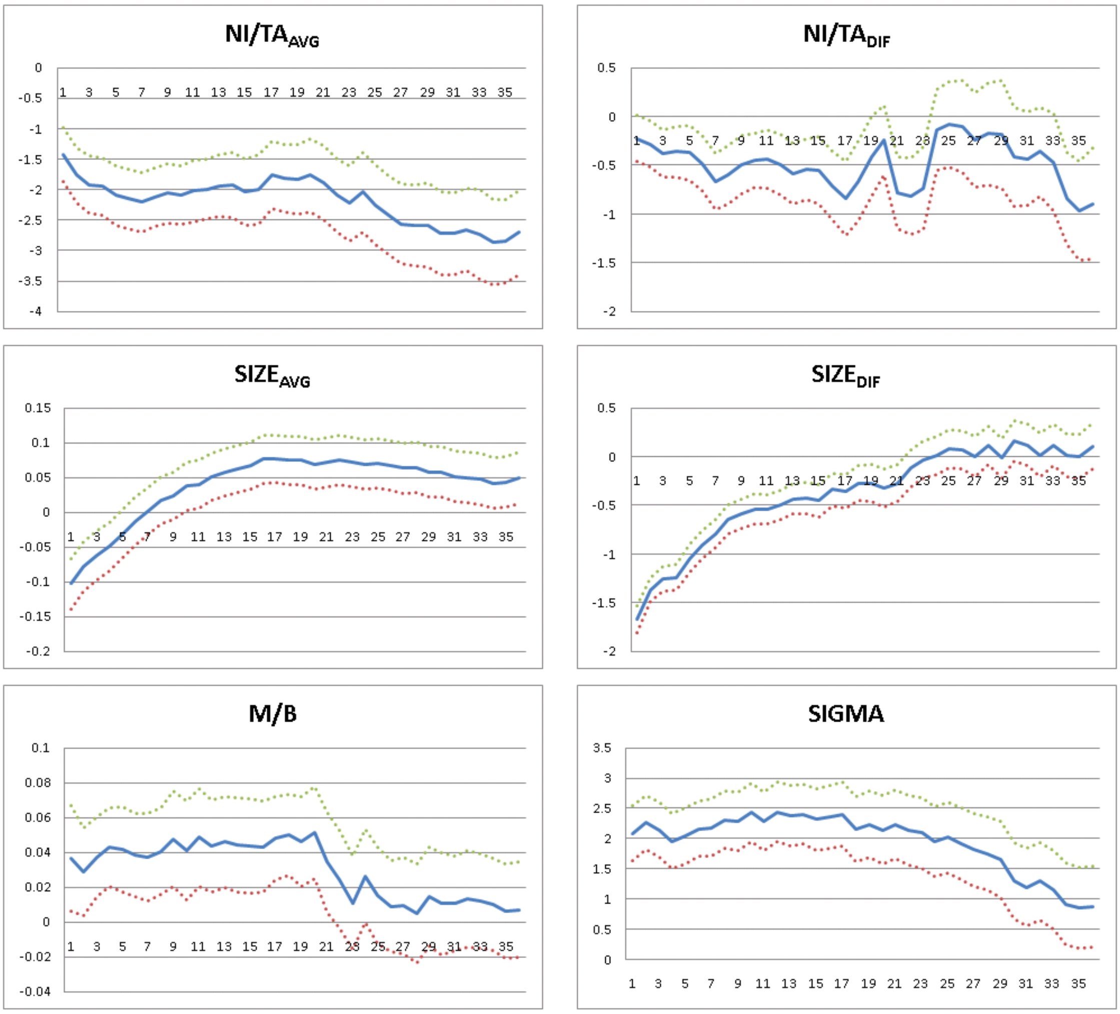

Figure 2. Aggregate number of defaults

This figure shows the observed (bars) and predicted (line) aggregate number of defaults for different prediction horizons. At each month-end, we compute the expected number of defaults in 1 month, 3 months, 6 months, 12 months, 24 months, 36 months and compare them with the observed values in the intended periods.

