Multiperiod Corporate Default Prediction – A Forward Intensity Approach

Jin-Chuan Duan, Jie Sun and Tao Wang*

(First Version: July 14, 2010; This Version: August 21, 2011)

Abstract

A forward intensity model for the prediction of corporate defaults over different future periods is proposed. Maximum pseudo-likelihood analysis is then conducted on a large sample of the US industrial and financial firms spanning the period 1991-2010 on a monthly basis. Several commonly used factors and firm-specific attributes are shown to be useful for prediction at both short and long horizons. Our implementation also factors in momentum in some variables and documents their importance in default prediction. 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 for an analysis of 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 Oversea-Chinese Banking Corporation Limited. Email: sunj3@ocbc.com. Wang is with Department of Finance, National University of Singapore. Email: wangtao@nus.edu.sg.

1 Introduction

Understanding the determinants of default 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. Naturally, default analysis should address the term structure effect. Major credit rating agencies usually provide both short-term and long-term credit ratings for corporates, but the information provided by them typically lacks granularity. The academic literature has also been lacking as far as the term structure of default probabilities is concerned. Credit risk modeling can be grouped into two large categories – structural and reduced-form approaches. In this paper, we model default term structures by 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, attempting 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) (DSW, henceforth) proposed a doubly stochastic Poisson intensity model to describe the occurrences of default. The state variables governing the Poisson intensities in their model are assumed to follow a specific high-dimensional vector autoregressive process. Even if one restricts the consideration to a small set of firm attributes, a large number of firms will simply make the overall dimension of the state variables very high and render the DSW model's practical implementation difficult. The need to specify time dynamics of the state variables in the DSW model is related to multiperiod default prediction, i.e, generating the term structure of default probabilities. Modelling a high-dimension time series faces challenges in two aspects. First, it is difficult to come up with a model that can adequately capture the joint behavior of the attributes for many firms. Second, implementing the model needs values for the parameters governing the autoregressive process. Parameter estimation, however, inevitably requires one to compute the inverse and determinant of the state variables' variance-covariance matrix. When the dimension is high, computational obstacles (computing time and numerical accuracy) quickly surface. For example, the DSW model employs two common state variables for all firms and two firm-specific attributes for each of the firms under consideration. If a sample contains 12,000 firms, which is roughly the sample size of our empirical study, the dimension of the state variables in the DSW model becomes 24,002, and the estimation task becomes numerically challenging.

We propose a new reduced-form approach based on a forward intensity construction to estimate a firm's default probabilities for different periods ahead. Like the DSW model, our construction takes into account both defaults/bankruptcies and other types of firm exits such as mergers and acquisitions. However, our method which relies on estimating forward default probabilities can produce the term structure of default probabilities without explicitly modelling and estimating the high-dimensional state variable process. Our default prediction over multiple periods solely uses the known data at the time of performing prediction, and thus circumvents the difficult task of specifying and estimating the time dynamics for covariates. Our forward intensity approach can be implemented by maximizing a pseudo-likelihood function constructed with overlapping data to utilize the available data to the fuller extent. Like the DSW model, the pseudo-likelihood function can also be decomposed to default and other exit components, making it less numerically intensive in estimation. Unique to our approach is the nature of the pseudo-likelihood function which makes it also decomposable for different forward periods. This decomposability in effect turns the estimation of the forward default parameters in a later period totally unrelated to the parameters governing the earlier periods. Thus, the forward intensity model's estimation becomes a non-sequential numerical optimization problem, and is naturally amenable to a parallel computing implementation.

In addition to computational efficiency, we conjecture that the forward intensity approach is more robust than DSW model especially for long-horizon prediction.[1] The DSW model is essentially an iterative and indirect method where one needs to generate future random values for the covariates so as to compute the probability of default. If the dynamics of the covariates are mis-specified, generating future values for the covariates multiple steps ahead may cause serious biases. In contrast, the forward intensity approach is a direct projection of current event realizations on past data, and is therefore more likely to be robust to model mis-specification.

Our empirical analysis uses a large sample of the US exchange-listed companies (both industrial and financial) covering more than 12,000 firms and over one million firm-month observations for the period from 1991 to 2010. We examine the effects of several commonly used macroeconomic factors and firm-specific attributes on companies' one-month forward default probabilities from an immediate start to as long as 36 months ahead. We find that a firm's leverage, liquidity, profitability and volatility are four

[1]We thank an anonymous referee for pointing this out.

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 by their 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-lived except for distance-to-default.

Our forward intensity approach actually coincides with the DSW model 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. Statistic tests 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 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%.[2] 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 study portfolio behavior. We are able to show

[2]In the literature, the receiver operating characteristic (ROC) is sometimes used as a performance metric. According to Sobehart and Keenan (2001), there is a simple monotonic relationship between accuracy ratio and ROC; that is, accuracy ratio equals $2(ROC - 0.5)$. A totally uninformative model thus has an ROC equal to 0.5.

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 7%. Interestingly, the peak of the forward default probability curve moved to around the fifth to sixth month, which is very close to its actual bankruptcy filing month.

To compare the forward intensity approach with the DSW model, we conduct additional prediction accuracy analysis. We employ the four covariates used in Duffie, *et al* (2007) for both methods. For the DSW model implementation, we also use the same vector autoregressive model as in that model. Because of the total number of covariates involved is too large to estimate, we reduce the number of firms in our sample by random sampling to 3,000 non-financial firms, a size comparable to the DSW implementation. We restrict the comparison analysis to non-financial firms because Duffie, *et al* (2007), like the great majority of the default/bankruptcy literature, excluded financial firms from their study. Our results show that both methods perform similarly for short-horizon predictions. When we increase the prediction horizon beyond one year, the DSW model outperforms in-sample but seriously underperforms out-of-sample. The inconsistent in-sample and out-of-sample performances of the DSW model is likely due to over-fitting arising from its use of a large number of parameters in modelling the joint dynamics of the covariates.

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 combined 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 forward starting time 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

$$f_{it}(\tau) = \exp\left(\alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \alpha_2(\tau)x_{it,2} + \cdots \alpha_k(\tau)x_{it,k}\right) \tag{5}$$

$$g_{it}(\tau) = f_{it}(\tau) + \exp\left(\beta_0(\tau) + \beta_1(\tau)x_{it,1} + \beta_2(\tau)x_{it,2} + \cdots \beta_k(\tau)x_{it,k}\right) \tag{6}$$

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.

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 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 $t = 0, 1, 2, \cdots$ and $\tau = 0, 1, 2, \cdots$ as time sequences with an increment of one month (and so are τ_{Di} and τ_{Ci}). 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 $[t+\tau, t+\tau+1]$.

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 $[t+\tau, t+\tau+1]$:

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

2. Forward combined exit probability at time t for the period $[t+\tau, t+\tau+1]$:

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

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

$$P_t(t < \tau_{Di} = \tau_{Ci} \leq t+\tau) = \sum_{s=0}^{\tau-1} e^{-\psi_{it}(s)s\Delta t}\left(1 - e^{-f_{it}(s)\Delta t}\right) \tag{9}$$

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

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

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 from 0 to T measured in months. 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 months with each equal to $\Delta t = 1/12$.

We assume 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. We denote the model's parameter set by $\alpha = \{\alpha(0), \alpha(1), \cdots, \alpha(\tau-1)\}$ and $\beta = \{\beta(0), \beta(1), \cdots, \beta(\tau-1)\}$. The pseudo-likelihood function for prediction horizon τ can be expressed as

α (0) 是一串参数

$$\mathscr{L}_\tau(\alpha, \beta; \tau_C, \tau_D, X) = \prod_{i=1}^{N} \prod_{t=0}^{T-1} \mathscr{L}_{\tau,i,t}(\alpha, \beta) \tag{11}$$

where

$$\begin{aligned}\mathscr{L}_{\tau,i,t}(\alpha, \beta) =& 1_{\{t_{0i}\le t, \tau_{Ci}>t+\tau\}} P_t(\tau_{Ci}; \tau_{Ci} > t+\tau) + 1_{\{t_{0i}\le t, \tau_{Di}=\tau_{Ci}\le t+\tau\}} P_t(\tau_{Di}; \tau_{Di} = \tau_{Ci} \le t+\tau) \\ &+ 1_{\{t_{0i}\le t, \tau_{Di}\ne\tau_{Ci}, \tau_{Ci}\le t+\tau\}} P_t(\tau_{Ci}; \tau_{Di} \ne \tau_{Ci} \& \tau_{Ci} \le t+\tau) + 1_{\{t_{0i}>t\}} + 1_{\{\tau_{Ci}\le t\}}\end{aligned}$$

and

$$P_t(\tau_{Ci}; \tau_{Ci} > t+\tau) = \exp\left[-\sum_{s=0}^{\tau-1} g_{it}(s)\Delta t\right]$$

$$P_t(\tau_{Di}; \tau_{Di} = \tau_{Ci} \le t+\tau)$$

$$= \begin{cases} 1 - \exp\left[-f_{it}(0)\Delta t\right], & \text{if } \tau_{Di} = t+1 \\ \exp\left[-\sum_{s=0}^{\tau_{Di}-t-2} g_{it}(s)\Delta t\right] \left\{1 - \exp\left[-f_{it}(\tau_{Di}-t-1)\Delta t\right]\right\}, & \text{if } t+1 < \tau_{Di} \le t+\tau \end{cases}$$

$$P_t(\tau_{Ci}; \tau_{Di} \ne \tau_{Ci} \& \tau_{Ci} \le t+\tau)$$

$$= \begin{cases} \exp[-f_{it}(0)\Delta t] - \exp[-g_{it}(0)\Delta t], & \text{if } \tau_{Ci} = t+1 \\ \exp\left[-\sum_{s=0}^{\tau_{Ci}-t-2} g_{it}(s)\Delta t\right] \times & \\ \{\exp[-f_{it}(\tau_{Ci}-t-1)\Delta t] - \exp[-g_{it}(\tau_{Ci}-t-1)\Delta t]\}, & \text{if } t+1 < \tau_{Ci} \le t+\tau \end{cases}$$

with $\Delta t = 1/12$, $f_{it}(s)$ and $g_{it}(s)$ as defined in equations (5) and (6). The first term on the right-hand side of the pseudo-likelihood function 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 the pseudo-likelihood to 1, which is transformed to 0 in the log-pseudo-likelihood function.