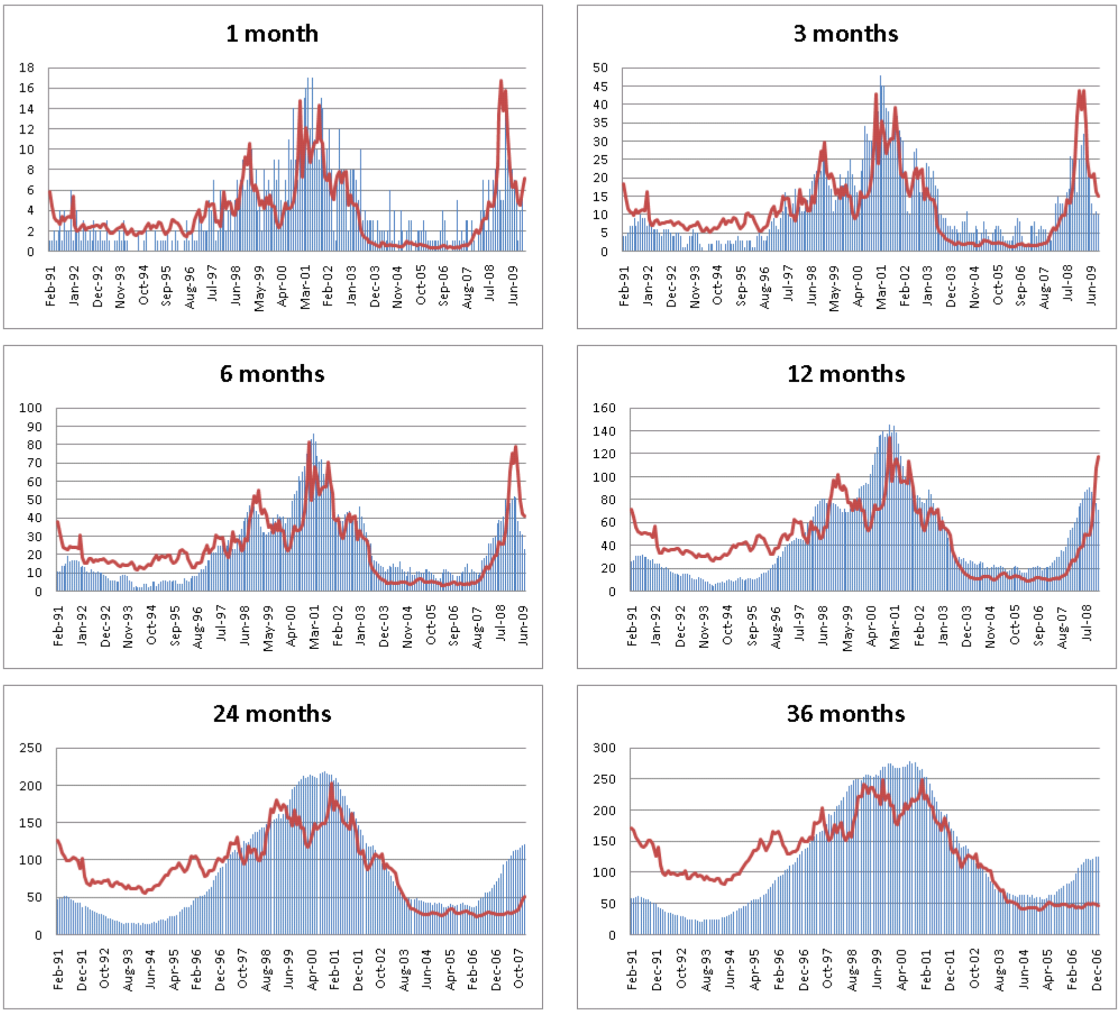

Figure 3A. In-sample cumulative accuracy profiles

This figure shows the in-sample cumulative accuracy profiles (power curves) based on all firms and the entire sample period (1991 to 2009) for different prediction horizons.

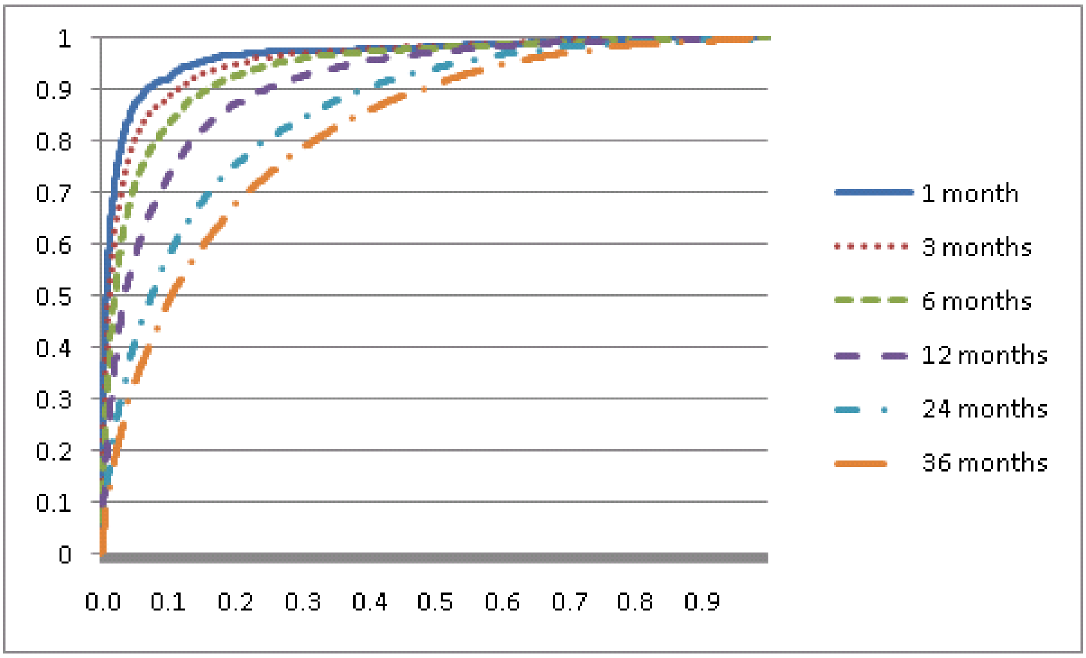

Figure 3B. Out-of-sample (cross-section) cumulative accuracy profiles

This figure shows the out-of-sample cumulative accuracy profiles (power curves) over the entire sample period (1991-2009) for different prediction horizons. We divide the firms equally into two groups: estimation group and evaluation group. We estimate the parameters based on the estimation group and then evaluate the prediction accuracy using the evaluation group.

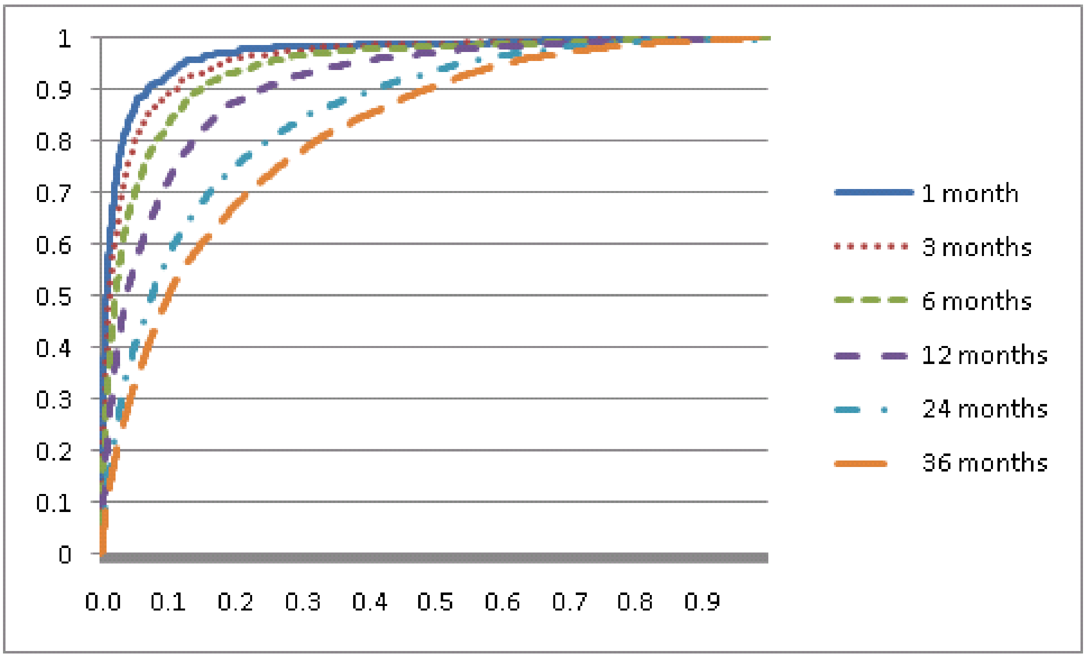

Figure 3C. Out-of-sample (over time) cumulative accuracy profiles

This figure shows the out-of-sample cumulative accuracy profiles (power curves) for the sample period (2001-2009) for different prediction horizons. We re-estimate the model at each month-end starting from the first month of 2001 and using only the data available at the of estimation.

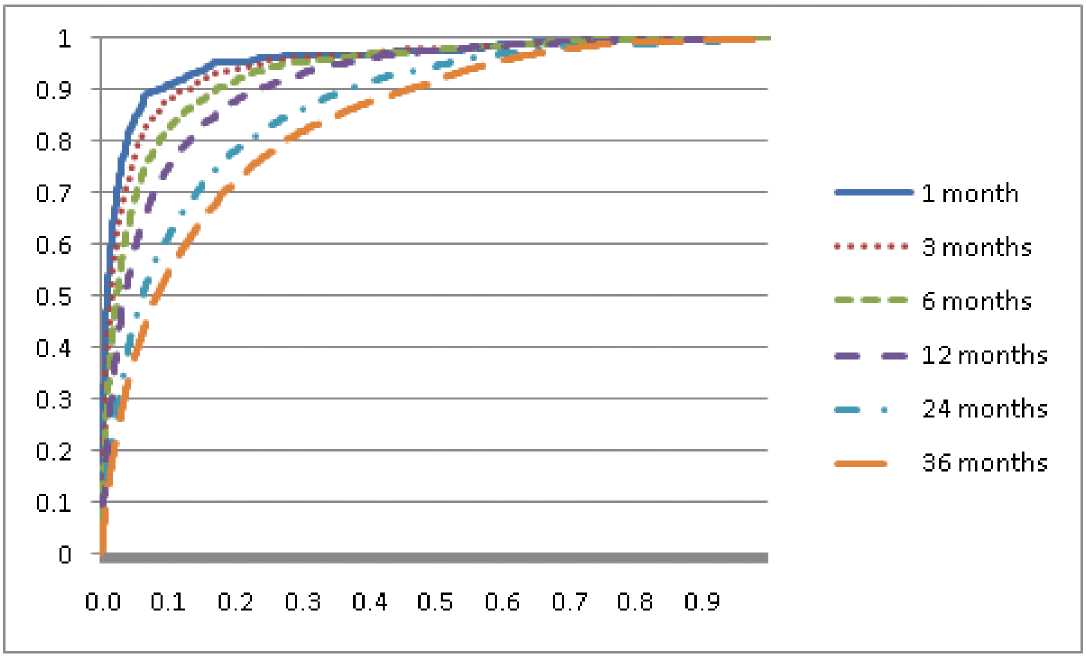

Figure 4. Lehman Brothers' term structure of forward and cumulative default probabilities

This figure shows the estimated term structure of forward default probabilities and that of cumulative default probabilities for Lehman Brothers, Merrill Lynch, Bank of America as well as the average values of the financial sector at 36 months, 24 months, 12 months and 3 months before Lehman Brothers' bankruptcy filing date (September 15, 2008).

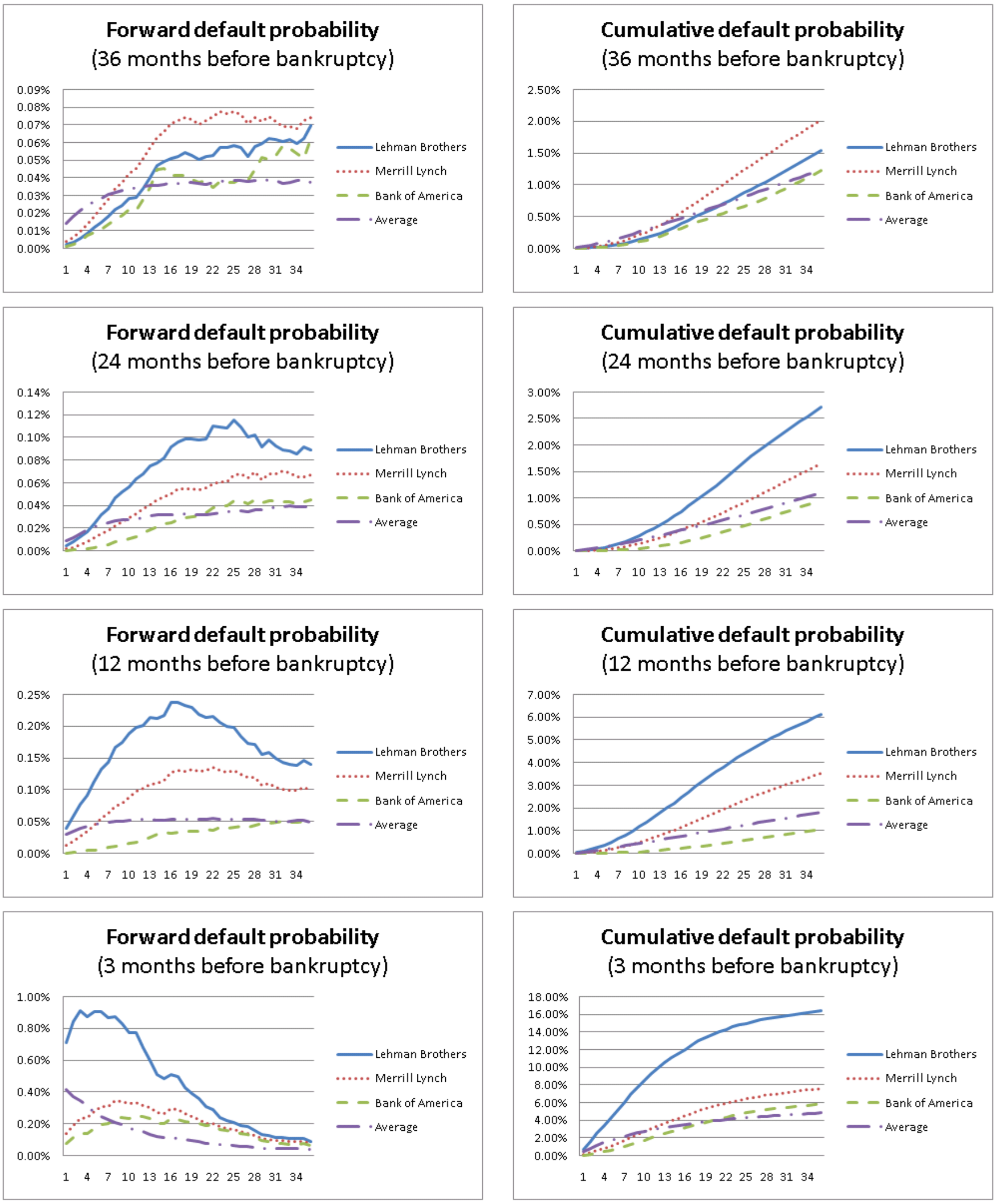

Table 1. Number of defaults and other exits

Total number of active firms, defaults/bankruptcies and other exits for each year over the sample period 1991-2009. The number of active firms is computed by averaging over the number of active firms across all months of the year.