The pseudo-likelihood function $\mathscr{L}_\tau$ can be numerically maximized to obtain estimates $\hat{\alpha}$ and $\hat{\beta}$. Note that when $\tau > 1$, the above pseudo-likelihood is constructed with observations from overlapped periods. As an example of $\tau = 2$ at a particular time point, default over the period that starts one period ahead will be correlated with the similar default event corresponding to the next time point due to an overlapping common period. Because of 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 (11) is the product of separate terms involving α and β, we can maximize its two components separately to obtain $\hat{\alpha}$ and $\hat{\beta}$, which is similar to Proposition 2 in Duffie, *et al* (2007).

Moreover, the pseudo-likelihood function for α (or β) can be further decomposed to separate terms involving $\alpha(\tau)$ (or $\beta(\tau)$) corresponding to different τ'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

$$\mathscr{L}(\alpha(s)) = \prod_{i=1}^{N} \prod_{t=0}^{T-s-1} \mathscr{L}_{i,t}(\alpha(s)), \quad s = 0, 1, \cdots, \tau - 1 \tag{12}$$

$$\mathscr{L}(\beta(s)) = \prod_{i=1}^{N} \prod_{t=0}^{T-s-1} \mathscr{L}_{i,t}(\beta(s)), \quad s = 0, 1, \cdots, \tau - 1 \tag{13}$$

where

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

$$
\begin{aligned}
\mathscr{L}_{i,t}(\beta(s)) =& 1_{\{t_{0i}\le t, \tau_{Ci}>t+s+1\}} \exp\{-[g_{it}(s) - f_{it}(s)]\Delta t\} \\
&+ 1_{\{t_{0i}\le t, \tau_{Di}=\tau_{Ci}=t+s+1\}} \\
&+ 1_{\{t_{0i}\le t, \tau_{Di}\neq\tau_{Ci}, \tau_{Ci}=t+s+1\}} \{1 - \exp\left[-(g_{it}(s) - f_{it}(s))\Delta t\right]\} \\
&+ 1_{\{t_{0i}>t\}} + 1_{\{\tau_{Ci}<t+s+1\}}
\end{aligned}
$$

Note that

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

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 2010. 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.[3] 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 12,225 companies (including financial firms) giving rise to 1,066,337 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 0.3% and 3.2% of the firms in each sample year. Other forms of exit are significantly higher, ranging from 5.1% to 13.7%.

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.

[3] 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.

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: logarithm 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. 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 "level" to reflect the recent level of the variable. We also calculate the difference between its current value and the 12-month moving average, and denote it by the subscript "trend". The "trend" 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 99.5 percentile value. Similarly, all values are subject to the floor at the 0.5 percentile value. The summary statistics and correlation matrix for the firm-specific attributes are reported in Tables 2-3.

The massive US governmental interventions during the 2008-09 financial crisis are likely to have significant impact on the default probabilities of US companies. To factor in the possible bailout effects, we add a common term $\lambda \exp\{-\delta(t - t_B)\}1_{t>t_B}$ to the forward default intensity function where t_B denotes the end of August 2008. This bailout dummy variable is set to zero until September 2008 when the US government starts to bail out AIG. The coefficient λ as well as the decay rate δ can be estimated by maximizing the pseudo-likelihood function. Since our data set has limited number of time periods after September 2008, we only employ such specification for the forward default intensity functions with horizons less than one year. Specifically, we employ the following forward default intensity function in the empirical study:

$$f_{it}(\tau) = \exp\{\lambda(\tau)\exp[-\delta(\tau)(t - t_B)]1_{t>t_B} + \alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \cdots \alpha_k(\tau)x_{it,k}\}$$

where $\tau = 0, 1, 2, \cdots, 11$.

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 0 month to 35 months. To show the impact of various factors/attributes on firms' default probabilities, we plot in Figure 1 the estimated coefficients corresponding to different forward starting times. Also plotted is the 90% confidence interval for each variable used in the forward default intensity function.

The bailout coefficients are negative and significant for all points on the term structure that we have computed (up to $\tau = 11$), which suggests that firms' default probabilities have been lower than they would have been if the government did not intervene post-Lehman Brothers' bankruptcy filing in September 2008. The exponential decay rates are positive as expected, but the estimates are insignificant possibly due to the limited data after September 2008.

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 may simply serve 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 benefit from a bailout by government simply because they may be "too big to fail". Our results show that forward default intensities decrease with size in the short run but 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 most of the prediction horizons, which is consistent with Campbell, *et al* (2008) although the estimations are usually insignificant. 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 how market-to-book asset ratio affects a company's default probability exactly and we 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 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. We repeat 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 bars depict the observed numbers of defaults and the solid line corresponds to the in-sample predicted values where the parameters are estimated using the whole sample. We also employ an expanding window approach to generate the out-of-sample results that are represented by the dashed line. At each month-end starting from January 2001, we re-estimate the model using all the data available up to that time and compute the predicted number of defaults for different prediction horizons.

For shorter horizons, our in-sample predictions fit the subsequent realizations quite well. However, as the horizon increases, the solid line deviates from the bars, implying a deteriorating performance in the longer run. Generally speaking, our in-sample long-run results 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. The out-of-sample predictions are close to the in-sample ones for most periods except for the internet bubble burst period and the 2008-09 financial crisis. These two periods are, however, quite unique from a historical perspective, which makes it harder for the estimated model to anticipate the default behavior out-of-sample.

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 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 in-sample 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 73.44% and 65.69%, respectively.

We further examine the prediction accuracy for two sub-samples. One sub-sample comprises only financial firms (SIC between 6000 and 6999) and the other sub-sample includes all non-financial firms. The accuracy ratios for the non-financial sample are quite close to those of the full sample. The in-sample prediction for financial firms is actually more accurate for horizons no more than 2 years, but less accurate when the prediction horizon is increased to 3 years.

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. Similar to what we have done in the last section, at each month-end starting from January 2001, we re-estimate the model with 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 usually very close to the in-sample results and are even higher for longer prediction horizons.

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 US 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 default probability 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 1.7%. This result suggests that the market did not foresee any 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 0.3% while the 3-years cumulative default probability rose to 2.4%. 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 1.2% and the 3-year cumulative default probability rose significantly to 5.3%. 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 6 months. The 1-year cumulative default probability increased sharply to 6.6% which is about 30 times of the value 3 years earlier. The 3-year cumulative default probability climbed to 12.4%. 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.

5.5 Parameter smoothing

As shown in Figure 1, the shapes of the parameters' term structures seem to exhibit regularity. Simplification can be obtained by applying some function to describe these term structures. The resulting functions may also be used for interpolation/extrapolation. We conduct additional accuracy analysis using the smoothed parameters to determine whether smoothing causes deterioration in performance. We employ the method as in Nelson and Siegel (1987), which was originally developed to model the term structure of interest rates. We apply the method to the term structure of each parameter appearing in the forward intensity function. For each set of parameters, $\theta(\tau)$, $\tau = 0, 1, 2, \cdots$, we assume the parameter value is a function of the forward-starting time according to the following Nelson-Siegel function:

$$\theta(\tau) = \beta_0 + \beta_1 \frac{1 - \exp(-\tau/d)}{\tau/d} + \beta_2 \left\{ \frac{[1 - \exp(-\tau/d)]}{\tau/d} - \exp(-\tau/d) \right\}$$

We apply the parameter estimates obtained earlier and fit the above function by the least square critierion to obtain the parameters: β_0, β_1, β_2, and d. Instead of using the original parameter estimates, we apply the smoothed parameter values in conducting default predictions. The result of the accuracy ratio analysis based on the smoothed parameters are reported in Table 6 under "Full sample (smoothed)". It is clear that the results are very close to those without parameter smoothing. No material deterioration in prediction accuracy indicates that the forward intensity model can be implemented with a much smaller set of parameters; for example, in the case of 36 months and for each covariate, one achieves the reduction from 36 parameters to just four Nelson-Siegel parameters.

5.6 The forward intensity model vs. the DSW model

The forward intensity model has two advantages over the DSW method. First, to generate the default probabilities under the DSW model beyond the immediate period, a high dimensional time series model must be specified. A suitable model that can adequately reflect the joint dynamics of firms' attributes is hard to come by, however. Applying an *ad hoc* model will run the risk of seriously mis-specifying the system. Apart from the difficulty of model specification, estimating a very high dimensional time series model can be numerically challenging. It is obvious that the joint system will likely involve a high-dimensional variance-covariance matrix. Computing its inverse and determinant will be required for estimation, but the task will be complex and numerically expensive due to its high dimension. In contrast, the forward intensity approach completely bypasses the model specification and estimation for the covariates.

There is a basis to conjecture that the forward intensity method will perform better vis-a-vis the DSW model. The DSW model relies on an indirect method of assessing default probability beyond the immediate period. It needs to simulate the joint dynamics of the covariates forward in order to compute the conditional default probabilities for a future period. Averaging is then performed over the conditional default probabilities to obtain the final default probability for the future period of interest. Its application requires an auxiliary system for the covariates, and the auxiliary system can potentially introduces serious model specification error and parameter uncertainty in light of its very high dimension. In comparison, the forward intensity approach forms a direct projection of the current default realizations on the past data. Therefore, this direct method is more likely to be robust to model mis-specification and is also free of parameter uncertainty introduced by the auxiliary system.

To study the performance of the forward intensity model versus the DSW model, we employ the same covariates as in Duffie, *et al* (2007); they are: the trailing one year S&P500 index return, the three month treasury rate, firms' distance-to-default[4], and firms' trailing one year stock return. We also use the same autoregressive model to

[4]The method that we use to compute distance-to-default is different from the DSW model as described previously.

describe the covariates as in the DSW model.[5] Duffie, *et al* (2007) used less than 3,000 firms in their study, and all were non-financial companies. But our sample is considerably larger and contains more than 12,000 companies including financial firms. Estimating the joint autoregressive model for our sample becomes numerically unworkable due to its exceedingly high dimension (i.e., 6,002 for 3000 firms vs. 24,002 for 12,000 firms). Therefore, we randomly select a sub-sample of 3,000 non-financial firms to make the sample size comparable in size and nature to that of Duffie, *et al* (2007). We stick to the four covariates as in the DSW model for two reasons. First, we would have to introduce some *ad hoc* specification for the covariates that are not already in the DSW model. Second, introducing even just one additional firm attribute would significantly increase the dimension of the joint system from 6,002 to 9,002, and for which we would also have difficulty of numerically estimating the joint dynamics.

Table 7 reports the accuracy ratios for both approaches on the randomly selected subsample. Panel A consists of the in-sample results where the parameters for both methods are estimated based on the entire sample period. The performance is close when the prediction horizon is short. The DSW model clearly outperforms the forward intensity approach when the prediction horizon is increased beyond six months. However, such superior in-sample performance of the DSW model is possibly due to the large number of parameters in the model.[6] Panel C of Table 7 reports the out-of-sample accuracy ratios. For the out-of-sample analysis, we follow Duffie, *et al* (2007) to estimate both models only once at the end of January 2001 using all the data available up to that time.[7] The performance of both models is still close for shorter horizons. But the DSW model begin to underperform for longer horizons and the prediction accuracy ratios also deteriorate significantly when comparing with its own in-sample accuracy ratios. Such results suggest an in-sample overfitting of the DSW method.

[5]The autoregressive model in DSW is built upon the three-month treasury rate, ten-year treasury rate, trailing one-year S&P500 index return, distance-to-default and logarithmic asset value. Since the covariates employed in their intensity model is trailing one-year equity return rather than asset return, we estimate the model using the logarithmic equity value rather than the logarithmic asset value. However, we also test the model based on the logarithmic asset value and deducing the needed equity value through the Black-Sholes formula. The main conclusions still hold.

[6]With 3,000 firms, two macroeconomic variables and two firm-specific variables, the number of parameters in the DSW model would exceed 6,000. However, the forward intensity approach only has 120, 240 and 360 parameters for one-year, two-year and three-year predictions, respectively. Moreover, increasing the number of firms in the sample will not affect the number of parameters for the forward intensity approach, but will further increase the number of parameters in the DSW model.

[7]In previous sections, we conduct the out-of-sample analysis by re-estimating the model every month starting from January 2001. Such frequent re-estimation strategy is, however, intractable for the DSW model because estimating the joint autoregressive model for the covariates is too computationally demanding.

The in-sample and out-of-sample comparison was made on different time periods. To control for the difference, we report a different set of in-sample performance results in Panel B of Table 7 where the parameters were obtained using the whole sample but the accuracy ratios are computed using the period 2001-2010 so as to coincide with the period for our out-of-sample analysis. Comparing Panels B and C, the results show that accuracy ratios produced by the forward intensity approach are highly consistent in- and out-of-sample, but the DSW model gives rise to outcomes that are highly volatile.

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 also find that the forward intensity approach generate more robust default predictions as compared to the model in Duffie, *et al* (2007).

The literature on default/bankruptcy predictions often avoids financial firms. A typical sample selection criterion adopted in this literature is to exclude financial firms. Given the importance of financial sector to the economy, such exclusion is clearly undesirable. We have shown that financial firms can be successfully lumped together with non-financial firms by applying Duan's (2010) way of estimating distance-to-default. In fact, we show that the resulting default prediction model common for financial and non-financial firms can lead to comparable performance for financial as well as non-financial sectors.

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 also offers 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 (11). The parameter set is denoted by θ and its true value is θ_0. $\log \mathscr{L}_N(y,\theta)$ is the log-pseudo-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*

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

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 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 $E\left[g_i(y,\theta)\right] = 0$. Let $\{y_i\}$ be a sequence of independent and not necessarily identically distributed random vectors such that $E\left[\sup_{\theta\in\Theta} |g_i(y_i,\theta)|^{1+\delta}\right] \leq M < \infty$ 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 $g_i(y,\theta) = \sum_{t=0}^{T-1} \log \mathscr{L}_{\tau,i,t}(\theta) - E_{\theta_0}\left[\sum_{t=0}^{T-1} \log \mathscr{L}_{\tau,i,t}(\theta)\right]$. $g_i(y,\theta)$ in a compact neighborhood of θ_0 satisfies all the conditions in Lemma 2 because of the assumptions. Therefore,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=0}^{T-1} \log \mathscr{L}_{\tau,i,t}(\theta) \rightarrow l(\theta) = \lim_{N\rightarrow\infty} N^{-1}\sum_{i=1}^{N} E_{\theta_0}\left[\sum_{t=0}^{T-1} \log \mathscr{L}_{\tau,i,t}(\theta)\right]$$

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

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

$$\frac{\partial \log \mathscr{L}_N(\hat{\theta})}{\partial \theta} = \frac{\partial \log \mathscr{L}_N(\theta_0)}{\partial \theta} + \frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial \theta \partial \theta'}(\hat{\theta} - \theta_0), \text{ where } \tilde{\theta} \text{ lies between } \theta_0 \text{ and } \hat{\theta}$$

$$\Rightarrow \hat{\theta} - \theta_0 = -\left(\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial \theta \partial \theta'}\right)^{-1}\left(\frac{1}{N}\frac{\partial \log \mathscr{L}_N(\theta_0)}{\partial \theta}\right)$$

Consider

$$\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial \theta \partial \theta'} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T-1}\frac{\partial^2 \log \mathscr{L}_{\tau,i,t}(\tilde{\theta})}{\partial \theta \partial \theta'} \rightarrow_p H(\tilde{\theta}) \;\; \textit{as } N \rightarrow \infty,$$

where $H(\theta) = \lim_{N\rightarrow\infty} \frac{1}{N}\sum_{i=1}^{N} E_{\theta_0}\sum_{t=0}^{T-1}\frac{\partial^2 \log \mathscr{L}_{\tau,i,t}(\theta)}{\partial \theta \partial \theta'}$. 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

$$\frac{1}{N}\frac{\partial^2 \log \mathscr{L}_N(\tilde{\theta})}{\partial \theta \partial \theta'} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T-1}\frac{\partial^2 \log \mathscr{L}_{\tau,i,t}(\tilde{\theta})}{\partial \theta \partial \theta'} \rightarrow_p H(\theta_0) \;\; \textit{as } N \rightarrow \infty.$$

Therefore,

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

where $\left\{ \sum_{t=0}^{T-1} \frac{\partial \log \mathscr{L}_{\tau,i,t}(\theta_0)}{\partial \theta}, i = 1, 2, \cdots, N \right\}$ 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

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

The asymptotic variance can thus be approximated by

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

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

$$E_t = V_t N(d_t) - e^{-r(T-t)} L N(d_t - \sigma\sqrt{T-t}) \tag{14}$$

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

$$d_t = \frac{\ln(V_t/L) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}. \tag{15}$$

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 the unknown fraction 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

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

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 (14), $\hat{d}_t$ is computed using equation (15) 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] Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449-470.

[25] Nelson, C.R. and A.F. Siegel, 1987, Parsimonious modeling of yield curves. *Journal of Business* 60, 473-489.

[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] Sobehart, J. and S. Keenan, 2001, Measuring default accurately, *RISK* 14, 31-33.

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

[31] 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. Bailout is the exponential decaying term for the bailout effect, 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 "level" denotes the average in the preceding 12 months, "trend" 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.

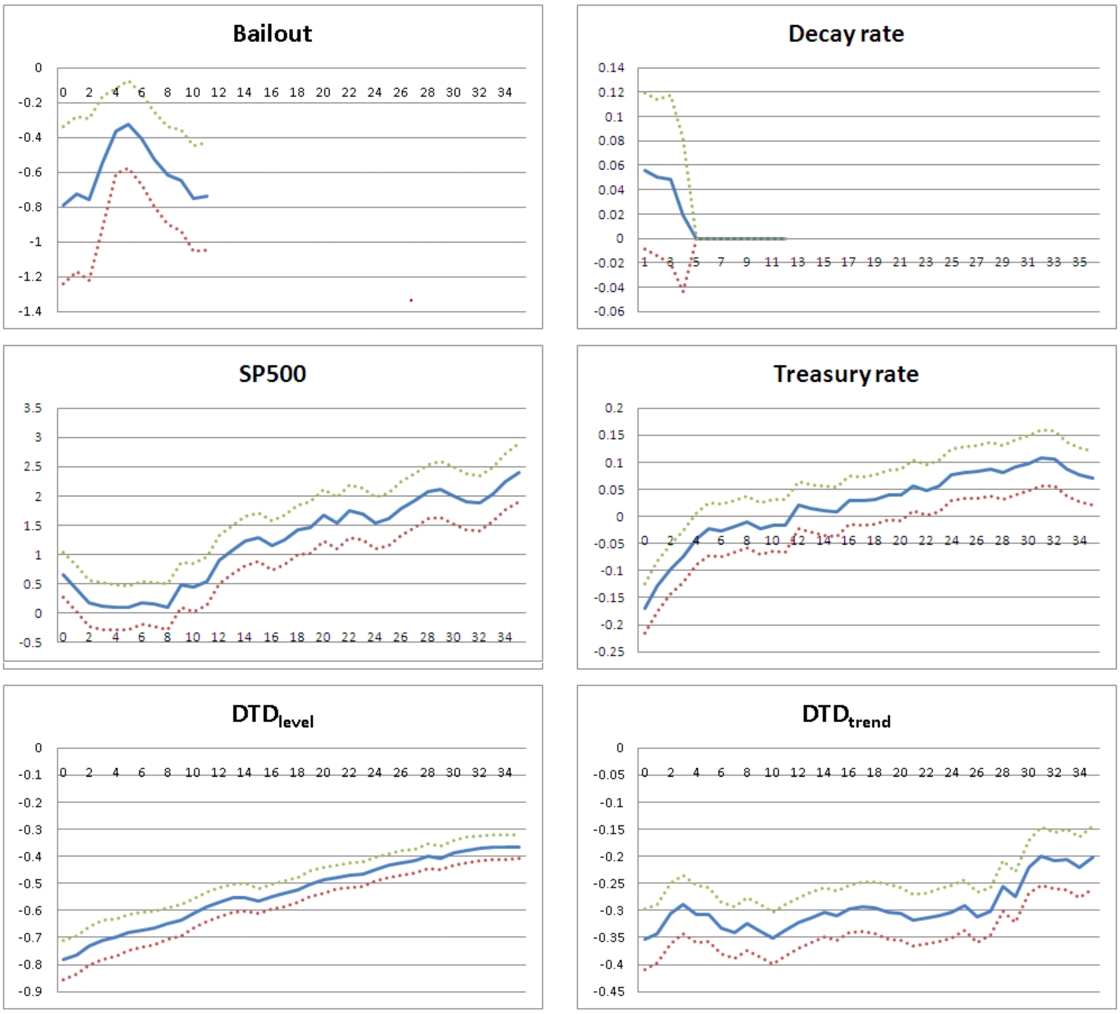

Figure 1. Parameter estimates for the forward default intensity function (Cont'd)