Year	Active Firms	Defaults	(%)	Other Exit	(%)
1991	4031	26	0.65%	259	6.43%
1992	4029	24	0.60%	334	8.29%
1993	4203	14	0.33%	217	5.16%
1994	4442	10	0.23%	282	6.35%
1995	5079	9	0.18%	406	7.99%
1996	5473	12	0.22%	477	8.72%
1997	5659	43	0.76%	557	9.84%
1998	5725	65	1.14%	749	13.08%
1999	5444	76	1.40%	742	13.63%
2000	5109	96	1.88%	623	12.19%
2001	4936	158	3.20%	584	11.83%
2002	4700	85	1.81%	399	8.49%
2003	4363	65	1.49%	369	8.46%
2004	4101	25	0.61%	306	7.46%
2005	3947	24	0.61%	293	7.42%
2006	3877	18	0.46%	285	7.35%
2007	3786	21	0.55%	358	9.46%
2008	3694	60	1.62%	290	7.85%
2009	3598	73	2.03%	245	6.81%

Table 2. Summary statistics of firm-specific attributes

Summary statistics for the firm-specific attributes. DTD is the distance-to-default, CASH/TA is cash and short-term investments over the total assets, NI/TA is the net income over the total assets, SIZE is log of firm's market equity value over the average market equity value of an S&P500 company, M/B is the market to book ratio, SIGMA is the 1-year idiosyncratic volatility. The subscript AVG denotes the average in the previous 12 months, DIF denotes the difference between current value and its previous 12-month average.

	Mean	Std.	Min	25% Pctl	Median	75% Pctl	Max
DTD_{AVG}	3.610	2.754	-1.555	1.629	3.152	5.076	14.361
DTD_{DIF}	-0.009	1.746	-8.417	-0.983	0.000	0.980	9.046
$CASH/TA_{AVG}$	0.169	0.212	0.000	0.025	0.072	0.233	0.979
$CASH/TA_{DIF}$	-0.005	0.067	-0.455	-0.020	-0.002	0.013	0.434
NI/TA_{AVG}	-0.011	0.069	-0.793	-0.008	0.005	0.017	0.127
NI/TA_{DIF}	-0.002	0.071	-0.941	-0.007	0.000	0.007	0.611
$SIZE_{AVG}$	-4.372	2.042	-9.209	-5.865	-4.525	-3.046	2.277
$SIZE_{DIF}$	-0.038	0.378	-2.144	-0.204	-0.023	0.150	1.708
M/B	2.000	2.303	0.305	1.030	1.302	2.072	32.813
SIGMA	0.141	0.099	0.027	0.073	0.114	0.177	0.571

Table 3. Correlation matrix for firm-specific attributes

The correlation matrix for firm-specific attributes. DTD is the distance-to-default, CASH/TA is cash and short-term investments over the total assets, NI/TA is the net income over the total assets, SIZE is log of firm's market equity value over the average market equity value of an S&P500 company, M/B is the market to book ratio, SIGMA is the 1-year idiosyncratic volatility. The subscript AVG denotes the average in the previous 12 months, DIF denotes the difference between current value and its previous 12-month average.

	DTD_{AVG}	DTD_{DIF}	$CASH/TA_{AVG}$	$CASH/TA_{DIF}$	NI/TA_{AVG}	NI/TA_{DIF}	$SIZE_{AVG}$	$SIZE_{DIF}$	M/B	SIGMA
DTD_{AVG}	1.000	-0.174	0.028	0.038	0.252	0.030	0.499	0.150	0.189	-0.449
DTD_{DIF}	-0.174	1.000	0.019	0.020	-0.017	0.034	-0.044	0.433	0.092	0.012
$CASH/TA_{AVG}$	0.028	0.019	1.000	-0.156	-0.291	-0.023	-0.084	-0.030	0.337	0.231
$CASH/TA_{DIF}$	0.038	0.020	-0.156	1.000	0.050	0.088	0.026	0.103	-0.027	-0.024
NI/TA_{AVG}	0.252	-0.017	-0.291	0.050	1.000	-0.115	0.271	0.091	-0.261	-0.403
NI/TA_{DIF}	0.030	0.034	-0.023	0.088	-0.115	1.000	0.016	0.128	-0.019	-0.015
$SIZE_{AVG}$	0.499	-0.044	-0.084	0.026	0.271	0.016	1.000	0.066	0.101	-0.417
$SIZE_{DIF}$	0.150	0.433	-0.030	0.103	0.091	0.128	0.066	1.000	0.294	0.002
M/B	0.189	0.092	0.337	-0.027	-0.261	-0.019	0.101	0.294	1.000	0.199
SIGMA	-0.449	0.012	0.231	-0.024	-0.403	-0.015	-0.417	0.002	0.199	1.000

Table 4. Maximum pseudo-likelihood estimates for $\alpha(\tau)$

The maximum pseudo-likelihood estimates of $\alpha(\tau)$ for different prediction horizons. SP500 is the trailing 1-year S&P500 index return, Treasury rate is the 3-month US Treasury rate, DTD is the distance-to-default, CASH/TA is cash and short-term investments over the total assets, NI/TA is the net income over the total assets, SIZE is log of firm's market equity value over the average market equity value of an S&P500 company, M/B is the market to book ratio, SIGMA is the 1-year idiosyncratic volatility. The subscript AVG denotes the average in the previous 12 months, DIF denotes the difference between current value and its previous 12-month average. *** denotes significance at 1%, ** denotes significance at 5% and * denotes significance at 10%.

Panel A: Maximum pseudo-likelihood estimates for $\alpha(\tau)$ (1-12 months)												
	1	2	3	4	5	6	7	8	9	10	11	12
Intercept	-4.885***	-4.505***	-4.263***	-4.108***	-3.979***	-3.837***	-3.679***	-3.591***	-3.559***	-3.411***	-3.421***	-3.378***
	(0.173)	(0.175)	(0.174)	(0.176)	(0.178)	(0.178)	(0.182)	(0.179)	(0.181)	(0.184)	(0.185)	(0.184)
SP500	0.700***	0.714***	0.694***	0.630**	0.511*	0.625**	0.800***	0.749***	0.703***	1.124***	1.045***	1.078***
	(0.234)	(0.251)	(0.257)	(0.267)	(0.270)	(0.267)	(0.262)	(0.260)	(0.259)	(0.259)	(0.266)	(0.271)
Treasury rate	-0.145***	-0.120***	-0.101***	-0.082***	-0.052*	-0.048	-0.054*	-0.039	-0.028	-0.055*	-0.043	-0.044
	(0.026)	(0.027)	(0.028)	(0.029)	(0.030)	(0.030)	(0.030)	(0.030)	(0.029)	(0.029)	(0.030)	(0.030)
DTD_{AVG}	-0.990***	-0.998***	-0.959***	-0.898***	-0.878***	-0.859***	-0.835***	-0.827***	-0.799***	-0.772***	-0.746***	-0.706***
	(0.059)	(0.050)	(0.047)	(0.045)	(0.045)	(0.043)	(0.041)	(0.039)	(0.038)	(0.039)	(0.037)	(0.036)
DTD_{DIF}	-0.560***	-0.607***	-0.578***	-0.513***	-0.529***	-0.547***	-0.551***	-0.566***	-0.556***	-0.543***	-0.551***	-0.526***
	(0.068)	(0.057)	(0.055)	(0.055)	(0.054)	(0.051)	(0.048)	(0.047)	(0.046)	(0.047)	(0.044)	(0.043)
$CASH/TA_{AVG}$	-1.583***	-1.549***	-1.510***	-1.488***	-1.344***	-1.204***	-1.157***	-1.086***	-1.075***	-0.990***	-0.949***	-0.994***
	(0.245)	(0.241)	(0.238)	(0.237)	(0.231)	(0.227)	(0.222)	(0.221)	(0.218)	(0.215)	(0.212)	(0.216)
$CASH/TA_{DIF}$	-1.458***	-1.639***	-1.522***	-1.619***	-1.499***	-1.440***	-1.164***	-1.152**	-1.166**	-1.017**	-0.730	-0.766
	(0.449)	(0.429)	(0.424)	(0.410)	(0.422)	(0.426)	(0.451)	(0.475)	(0.496)	(0.497)	(0.511)	(0.493)
NI/TA_{AVG}	-1.432***	-1.765***	-1.918***	-1.950***	-2.094***	-2.150***	-2.208***	-2.126***	-2.056***	-2.089***	-2.021***	-1.991***
	(0.268)	(0.279)	(0.284)	(0.291)	(0.296)	(0.295)	(0.292)	(0.293)	(0.298)	(0.296)	(0.307)	(0.296)
NI/TA_{DIF}	-0.229	-0.284*	-0.381**	-0.362**	-0.372**	-0.480***	-0.668***	-0.597***	-0.496***	-0.451***	-0.440**	-0.498***
	(0.144)	(0.145)	(0.149)	(0.158)	(0.171)	(0.171)	(0.174)	(0.180)	(0.181)	(0.169)	(0.182)	(0.188)
$SIZE_{AVG}$	-0.103***	-0.078***	-0.062***	-0.050**	-0.030	-0.013	0.001	0.017	0.023	0.037*	0.040*	0.051**
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$SIZE_{DIF}$	-1.670***	-1.366***	-1.258***	-1.238***	-1.044***	-0.908***	-0.791***	-0.646***	-0.588***	-0.536***	-0.542***	-0.493***
	(0.084)	(0.076)	(0.079)	(0.082)	(0.087)	(0.088)	(0.088)	(0.090)	(0.092)	(0.093)	(0.093)	(0.091)
M/B	0.037**	0.029*	0.037***	0.043***	0.042***	0.038***	0.037**	0.041***	0.048***	0.041**	0.049***	0.044***
	(0.018)	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)	(0.017)	(0.016)
SIGMA	2.090***	2.265***	2.144***	1.961***	2.047***	2.163***	2.183***	2.311***	2.288***	2.436***	2.287***	2.439***
	(0.277)	(0.272)	(0.278)	(0.279)	(0.278)	(0.279)	(0.283)	(0.288)	(0.290)	(0.290)	(0.291)	(0.297)

Panel B: Maximum pseudo-likelihood estimates for $\alpha(\tau)$ (13-24 months)

	13	14	15	16	17	18	19	20	21	22	23	24
Intercept	-3.340***	-3.362***	-3.377***	-3.341***	-3.469***	-3.512***	-3.616***	-3.780***	-3.872***	-3.899***	-3.926***	-3.965***
	(0.185)	(0.189)	(0.190)	(0.185)	(0.182)	(0.183)	(0.186)	(0.195)	(0.197)	(0.194)	(0.198)	(0.194)
SP500	1.085***	1.127***	1.218***	1.245***	1.195***	1.252***	1.425***	1.369***	1.620***	1.444***	1.656***	1.617***
	(0.268)	(0.263)	(0.266)	(0.267)	(0.275)	(0.271)	(0.284)	(0.292)	(0.296)	(0.293)	(0.293)	(0.292)
Treasury rate	-0.038	-0.032	-0.028	-0.021	0.002	0.015	0.026	0.048	0.058*	0.079***	0.077**	0.088***
	(0.030)	(0.029)	(0.030)	(0.029)	(0.028)	(0.028)	(0.028)	(0.029)	(0.030)	(0.030)	(0.031)	(0.031)
DTD_{AVG}	-0.679***	-0.633***	-0.609***	-0.614***	-0.587***	-0.569***	-0.548***	-0.520***	-0.493***	-0.487***	-0.471***	-0.469***
	(0.036)	(0.037)	(0.036)	(0.034)	(0.033)	(0.033)	(0.033)	(0.034)	(0.033)	(0.031)	(0.031)	(0.030)
DTD_{DIF}	-0.493***	-0.430***	-0.403***	-0.415***	-0.385***	-0.365***	-0.363***	-0.344***	-0.330***	-0.357***	-0.343***	-0.345***
	(0.041)	(0.043)	(0.040)	(0.040)	(0.040)	(0.042)	(0.043)	(0.046)	(0.046)	(0.042)	(0.040)	(0.041)
$CASH/TA_{AVG}$	-0.992***	-0.943***	-0.938***	-0.852***	-0.763***	-0.722***	-0.710***	-0.665***	-0.594***	-0.542**	-0.470**	-0.501**
	(0.214)	(0.210)	(0.210)	(0.207)	(0.202)	(0.204)	(0.205)	(0.207)	(0.208)	(0.211)	(0.210)	(0.212)
$CASH/TA_{DIF}$	-1.060**	-1.339***	-1.060**	-1.183**	-0.482	-1.036**	-1.125**	-1.289***	-1.336***	-1.321***	-1.396***	-1.295***
	(0.501)	(0.471)	(0.470)	(0.487)	(0.491)	(0.501)	(0.496)	(0.490)	(0.489)	(0.475)	(0.491)	(0.491)
NI/TA_{AVG}	-1.937***	-1.929***	-2.040***	-2.001***	-1.758***	-1.819***	-1.836***	-1.764***	-1.895***	-2.095***	-2.226***	-2.040***
	(0.307)	(0.324)	(0.331)	(0.350)	(0.332)	(0.337)	(0.349)	(0.365)	(0.369)	(0.361)	(0.372)	(0.395)
NI/TA_{DIF}	-0.587***	-0.541***	-0.554***	-0.710***	-0.842***	-0.663***	-0.412*	-0.243	-0.784***	-0.813***	-0.739***	-0.139
	(0.189)	(0.189)	(0.211)	(0.215)	(0.232)	(0.251)	(0.248)	(0.219)	(0.226)	(0.238)	(0.260)	(0.251)
$SIZE_{AVG}$	0.058***	0.063***	0.067***	0.076***	0.077***	0.074***	0.075***	0.069***	0.072***	0.075***	0.072***	0.068***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)
$SIZE_{DIF}$	-0.436***	-0.426***	-0.454***	-0.337***	-0.356***	-0.272**	-0.270**	-0.324***	-0.273**	-0.119	-0.030	0.011
	(0.094)	(0.099)	(0.099)	(0.100)	(0.106)	(0.107)	(0.113)	(0.117)	(0.116)	(0.116)	(0.117)	(0.121)
M/B	0.046***	0.044***	0.044***	0.043***	0.048***	0.050***	0.046***	0.051***	0.034**	0.025	0.011	0.026
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)
SIGMA	2.385***	2.405***	2.316***	2.353***	2.404***	2.156***	2.237***	2.142***	2.237***	2.139***	2.100***	1.945***
	(0.304)	(0.302)	(0.313)	(0.315)	(0.318)	(0.329)	(0.334)	(0.342)	(0.343)	(0.349)	(0.356)	(0.350)