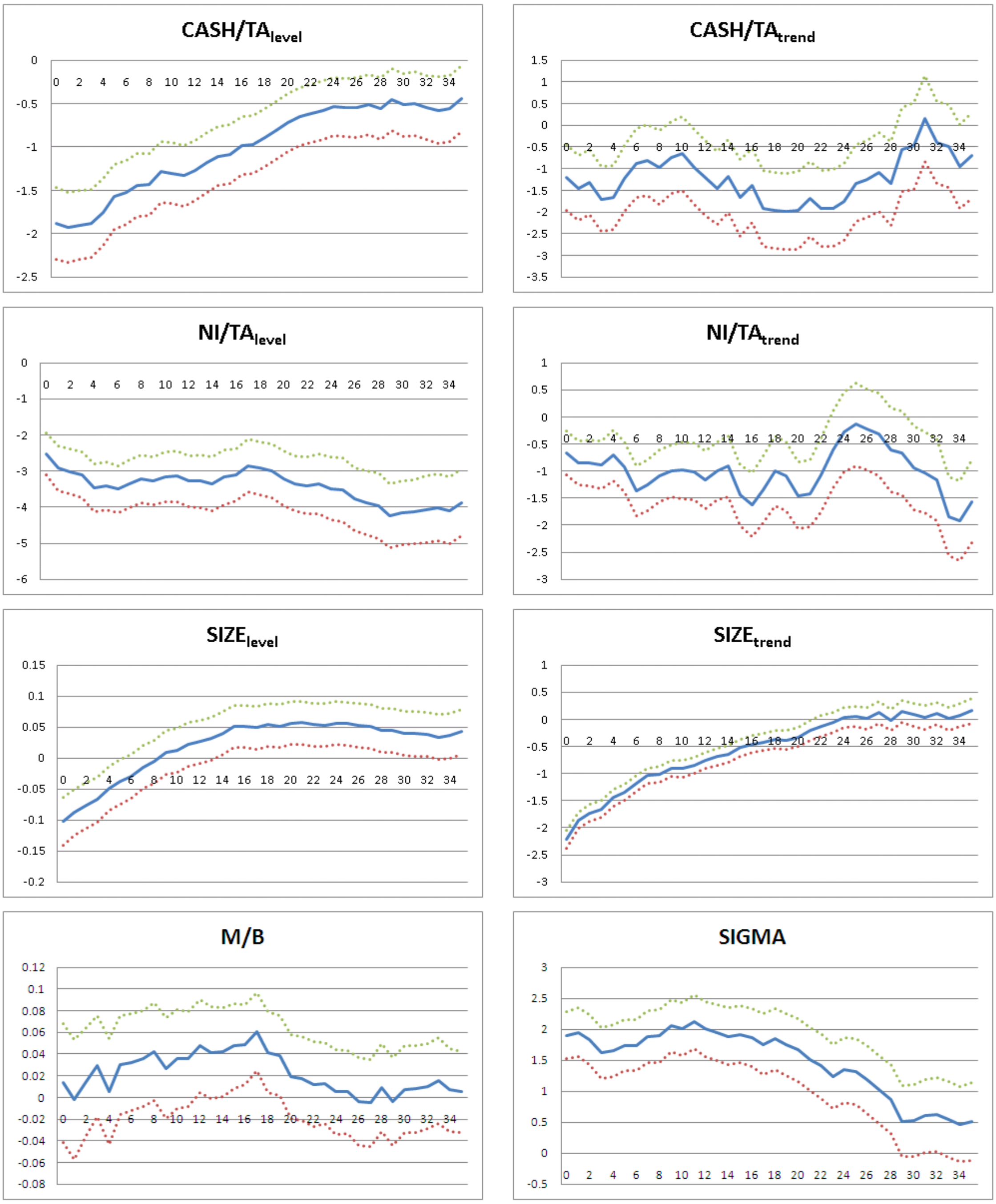

Figure 2. Aggregate number of defaults

This figure shows the observed (bars), in-sample predicted (solid line) and out-of-sample predicted (dashed 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. The in-sample results are generated using the parameters estimated over the whole sample. At each month-end starting from January 2001, we re-estimate the model using all the data available up to that time and compute the out-of-sample predicted number of defaults for different prediction horizons.

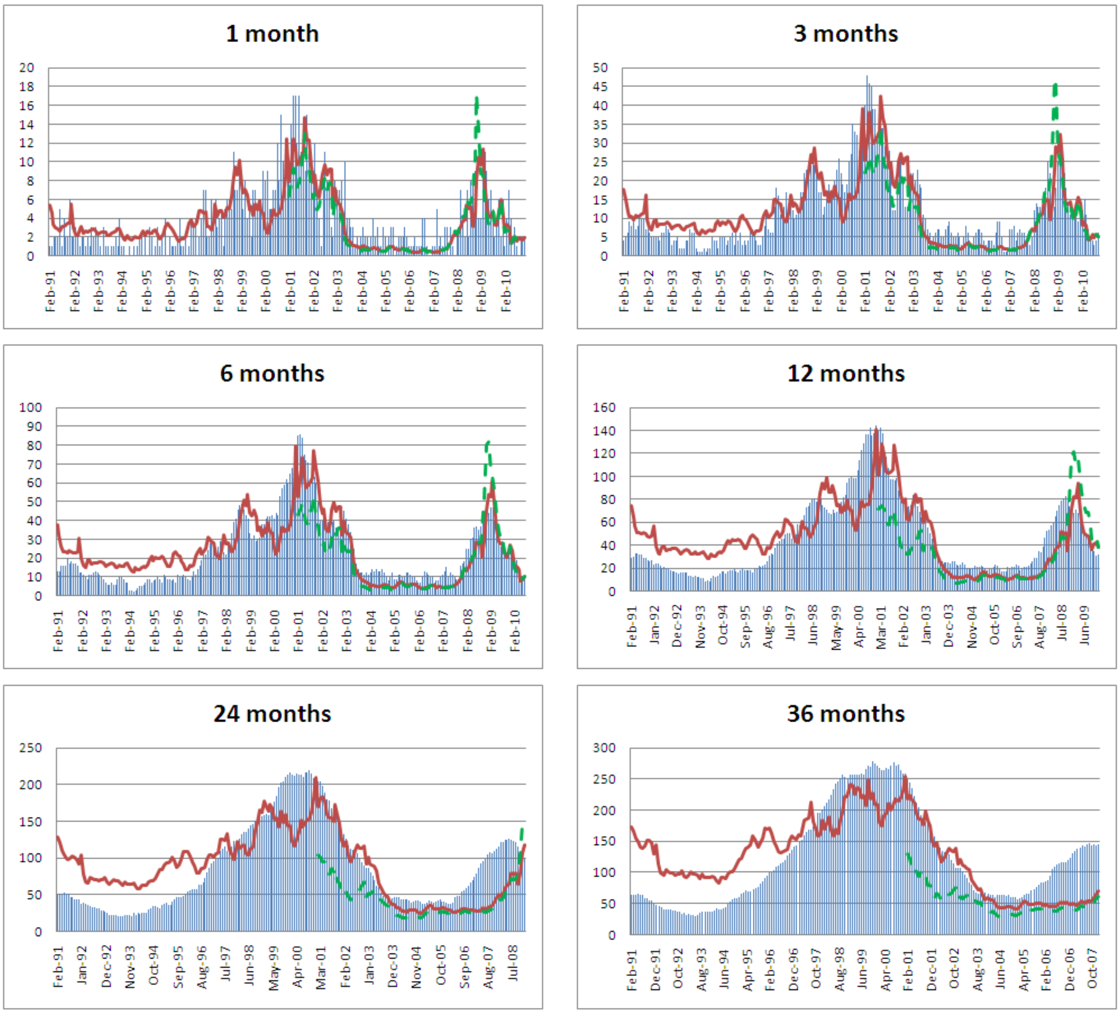

Figure 3A. In-sample cumulative accuracy profiles

This figure shows the in-sample cumulative accuracy profiles (power curves) based on all firms and the entire sample period (1991 to 2010) for different prediction horizons.