Panel C: Maximum pseudo-likelihood estimates for $\alpha(\tau)$ (25-36 months)

	25	26	27	28	29	30	31	32	33	34	35	36
Intercept	-4.079***	-4.112***	-4.144***	-4.163***	-4.214***	-4.136***	-4.188***	-4.269***	-4.193***	-4.150***	-4.096***	-4.060***
	(0.196)	(0.200)	(0.206)	(0.207)	(0.214)	(0.212)	(0.212)	(0.214)	(0.214)	(0.215)	(0.213)	(0.213)
SP500	1.486***	1.531***	1.708***	1.769***	1.955***	2.005***	1.831***	1.689***	1.723***	1.912***	2.105***	2.289***
	(0.288)	(0.288)	(0.299)	(0.298)	(0.298)	(0.309)	(0.311)	(0.306)	(0.299)	(0.295)	(0.303)	(0.314)
Treasury rate	0.115***	0.117***	0.112***	0.121***	0.107***	0.109***	0.118***	0.126***	0.115***	0.097***	0.083***	0.075**
	(0.031)	(0.031)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.031)	(0.031)	(0.031)	(0.031)
DTD_{AVG}	-0.448***	-0.430***	-0.423***	-0.415***	-0.389***	-0.401***	-0.385***	-0.367***	-0.370***	-0.363***	-0.361***	-0.361***
	(0.029)	(0.030)	(0.030)	(0.030)	(0.031)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
DTD_{DIF}	-0.326***	-0.303***	-0.322***	-0.308***	-0.247***	-0.268***	-0.225***	-0.183***	-0.204***	-0.204***	-0.200***	-0.175***
	(0.040)	(0.039)	(0.039)	(0.037)	(0.038)	(0.038)	(0.038)	(0.040)	(0.039)	(0.041)	(0.044)	(0.044)
$CASH/TA_{AVG}$	-0.459**	-0.479**	-0.504**	-0.455**	-0.526**	-0.443**	-0.454**	-0.463**	-0.455**	-0.454**	-0.444*	-0.384*
	(0.210)	(0.211)	(0.215)	(0.213)	(0.219)	(0.222)	(0.225)	(0.225)	(0.228)	(0.231)	(0.233)	(0.233)
$CASH/TA_{DIF}$	-1.169**	-0.878*	-0.807	-0.618	-1.272**	-0.636	-0.488	-0.026	-0.303	-0.205	-0.730	-0.829
	(0.488)	(0.520)	(0.506)	(0.539)	(0.559)	(0.563)	(0.554)	(0.555)	(0.529)	(0.529)	(0.573)	(0.601)
NI/TA_{AVG}	-2.247***	-2.418***	-2.563***	-2.592***	-2.579***	-2.718***	-2.720***	-2.658***	-2.735***	-2.860***	-2.848***	-2.705***
	(0.396)	(0.389)	(0.396)	(0.402)	(0.419)	(0.418)	(0.408)	(0.409)	(0.443)	(0.426)	(0.414)	(0.421)
NI/TA_{DIF}	-0.076	-0.101	-0.238	-0.173	-0.187	-0.412	-0.432	-0.360	-0.466	-0.843***	-0.964***	-0.893***
	(0.265)	(0.287)	(0.295)	(0.319)	(0.336)	(0.309)	(0.291)	(0.277)	(0.306)	(0.288)	(0.307)	(0.347)
$SIZE_{AVG}$	0.070***	0.067***	0.063***	0.065***	0.058***	0.058***	0.051**	0.050**	0.048**	0.042*	0.043*	0.050**
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$SIZE_{DIF}$	0.081	0.071	0.000	0.114	-0.012	0.163	0.119	0.015	0.120	0.009	0.002	0.108
	(0.119)	(0.118)	(0.125)	(0.122)	(0.122)	(0.124)	(0.128)	(0.137)	(0.129)	(0.132)	(0.136)	(0.142)
M/B	0.015	0.009	0.009	0.005	0.015	0.011	0.011	0.013	0.012	0.010	0.006	0.007
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)
SIGMA	2.017***	1.917***	1.814***	1.755***	1.657***	1.307***	1.194***	1.302***	1.153***	0.917**	0.855**	0.874**
	(0.352)	(0.362)	(0.366)	(0.371)	(0.379)	(0.382)	(0.390)	(0.389)	(0.398)	(0.412)	(0.407)	(0.409)

Table 5. Maximum pseudo-likelihood estimates for $\beta(\tau)$

This table presents the maximum pseudo-likelihood estimates of $\beta(\tau)$ for different horizons. SP500 is the trailing 1-year S&P500 index return, Treasury rate is the 3-month US Treasury rate, DTD is the distance-to-default, CASH/TA is cash and short-term investments over the total assets, NI/TA is the net income over the total assets, SIZE is log of firm's market equity value over the average market equity value of an S&P500 company, M/B is the market to book ratio, SIGMA is the 1-year idiosyncratic volatility. The subscript AVG denotes the average in the previous 12 months, DIF denotes the difference between current value and its previous 12-month average. *** denotes significance at 1%, ** denotes significance at 5% and * denotes significance at 10%.