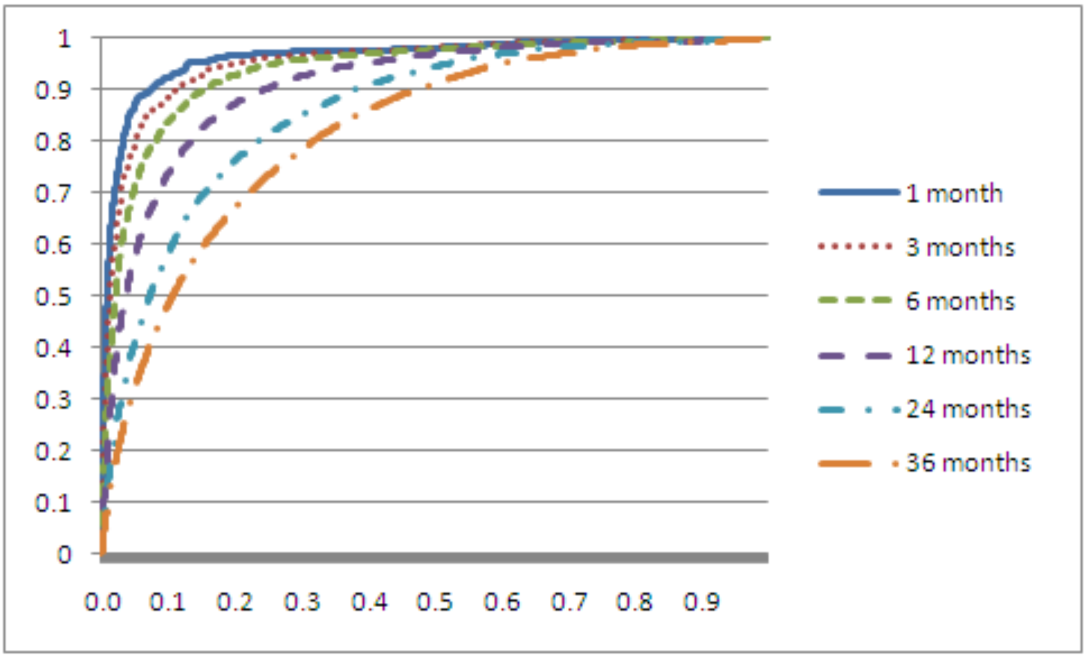

Figure 3B. Out-of-sample (cross-section) cumulative accuracy profiles

This figure shows the out-of-sample cumulative accuracy profiles (power curves) over the entire sample period (1991-2010) 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.

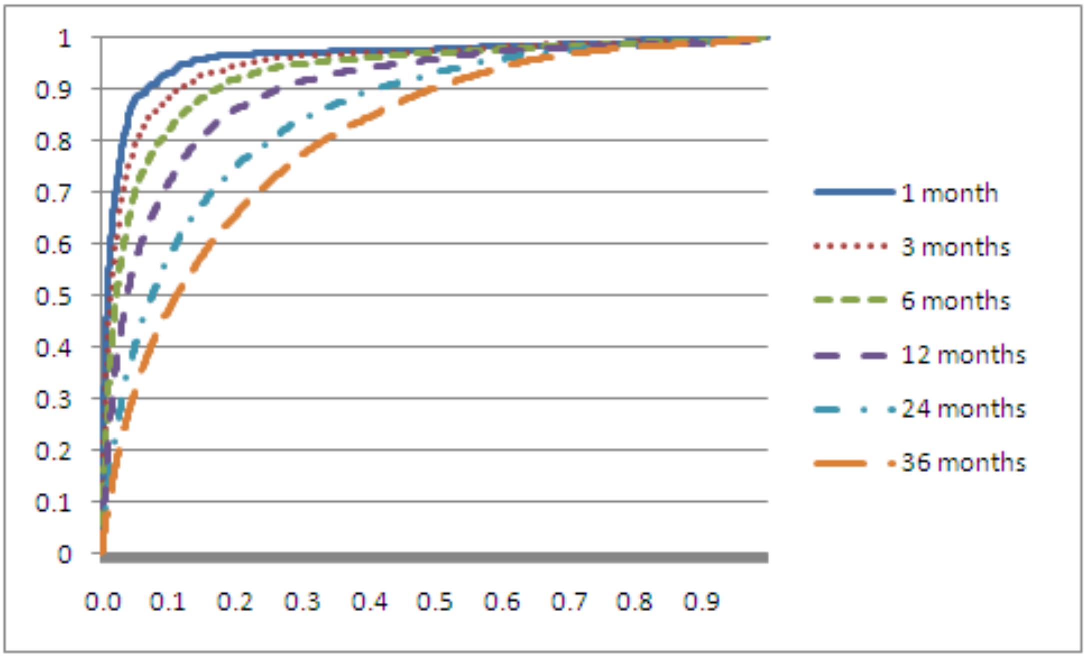

Figure 3C. Out-of-sample (over time) cumulative accuracy profiles

This figure shows the out-of-sample cumulative accuracy profiles (power curves) for the sample period (2001-2010) 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 that time for estimation.

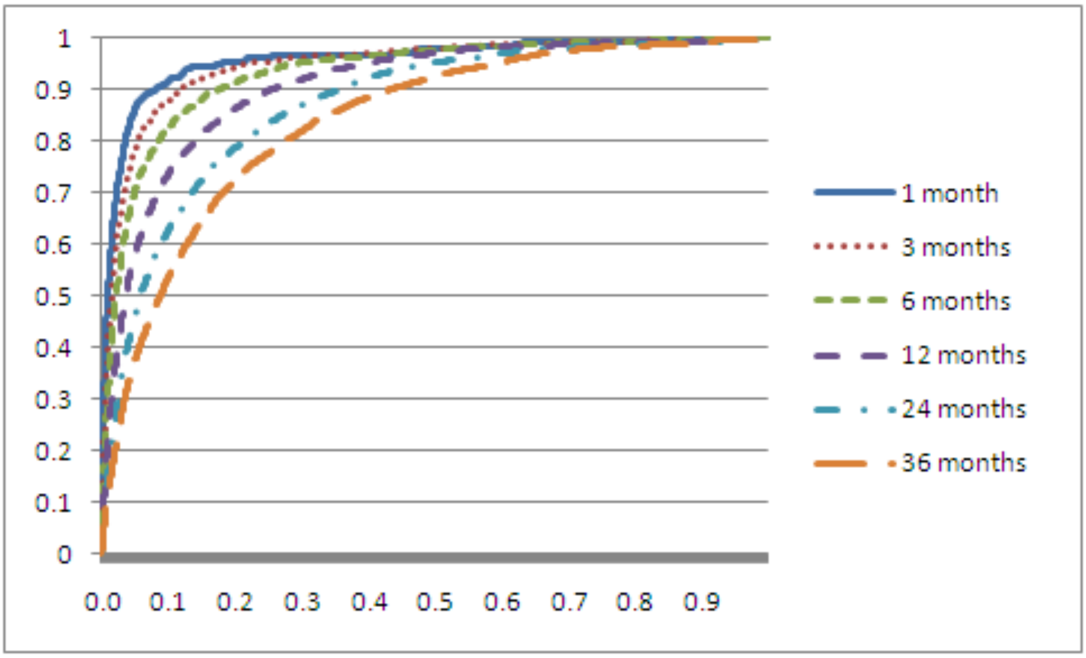

Figure 4. Lehman Brothers' term structure of forward and cumulative default probabilities

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).