Panel A: Maximum pseudo-likelihood estimates for $\beta(\tau)$ (1-12 months)												
	1	2	3	4	5	6	7	8	9	10	11	12
Intercept	-4.314***	-4.192***	-4.070***	-3.991***	-3.895***	-3.817***	-3.750***	-3.725***	-3.701***	-3.661***	-3.642***	-3.622***
	(0.059)	(0.060)	(0.061)	(0.062)	(0.063)	(0.065)	(0.065)	(0.066)	(0.067)	(0.068)	(0.068)	(0.069)
SP500	0.071	0.187**	0.336***	0.332***	0.384***	0.514***	0.569***	0.621***	0.642***	0.729***	0.644***	0.684***
	(0.079)	(0.080)	(0.081)	(0.081)	(0.083)	(0.083)	(0.084)	(0.085)	(0.086)	(0.086)	(0.087)	(0.089)
Treasury rate	0.049***	0.044***	0.038***	0.042***	0.039***	0.035***	0.037***	0.038***	0.041***	0.040***	0.046***	0.044***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
$\mathrm{DTD}_{\mathrm{AVG}}$	0.067***	0.052***	0.040***	0.032***	0.022***	0.011*	0.004	-0.002	-0.007	-0.013*	-0.018**	-0.021***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$\mathrm{DTD}_{\mathrm{DIF}}$	0.181***	0.157***	0.140***	0.120***	0.097***	0.067***	0.048***	0.031***	0.025**	0.012	0.005	-0.000
	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
$\mathrm{CASH/TA}_{\mathrm{AVG}}$	-0.599***	-0.526***	-0.445***	-0.361***	-0.285***	-0.239***	-0.182***	-0.155**	-0.132**	-0.101*	-0.092	-0.083
	(0.059)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.061)	(0.061)	(0.062)	(0.063)
$\mathrm{CASH/TA}_{\mathrm{DIF}}$	-0.489***	-0.626***	-0.609***	-0.670***	-0.671***	-0.649***	-0.663***	-0.568***	-0.432***	-0.367**	-0.425**	-0.453***
	(0.141)	(0.146)	(0.148)	(0.152)	(0.155)	(0.155)	(0.158)	(0.159)	(0.161)	(0.162)	(0.165)	(0.167)
$\mathrm{NI/TA}_{\mathrm{AVG}}$	-2.039***	-2.028***	-2.111***	-2.172***	-2.225***	-2.244***	-2.302***	-2.293***	-2.251***	-2.233***	-2.198***	-2.122***
	(0.123)	(0.130)	(0.132)	(0.134)	(0.137)	(0.140)	(0.143)	(0.145)	(0.146)	(0.147)	(0.148)	(0.149)
$\mathrm{NI/TA}_{\mathrm{DIF}}$	-1.169***	-1.222***	-1.073***	-0.949***	-0.815***	-0.637***	-0.591***	-0.531***	-0.545***	-0.511***	-0.515***	-0.441***
	(0.078)	(0.082)	(0.087)	(0.089)	(0.094)	(0.097)	(0.098)	(0.103)	(0.107)	(0.108)	(0.111)	(0.116)
$\mathrm{SIZE}_{\mathrm{AVG}}$	-0.222***	-0.218***	-0.216***	-0.215***	-0.213***	-0.212***	-0.208***	-0.206***	-0.202***	-0.198***	-0.196***	-0.195***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$\mathrm{SIZE}_{\mathrm{DIF}}$	-0.568***	-0.549***	-0.588***	-0.650***	-0.685***	-0.688***	-0.680***	-0.659***	-0.604***	-0.541***	-0.510***	-0.483***
	(0.033)	(0.034)	(0.036)	(0.035)	(0.035)	(0.036)	(0.037)	(0.038)	(0.039)	(0.039)	(0.039)	(0.040)
M/B	-0.042***	-0.048***	-0.057***	-0.068***	-0.073***	-0.068***	-0.074***	-0.068***	-0.062***	-0.060***	-0.053***	-0.046***
	(0.007)	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)
SIGMA	2.726***	2.557***	2.251***	1.928***	1.624***	1.353***	1.124***	1.046***	1.033***	0.990***	0.878***	0.842***
	(0.102)	(0.103)	(0.106)	(0.109)	(0.113)	(0.117)	(0.121)	(0.124)	(0.126)	(0.128)	(0.130)	(0.131)

Panel B: Maximum pseudo-likelihood estimates for $\beta(\tau)$ (13-24 months)

	13	14	15	16	17	18	19	20	21	22	23	24
Intercept	-3.574***	-3.538***	-3.543***	-3.519***	-3.486***	-3.447***	-3.388***	-3.329***	-3.295***	-3.252***	-3.239***	-3.228***
	(0.069)	(0.070)	(0.071)	(0.071)	(0.072)	(0.072)	(0.072)	(0.072)	(0.073)	(0.073)	(0.074)	(0.074)
SP500	0.723***	0.772***	0.839***	0.804***	0.918***	0.904***	0.968***	1.057***	1.027***	1.030***	1.113***	1.104***
	(0.088)	(0.090)	(0.093)	(0.093)	(0.095)	(0.094)	(0.095)	(0.095)	(0.095)	(0.097)	(0.098)	(0.098)
Treasury rate	0.040***	0.037***	0.035***	0.033***	0.023**	0.020**	0.013	0.002	0.003	-0.002	-0.005	-0.004
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
DTD_{AVG}	-0.025***	-0.026***	-0.024***	-0.026***	-0.025***	-0.024***	-0.028***	-0.030***	-0.033***	-0.036***	-0.035***	-0.037***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
DTD_{DIF}	-0.005	-0.009	-0.006	-0.010	-0.007	0.001	-0.001	-0.005	-0.010	-0.017	-0.012	-0.014
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
$\text{CASH/TA}_{\text{AVG}}$	-0.069	-0.045	-0.032	-0.024	-0.026	-0.007	0.018	0.024	0.041	0.051	0.061	0.071
	(0.063)	(0.064)	(0.064)	(0.065)	(0.066)	(0.066)	(0.067)	(0.067)	(0.068)	(0.069)	(0.070)	(0.070)
$\text{CASH/TA}_{\text{DIF}}$	-0.488***	-0.340**	-0.379**	-0.493***	-0.621***	-0.639***	-0.544***	-0.485***	-0.491***	-0.466**	-0.614***	-0.504***
	(0.167)	(0.165)	(0.166)	(0.168)	(0.173)	(0.175)	(0.177)	(0.182)	(0.185)	(0.184)	(0.184)	(0.188)
$\text{NI/TA}_{\text{AVG}}$	-2.130***	-2.100***	-2.071***	-1.985***	-1.956***	-1.976***	-1.925***	-1.902***	-1.897***	-1.852***	-1.871***	-1.906***
	(0.151)	(0.155)	(0.159)	(0.162)	(0.163)	(0.164)	(0.169)	(0.173)	(0.175)	(0.174)	(0.178)	(0.177)
$\text{NI/TA}_{\text{DIF}}$	-0.492***	-0.529***	-0.648***	-0.572***	-0.568***	-0.482***	-0.355***	-0.404***	-0.434***	-0.658***	-0.472***	-0.475***
	(0.117)	(0.120)	(0.119)	(0.124)	(0.126)	(0.134)	(0.137)	(0.144)	(0.146)	(0.146)	(0.152)	(0.147)
SIZE_{AVG}	-0.191***	-0.188***	-0.187***	-0.184***	-0.182***	-0.180***	-0.173***	-0.167***	-0.163***	-0.162***	-0.159***	-0.158***
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
SIZE_{DIF}	-0.448***	-0.419***	-0.421***	-0.390***	-0.426***	-0.434***	-0.378***	-0.379***	-0.323***	-0.292***	-0.286***	-0.272***
	(0.040)	(0.040)	(0.041)	(0.041)	(0.042)	(0.043)	(0.044)	(0.044)	(0.044)	(0.045)	(0.046)	(0.047)
M/B	-0.045***	-0.047***	-0.046***	-0.043***	-0.042***	-0.043***	-0.041***	-0.039***	-0.040***	-0.038***	-0.042***	-0.038***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
SIGMA	0.800***	0.782***	0.815***	0.824***	0.785***	0.654***	0.706***	0.731***	0.719***	0.658***	0.673***	0.594***
	(0.133)	(0.135)	(0.137)	(0.138)	(0.141)	(0.145)	(0.148)	(0.151)	(0.152)	(0.155)	(0.157)	(0.159)