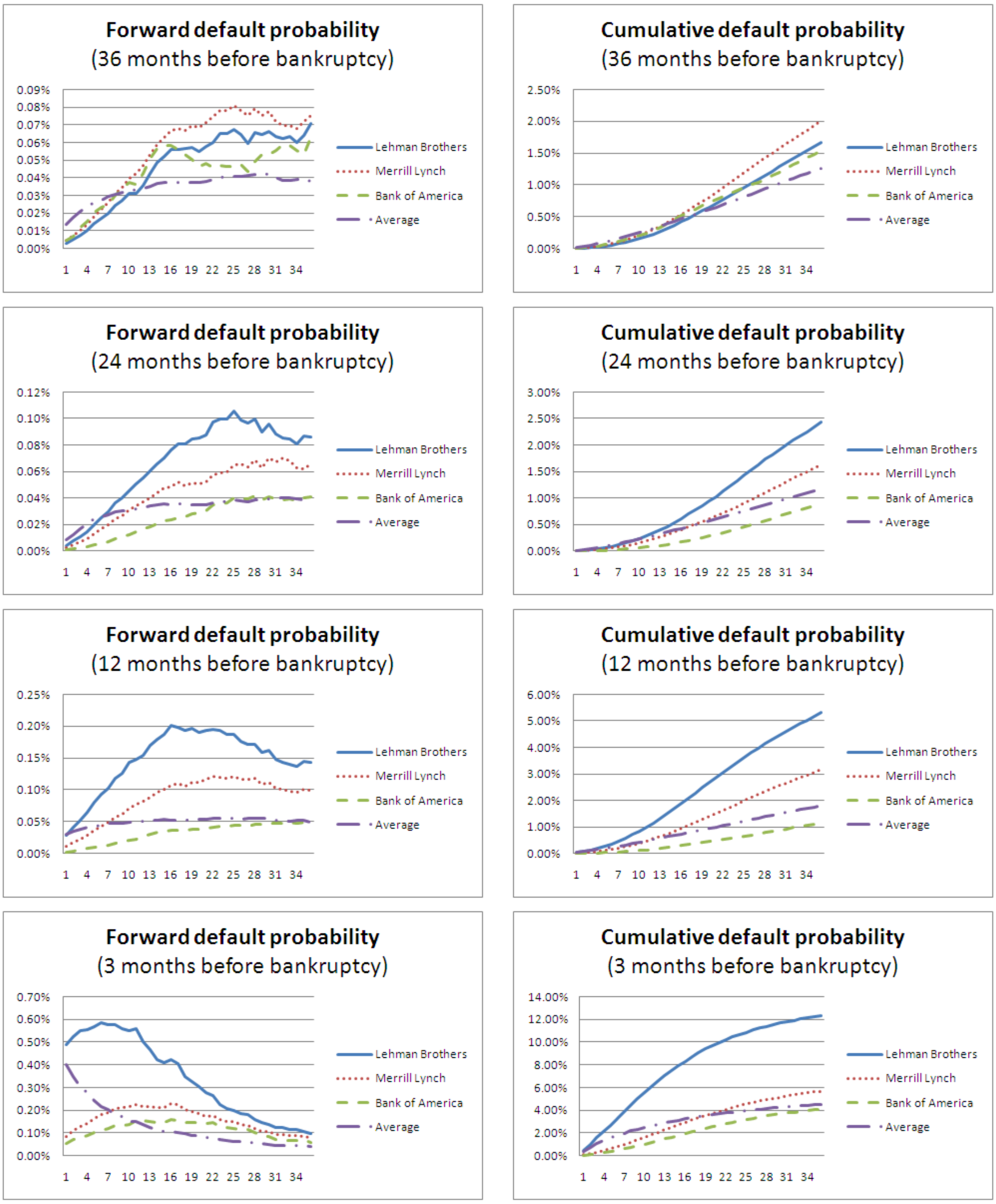

Table 1. Number of defaults and other exits

Total number of active firms, defaults/bankruptcies and other exits for each year over the sample period 1991-2010. 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/Bankruptcies	(%)	Other Exit	(%)
1991	4018	28	0.70%	258	6.42%
1992	4016	23	0.57%	333	8.29%
1993	4194	16	0.38%	213	5.08%
1994	4428	13	0.29%	281	6.35%
1995	5064	18	0.36%	396	7.82%
1996	5453	16	0.29%	472	8.66%
1997	5641	46	0.82%	553	9.80%
1998	5705	66	1.16%	749	13.13%
1999	5423	73	1.35%	745	13.74%
2000	5086	102	2.01%	616	12.11%
2001	4912	156	3.18%	583	11.87%
2002	4675	80	1.71%	403	8.62%
2003	4336	62	1.43%	370	8.53%
2004	4073	24	0.59%	306	7.51%
2005	3920	23	0.59%	292	7.45%
2006	3852	19	0.49%	283	7.35%
2007	3764	21	0.56%	353	9.38%
2008	3672	57	1.55%	290	7.90%
2009	3581	70	1.95%	246	6.87%
2010	3385	32	0.95%	228	6.74%

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 "level" denotes the average in the previous 12 months, "trend" denotes the difference between current value and its previous 12-month average.

	Mean	Std	Min	25%	Median	75%	Max
DTD_{level}	3.630	2.875	-1.364	1.621	3.139	5.056	16.582
DTD_{trend}	0.021	1.796	-7.298	-0.962	0.040	1.045	6.255
$CASH/TA_{level}$	0.170	0.212	0.000	0.026	0.073	0.235	0.931
$CASH/TA_{trend}$	-0.004	0.063	-0.291	-0.019	-0.002	0.014	0.271
NI/TA_{level}	-0.010	0.062	-0.411	-0.008	0.005	0.016	0.080
NI/TA_{trend}	-0.001	0.054	-0.368	-0.007	0.000	0.007	0.268
$SIZE_{level}$	-4.359	2.033	-8.654	-5.854	-4.509	-3.030	1.437
$SIZE_{trend}$	-0.035	0.362	-1.473	-0.200	-0.021	0.150	1.180
M/B	1.953	1.921	0.455	1.029	1.301	2.064	15.561
SIGMA	0.142	0.104	0.024	0.073	0.114	0.177	0.718

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 "level" denotes the average in the previous 12 months, "trend" denotes the difference between current value and its previous 12-month average.

	DTD_{level}	DTD_{trend}	$CASH/TA_{level}$	$CASH/TA_{trend}$	NI/TA_{level}	NI/TA_{trend}	$SIZE_{level}$	$SIZE_{trend}$	M/B	SIGMA
DTD_{level}	1.000	-0.197	0.037	0.032	0.256	0.025	0.474	0.140	0.210	-0.421
DTD_{trend}	-0.197	1.000	0.010	0.029	-0.013	0.046	-0.027	0.432	0.093	0.012
$CASH/TA_{level}$	0.037	0.010	1.000	-0.151	-0.315	-0.024	-0.084	-0.030	0.370	0.220
$CASH/TA_{trend}$	0.032	0.029	-0.151	1.000	0.056	0.098	0.029	0.107	-0.026	-0.024
NI/TA_{level}	0.256	-0.013	-0.315	0.056	1.000	-0.133	0.290	0.093	-0.246	-0.421
NI/TA_{trend}	0.025	0.046	-0.024	0.098	-0.133	1.000	0.012	0.143	0.002	-0.001
$SIZE_{level}$	0.474	-0.027	-0.084	0.029	0.290	0.012	1.000	0.069	0.120	-0.409
$SIZE_{trend}$	0.140	0.432	-0.030	0.107	0.093	0.143	0.069	1.000	0.303	0.012
M/B	0.210	0.093	0.370	-0.026	-0.246	0.002	0.120	0.303	1.000	0.197
SIGMA	-0.421	0.012	0.220	-0.024	-0.421	-0.001	-0.409	0.012	0.197	1.000

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

The maximum pseudo-likelihood estimates of $\alpha(\tau)$ for different prediction horizons. Bailout is the term to reflect the government intervention in September 2008 which is subject to an exponential decay rate, 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 the logarithm 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 "level" denotes the average in the previous 12 months, "trend" 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)												
	$\alpha(0)$	$\alpha(1)$	$\alpha(2)$	$\alpha(3)$	$\alpha(4)$	$\alpha(5)$	$\alpha(6)$	$\alpha(7)$	$\alpha(8)$	$\alpha(9)$	$\alpha(10)$	$\alpha(11)$
Bailout	-0.790***	-0.724***	-0.758***	-0.542**	-0.362**	-0.324**	-0.410***	-0.525***	-0.617***	-0.650***	-0.751***	-0.736***
	(0.274)	(0.270)	(0.281)	(0.230)	(0.148)	(0.153)	(0.159)	(0.165)	(0.170)	(0.177)	(0.184)	(0.188)
Decay rate	0.056	0.050	0.048	0.019	8.0×10^{-5}	$1.9.0 \times 10^{-5}$	3.6×10^{-6}	1.3×10^{-4}	1.6×10^{-4}	1.3×10^{-4}	1.2×10^{-4}	1.1×10^{-4}
	(0.039)	(0.039)	(0.042)	(0.038)	(3.3×10^{-4})	(6.5×10^{-5})	(5.7×10^{-5})	(1.4×10^{-4})	(3.5×10^{-4})	(1.7×10^{-4})	(1.4×10^{-4})	(8.1×10^{-5})
Intercept	-4.918***	-4.567***	-4.390***	-4.269***	-4.138***	-4.140***	-3.999***	-3.880***	-3.829***	-3.714***	-3.719***	-3.687***
	(0.205)	(0.206)	(0.209)	(0.210)	(0.207)	(0.208)	(0.209)	(0.204)	(0.198)	(0.202)	(0.200)	(0.197)
SP500	0.659***	0.411*	0.170	0.118	0.096	0.096	0.178	0.151	0.108	0.476**	0.438*	0.548**
	(0.232)	(0.239)	(0.241)	(0.244)	(0.230)	(0.226)	(0.224)	(0.229)	(0.235)	(0.234)	(0.248)	(0.254)
Treasury rate	-0.169***	-0.129***	-0.097***	-0.074**	-0.042	-0.023	-0.026	-0.019	-0.010	-0.023	-0.016	-0.017
	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)
DTD_{level}	-0.783***	-0.765***	-0.730***	-0.710***	-0.700***	-0.682***	-0.671***	-0.665***	-0.648***	-0.637***	-0.610***	-0.585***
	(0.043)	(0.042)	(0.043)	(0.044)	(0.041)	(0.040)	(0.040)	(0.038)	(0.035)	(0.035)	(0.033)	(0.033)
DTD_{trend}	-0.353***	-0.343***	-0.305***	-0.288***	-0.307***	-0.308***	-0.332***	-0.341***	-0.325***	-0.338***	-0.352***	-0.336***
	(0.034)	(0.033)	(0.034)	(0.033)	(0.032)	(0.030)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
$CASH/TA_{level}$	-1.883***	-1.928***	-1.899***	-1.885***	-1.754***	-1.576***	-1.525***	-1.443***	-1.434***	-1.285***	-1.304***	-1.333***
	(0.251)	(0.248)	(0.241)	(0.240)	(0.234)	(0.226)	(0.222)	(0.221)	(0.220)	(0.214)	(0.213)	(0.215)
$CASH/TA_{trend}$	-1.203**	-1.448***	-1.314***	-1.699***	-1.654***	-1.226***	-0.876*	-0.802	-0.974*	-0.742	-0.655	-0.970*
	(0.468)	(0.455)	(0.452)	(0.455)	(0.449)	(0.465)	(0.484)	(0.492)	(0.514)	(0.509)	(0.521)	(0.527)
NI/TA_{level}	-2.518***	-2.920***	-3.012***	-3.109***	-3.465***	-3.411***	-3.498***	-3.357***	-3.218***	-3.280***	-3.162***	-3.141***
	(0.353)	(0.378)	(0.372)	(0.383)	(0.402)	(0.402)	(0.397)	(0.398)	(0.402)	(0.406)	(0.418)	(0.425)
NI/TA_{trend}	-0.670***	-0.850***	-0.855***	-0.883***	-0.711**	-0.927***	-1.361***	-1.264***	-1.089***	-0.997***	-0.988***	-1.007***
	(0.244)	(0.245)	(0.258)	(0.268)	(0.280)	(0.277)	(0.279)	(0.288)	(0.296)	(0.296)	(0.314)	(0.322)
$SIZE_{level}$	-0.102***	-0.087***	-0.077***	-0.067***	-0.049**	-0.038*	-0.030	-0.014	-0.006	0.008	0.013	0.022
	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.022)	(0.021)	(0.021)
$SIZE_{trend}$	-2.212***	-1.862***	-1.733***	-1.652***	-1.444***	-1.355***	-1.185***	-1.043***	-1.009***	-0.901***	-0.911***	-0.849***
	(0.098)	(0.093)	(0.095)	(0.094)	(0.094)	(0.090)	(0.090)	(0.089)	(0.090)	(0.090)	(0.093)	(0.091)
M/B	0.013	-0.002	0.014	0.029	0.005	0.030	0.032	0.036	0.042	0.027	0.036	0.035
	(0.033)	(0.034)	(0.030)	(0.028)	(0.030)	(0.028)	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.027)
SIGMA	1.901***	1.953***	1.837***	1.617***	1.648***	1.742***	1.744***	1.876***	1.896***	2.059***	2.007***	2.121***
	(0.232)	(0.236)	(0.245)	(0.251)	(0.254)	(0.253)	(0.254)	(0.257)	(0.260)	(0.254)	(0.259)	(0.264)

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

	$\alpha(12)$	$\alpha(13)$	$\alpha(14)$	$\alpha(15)$	$\alpha(16)$	$\alpha(17)$	$\alpha(18)$	$\alpha(19)$	$\alpha(20)$	$\alpha(21)$	$\alpha(22)$	$\alpha(23)$
Bailout												
Decay rate												
Intercept	-3.869***	-3.800***	-3.725***	-3.627***	-3.692***	-3.709***	-3.712***	-3.789***	-3.767***	-3.793***	-3.776***	-3.765***
	(0.172)	(0.176)	(0.177)	(0.173)	(0.172)	(0.173)	(0.175)	(0.180)	(0.182)	(0.181)	(0.184)	(0.182)
SP500	0.912***	1.086***	1.227***	1.297***	1.153***	1.248***	1.428***	1.464***	1.674***	1.533***	1.742***	1.685***
	(0.250)	(0.253)	(0.254)	(0.252)	(0.256)	(0.256)	(0.258)	(0.265)	(0.272)	(0.270)	(0.270)	(0.269)
Treasury rate	0.021	0.015	0.011	0.009	0.030	0.029	0.032	0.040	0.040	0.057**	0.049*	0.056*
	(0.026)	(0.027)	(0.028)	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)
$\mathrm{DTD}_{\mathrm{level}}$	-0.570***	-0.554***	-0.552***	-0.565***	-0.549***	-0.538***	-0.522***	-0.501***	-0.489***	-0.477***	-0.469***	-0.466***
	(0.032)	(0.032)	(0.031)	(0.029)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.027)	(0.028)	(0.027)
$\mathrm{DTD}_{\mathrm{trend}}$	-0.323***	-0.313***	-0.302***	-0.309***	-0.297***	-0.293***	-0.295***	-0.303***	-0.306***	-0.319***	-0.315***	-0.309***
	(0.028)	(0.029)	(0.028)	(0.028)	(0.027)	(0.028)	(0.029)	(0.031)	(0.030)	(0.029)	(0.029)	(0.029)
$\mathrm{CASH/TA}_{\mathrm{level}}$	-1.273***	-1.178***	-1.105***	-1.084***	-0.986***	-0.970***	-0.892***	-0.813***	-0.720***	-0.655***	-0.617***	-0.576***
	(0.211)	(0.208)	(0.204)	(0.207)	(0.201)	(0.205)	(0.204)	(0.204)	(0.203)	(0.203)	(0.204)	(0.205)
$\mathrm{CASH/TA}_{\mathrm{trend}}$	-1.209**	-1.449***	-1.174**	-1.675***	-1.381***	-1.919***	-1.954***	-1.984***	-1.960***	-1.699***	-1.920***	-1.907***
	(0.527)	(0.514)	(0.512)	(0.541)	(0.520)	(0.535)	(0.532)	(0.531)	(0.549)	(0.530)	(0.530)	(0.533)
$\mathrm{NI/TA}_{\mathrm{level}}$	-3.280***	-3.277***	-3.354***	-3.175***	-3.095***	-2.851***	-2.916***	-2.995***	-3.207***	-3.352***	-3.404***	-3.357***
	(0.433)	(0.446)	(0.451)	(0.456)	(0.436)	(0.442)	(0.444)	(0.451)	(0.456)	(0.462)	(0.480)	(0.501)
$\mathrm{NI/TA}_{\mathrm{trend}}$	-1.167***	-0.997***	-0.914***	-1.445***	-1.627***	-1.339***	-0.999**	-1.093***	-1.450***	-1.413***	-1.068**	-0.604
	(0.326)	(0.324)	(0.340)	(0.344)	(0.358)	(0.379)	(0.395)	(0.385)	(0.372)	(0.379)	(0.416)	(0.433)
$\mathrm{SIZE}_{\mathrm{level}}$	0.026	0.031	0.040*	0.051**	0.051**	0.049**	0.054**	0.052**	0.056***	0.057***	0.054**	0.053**
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$\mathrm{SIZE}_{\mathrm{trend}}$	-0.761***	-0.691***	-0.644***	-0.521***	-0.456***	-0.418***	-0.373***	-0.385***	-0.327***	-0.208*	-0.122	-0.051
	(0.091)	(0.095)	(0.095)	(0.096)	(0.096)	(0.098)	(0.102)	(0.106)	(0.107)	(0.109)	(0.113)	(0.113)
M/B	0.047*	0.042	0.042*	0.047**	0.049**	0.060***	0.042*	0.039*	0.019	0.017	0.012	0.013
	(0.026)	(0.026)	(0.025)	(0.024)	(0.022)	(0.022)	(0.022)	(0.023)	(0.024)	(0.024)	(0.024)	(0.023)
SIGMA	2.006***	1.945***	1.885***	1.920***	1.864***	1.756***	1.846***	1.751***	1.666***	1.518***	1.408***	1.237***
	(0.271)	(0.273)	(0.281)	(0.282)	(0.288)	(0.297)	(0.300)	(0.306)	(0.308)	(0.311)	(0.317)	(0.318)

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

	$\alpha(24)$	$\alpha(25)$	$\alpha(26)$	$\alpha(27)$	$\alpha(28)$	$\alpha(29)$	$\alpha(30)$	$\alpha(31)$	$\alpha(32)$	$\alpha(33)$	$\alpha(34)$	$\alpha(35)$
Bailout												
Decay rate												
Intercept	-3.846***	-3.886***	-3.930***	-3.939***	-4.003***	-3.959***	-4.051***	-4.111***	-4.111***	-4.085***	-4.025***	-3.999***
	(0.183)	(0.185)	(0.191)	(0.192)	(0.197)	(0.197)	(0.200)	(0.203)	(0.203)	(0.205)	(0.206)	(0.208)
SP500	1.542***	1.608***	1.784***	1.919***	2.067***	2.116***	2.005***	1.903***	1.876***	2.037***	2.240***	2.400***
	(0.270)	(0.273)	(0.280)	(0.283)	(0.281)	(0.291)	(0.300)	(0.294)	(0.288)	(0.284)	(0.292)	(0.307)
Treasury rate	0.077***	0.081***	0.083***	0.088***	0.081***	0.091***	0.098***	0.108***	0.107***	0.088***	0.077**	0.071**
	(0.029)	(0.029)	(0.030)	(0.030)	(0.030)	(0.031)	(0.031)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)
DTD_{level}	-0.447***	-0.434***	-0.425***	-0.418***	-0.399***	-0.406***	-0.387***	-0.377***	-0.372***	-0.366***	-0.366***	-0.365***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)	(0.027)	(0.027)
DTD_{trend}	-0.302***	-0.290***	-0.312***	-0.301***	-0.255***	-0.275***	-0.220***	-0.200***	-0.208***	-0.206***	-0.220***	-0.203***
	(0.029)	(0.028)	(0.028)	(0.027)	(0.029)	(0.028)	(0.030)	(0.033)	(0.032)	(0.033)	(0.034)	(0.035)
$CASH/TA_{level}$	-0.535***	-0.550***	-0.543***	-0.508**	-0.557***	-0.452**	-0.515**	-0.499**	-0.546**	-0.577**	-0.559**	-0.446*
	(0.204)	(0.204)	(0.209)	(0.209)	(0.214)	(0.217)	(0.221)	(0.222)	(0.226)	(0.233)	(0.234)	(0.231)
$CASH/TA_{trend}$	-1.765***	-1.331**	-1.242**	-1.078*	-1.345**	-0.553	-0.475	0.155	-0.400	-0.479	-0.948	-0.697
	(0.536)	(0.545)	(0.541)	(0.553)	(0.586)	(0.588)	(0.608)	(0.595)	(0.571)	(0.586)	(0.587)	(0.603)
NI/TA_{level}	-3.480***	-3.520***	-3.778***	-3.891***	-3.963***	-4.231***	-4.163***	-4.137***	-4.061***	-4.007***	-4.090***	-3.873***
	(0.523)	(0.534)	(0.529)	(0.537)	(0.549)	(0.541)	(0.539)	(0.538)	(0.566)	(0.561)	(0.563)	(0.564)
NI/TA_{trend}	-0.283	-0.136	-0.232	-0.324	-0.607	-0.667	-0.946**	-1.027**	-1.163**	-1.839***	-1.917***	-1.559***
	(0.446)	(0.464)	(0.450)	(0.464)	(0.471)	(0.469)	(0.468)	(0.456)	(0.464)	(0.440)	(0.444)	(0.469)
$SIZE_{level}$	0.056***	0.056***	0.052**	0.051**	0.044**	0.045**	0.039*	0.039*	0.038*	0.034	0.036	0.042*
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$SIZE_{trend}$	0.036	0.057	0.016	0.124	-0.018	0.144	0.086	0.042	0.108	0.012	0.075	0.157
	(0.116)	(0.115)	(0.124)	(0.121)	(0.118)	(0.123)	(0.129)	(0.135)	(0.129)	(0.129)	(0.131)	(0.136)
M/B	0.005	0.005	-0.004	-0.005	0.009	-0.004	0.007	0.008	0.010	0.015	0.007	0.005
	(0.024)	(0.023)	(0.025)	(0.024)	(0.025)	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)
SIGMA	1.343***	1.316***	1.190***	1.031***	0.869**	0.518	0.528	0.603*	0.626*	0.543	0.470	0.509
	(0.316)	(0.327)	(0.331)	(0.335)	(0.340)	(0.349)	(0.355)	(0.356)	(0.362)	(0.372)	(0.371)	(0.383)