Panel C: Maximum pseudo-likelihood estimates for $\beta(\tau)$ (25-36 months)

	25	26	27	28	29	30	31	32	33	34	35	36
Intercept	-3.215***	-3.229***	-3.218***	-3.208***	-3.202***	-3.174***	-3.153***	-3.140***	-3.134***	-3.118***	-3.116***	-3.132***
	(0.075)	(0.075)	(0.076)	(0.076)	(0.076)	(0.076)	(0.077)	(0.077)	(0.077)	(0.078)	(0.079)	(0.079)
SP500	1.083***	1.011***	0.989***	1.025***	0.943***	0.982***	1.023***	0.998***	0.941***	0.917***	0.869***	0.866***
	(0.097)	(0.098)	(0.098)	(0.099)	(0.098)	(0.099)	(0.100)	(0.103)	(0.103)	(0.105)	(0.105)	(0.104)
Treasury rate	-0.004	0.000	0.002	0.001	0.005	0.003	-0.000	0.000	0.003	0.002	0.004	0.002
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
DTD_{AVG}	-0.035***	-0.035***	-0.038***	-0.042***	-0.043***	-0.045***	-0.046***	-0.045***	-0.046***	-0.052***	-0.052***	-0.049***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
DTD_{DIF}	-0.014	-0.005	-0.003	-0.002	-0.008	-0.012	-0.017	-0.024**	-0.023**	-0.027**	-0.033***	-0.024**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)
$CASH/TA_{AVG}$	0.097	0.089	0.136*	0.146**	0.137*	0.159**	0.167**	0.182**	0.180**	0.159**	0.158**	0.169**
	(0.071)	(0.071)	(0.072)	(0.072)	(0.073)	(0.073)	(0.074)	(0.074)	(0.074)	(0.076)	(0.076)	(0.077)
$CASH/TA_{DIF}$	-0.454**	-0.414**	-0.395**	-0.220	-0.253	-0.163	0.168	0.130	0.007	-0.065	-0.144	-0.268
	(0.183)	(0.182)	(0.186)	(0.190)	(0.196)	(0.200)	(0.205)	(0.212)	(0.215)	(0.210)	(0.210)	(0.217)
NI/TA_{AVG}	-1.957***	-1.883***	-1.843***	-1.737***	-1.769***	-1.752***	-1.758***	-1.747***	-1.746***	-1.626***	-1.688***	-1.628***
	(0.182)	(0.183)	(0.186)	(0.192)	(0.197)	(0.199)	(0.202)	(0.203)	(0.206)	(0.211)	(0.212)	(0.213)
NI/TA_{DIF}	-0.489***	-0.427***	-0.339**	-0.114	-0.324*	-0.416**	-0.520***	-0.393**	-0.338*	-0.356**	-0.401**	-0.337*
	(0.147)	(0.156)	(0.166)	(0.174)	(0.166)	(0.165)	(0.164)	(0.168)	(0.180)	(0.181)	(0.181)	(0.183)
$SIZE_{AVG}$	-0.157***	-0.155***	-0.154***	-0.154***	-0.153***	-0.149***	-0.146***	-0.143***	-0.142***	-0.143***	-0.145***	-0.145***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
$SIZE_{DIF}$	-0.265***	-0.285***	-0.251***	-0.257***	-0.233***	-0.194***	-0.173***	-0.162***	-0.152***	-0.138***	-0.114**	-0.108**
	(0.047)	(0.047)	(0.048)	(0.048)	(0.048)	(0.049)	(0.049)	(0.050)	(0.050)	(0.051)	(0.051)	(0.052)
M/B	-0.043***	-0.036***	-0.034***	-0.027***	-0.025***	-0.027***	-0.027***	-0.030***	-0.026***	-0.014*	-0.016**	-0.018**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
SIGMA	0.561***	0.590***	0.512***	0.434***	0.400**	0.426**	0.443***	0.459***	0.397**	0.311*	0.250	0.342*
	(0.161)	(0.162)	(0.164)	(0.166)	(0.166)	(0.167)	(0.169)	(0.171)	(0.172)	(0.174)	(0.177)	(0.179)

Table 6. Accuracy ratios

This table reports the accuracy ratios derived from the cumulative accuracy profiles based on rank orders. Panel A reports the in-sample results for all firms and the entire sample period (1991-2009). Panel B presents the out-of-sample (cross-section) results for the entire sample period (1991-2009) where we equally divide the firms into two groups: estimation group and evaluation group. We estimate the parameters based on the estimation group and then evaluate the prediction accuracy using the evaluation group. Panel C reports the out-of-sample (over time) results for the sample period (2001-2009). We re-estimate the model at each month-end starting from the first month of 2001 and using only the data available at the of estimation.

Panel A: In-sample result					
1 month	3 months	6 months	12 months	24 months	36 months
93.26%	91.34%	88.39%	82.92%	72.74%	65.80%
Panel B: Out-of-sample (cross-section) result					
1 month	3 months	6 months	12 months	24 months	36 months
93.89%	91.80%	88.62%	82.61%	72.14%	65.55%
Panel C: Out-of-sample (over time) result					
1 month	3 months	6 months	12 months	24 months	36 months
91.70%	90.04%	87.41%	83.77%	75.16%	69.78%