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 "level" denotes the average in the previous 12 months, "trend" 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)												
	$\beta(0)$	$\beta(1)$	$\beta(2)$	$\beta(3)$	$\beta(4)$	$\beta(5)$	$\beta(6)$	$\beta(7)$	$\beta(8)$	$\beta(9)$	$\beta(10)$	$\beta(11)$
Intercept	-4.178***	-4.108***	-4.029***	-3.971***	-3.901***	-3.837***	-3.779***	-3.762***	-3.733***	-3.688***	-3.654***	-3.621***
	(0.053)	(0.054)	(0.055)	(0.056)	(0.057)	(0.058)	(0.059)	(0.060)	(0.061)	(0.062)	(0.063)	(0.064)
SP500	0.033	0.112	0.233***	0.237***	0.297***	0.412***	0.484***	0.545***	0.582***	0.670***	0.626***	0.686***
	(0.072)	(0.072)	(0.072)	(0.073)	(0.075)	(0.075)	(0.076)	(0.077)	(0.078)	(0.080)	(0.084)	(0.085)
Treasury rate	0.055***	0.054***	0.051***	0.053***	0.049***	0.047***	0.046***	0.046***	0.048***	0.046***	0.047***	0.043***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
DTD_{level}	0.060***	0.052***	0.044***	0.039***	0.031***	0.022***	0.015**	0.011*	0.005	-0.001	-0.008	-0.011*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
DTD_{trend}	0.151***	0.128***	0.113***	0.099***	0.077***	0.052***	0.035***	0.026***	0.022**	0.013	0.003	-0.005
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
$CASH/TA_{level}$	-0.612***	-0.540***	-0.464***	-0.384***	-0.318***	-0.268***	-0.222***	-0.203***	-0.175***	-0.146**	-0.134**	-0.116*
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.061)	(0.061)	(0.062)	(0.063)
$CASH/TA_{trend}$	-0.523***	-0.673***	-0.691***	-0.668***	-0.684***	-0.711***	-0.732***	-0.643***	-0.457***	-0.394**	-0.468***	-0.511***
	(0.150)	(0.156)	(0.159)	(0.163)	(0.165)	(0.165)	(0.166)	(0.168)	(0.170)	(0.173)	(0.176)	(0.178)
NI/TA_{level}	-3.156***	-3.113***	-3.133***	-3.146***	-3.101***	-3.097***	-3.110***	-3.074***	-2.994***	-2.963***	-2.891***	-2.810***
	(0.156)	(0.160)	(0.161)	(0.164)	(0.167)	(0.171)	(0.173)	(0.176)	(0.179)	(0.182)	(0.183)	(0.185)
NI/TA_{trend}	-2.107***	-2.140***	-1.865***	-1.573***	-1.300***	-1.048***	-0.936***	-0.791***	-0.794***	-0.769***	-0.752***	-0.686***
	(0.126)	(0.134)	(0.140)	(0.142)	(0.146)	(0.149)	(0.152)	(0.159)	(0.162)	(0.163)	(0.167)	(0.173)
$SIZE_{level}$	-0.217***	-0.216***	-0.215***	-0.215***	-0.214***	-0.212***	-0.209***	-0.207***	-0.203***	-0.199***	-0.197***	-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)
$SIZE_{trend}$	-0.424***	-0.412***	-0.476***	-0.581***	-0.639***	-0.647***	-0.658***	-0.661***	-0.616***	-0.559***	-0.525***	-0.487***
	(0.036)	(0.037)	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)
M/B	-0.061***	-0.071***	-0.080***	-0.089***	-0.090***	-0.091***	-0.090***	-0.084***	-0.078***	-0.075***	-0.065***	-0.061***
	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)
SIGMA	2.198***	2.090***	1.872***	1.618***	1.390***	1.198***	0.997***	0.946***	0.935***	0.891***	0.792***	0.777***
	(0.087)	(0.088)	(0.091)	(0.095)	(0.098)	(0.102)	(0.107)	(0.110)	(0.113)	(0.115)	(0.117)	(0.119)

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

	$\beta(12)$	$\beta(13)$	$\beta(14)$	$\beta(15)$	$\beta(16)$	$\beta(17)$	$\beta(18)$	$\beta(19)$	$\beta(20)$	$\beta(21)$	$\beta(22)$	$\beta(23)$
Intercept	-3.581***	-3.539***	-3.531***	-3.516***	-3.483***	-3.445***	-3.389***	-3.337***	-3.318***	-3.288***	-3.267***	-3.261***
	(0.065)	(0.066)	(0.067)	(0.067)	(0.068)	(0.068)	(0.069)	(0.068)	(0.069)	(0.070)	(0.070)	(0.071)
SP500	0.703***	0.739***	0.796***	0.754***	0.858***	0.849***	0.919***	0.995***	0.946***	0.959***	1.053***	1.049***
	(0.085)	(0.087)	(0.089)	(0.089)	(0.091)	(0.090)	(0.091)	(0.091)	(0.092)	(0.094)	(0.095)	(0.096)
Treasury rate	0.041***	0.037***	0.033***	0.032***	0.023**	0.020**	0.013	0.002	0.004	-0.001	-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)
$\mathrm{DTD}_{\mathrm{level}}$	-0.013**	-0.015**	-0.013*	-0.014**	-0.014*	-0.014**	-0.018**	-0.020***	-0.021***	-0.024***	-0.024***	-0.026***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$\mathrm{DTD}_{\mathrm{trend}}$	-0.005	-0.008	-0.005	-0.009	-0.009	-0.003	-0.004	-0.010	-0.010	-0.014	-0.010	-0.012
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
$\mathrm{CASH/TA}_{\mathrm{level}}$	-0.105*	-0.079	-0.070	-0.048	-0.042	-0.020	0.003	0.004	0.017	0.025	0.039	0.047
	(0.063)	(0.064)	(0.064)	(0.065)	(0.065)	(0.066)	(0.067)	(0.067)	(0.068)	(0.069)	(0.069)	(0.070)
$\mathrm{CASH/TA}_{\mathrm{trend}}$	-0.525***	-0.383**	-0.448**	-0.483***	-0.616***	-0.576***	-0.525***	-0.532***	-0.565***	-0.520***	-0.678***	-0.572***
	(0.178)	(0.177)	(0.177)	(0.180)	(0.183)	(0.187)	(0.186)	(0.193)	(0.196)	(0.195)	(0.196)	(0.199)
$\mathrm{NI/TA}_{\mathrm{level}}$	-2.823***	-2.805***	-2.757***	-2.630***	-2.563***	-2.580***	-2.490***	-2.438***	-2.357***	-2.301***	-2.282***	-2.299***
	(0.187)	(0.191)	(0.196)	(0.200)	(0.202)	(0.204)	(0.208)	(0.212)	(0.215)	(0.216)	(0.219)	(0.218)
$\mathrm{NI/TA}_{\mathrm{trend}}$	-0.790***	-0.886***	-1.016***	-0.798***	-0.755***	-0.723***	-0.561***	-0.547***	-0.457**	-0.748***	-0.561***	-0.575***
	(0.173)	(0.177)	(0.178)	(0.183)	(0.185)	(0.189)	(0.193)	(0.201)	(0.208)	(0.210)	(0.217)	(0.217)
$\mathrm{SIZE}_{\mathrm{level}}$	-0.191***	-0.188***	-0.187***	-0.184***	-0.183***	-0.181***	-0.174***	-0.168***	-0.165***	-0.164***	-0.161***	-0.161***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
$\mathrm{SIZE}_{\mathrm{trend}}$	-0.457***	-0.417***	-0.426***	-0.398***	-0.429***	-0.423***	-0.370***	-0.369***	-0.329***	-0.312***	-0.308***	-0.292***
	(0.041)	(0.041)	(0.042)	(0.042)	(0.043)	(0.044)	(0.044)	(0.045)	(0.045)	(0.046)	(0.047)	(0.047)
M/B	-0.060***	-0.063***	-0.063***	-0.060***	-0.059***	-0.059***	-0.057***	-0.054***	-0.052***	-0.050***	-0.052***	-0.047***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
SIGMA	0.741***	0.723***	0.761***	0.792***	0.755***	0.622***	0.690***	0.743***	0.760***	0.701***	0.707***	0.644***
	(0.120)	(0.122)	(0.124)	(0.125)	(0.127)	(0.132)	(0.134)	(0.137)	(0.138)	(0.140)	(0.142)	(0.144)

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

	$\beta(24)$	$\beta(25)$	$\beta(26)$	$\beta(27)$	$\beta(28)$	$\beta(29)$	$\beta(30)$	$\beta(31)$	$\beta(32)$	$\beta(33)$	$\beta(34)$	$\beta(35)$
Intercept	-3.262***	-3.273***	-3.279***	-3.268***	-3.265***	-3.238***	-3.212***	-3.189***	-3.186***	-3.176***	-3.172***	-3.173***
	(0.072)	(0.072)	(0.073)	(0.073)	(0.074)	(0.074)	(0.075)	(0.075)	(0.076)	(0.077)	(0.077)	(0.078)
SP500	1.017***	0.965***	0.958***	0.997***	0.908***	0.951***	0.984***	0.963***	0.905***	0.881***	0.829***	0.830***
	(0.096)	(0.097)	(0.098)	(0.099)	(0.098)	(0.099)	(0.101)	(0.103)	(0.104)	(0.105)	(0.106)	(0.105)
Treasury rate	-0.001	0.003	0.006	0.005	0.009	0.006	0.003	0.002	0.005	0.004	0.006	0.004
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
$\mathrm{DTD}_{\mathrm{level}}$	-0.025***	-0.024***	-0.025***	-0.029***	-0.031***	-0.033***	-0.036***	-0.037***	-0.038***	-0.041***	-0.042***	-0.041***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$\mathrm{DTD}_{\mathrm{trend}}$	-0.010	-0.002	0.004	0.003	-0.003	-0.006	-0.012	-0.020*	-0.019*	-0.022*	-0.028**	-0.022**
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
$\mathrm{CASH/TA}_{\mathrm{level}}$	0.066	0.069	0.116	0.135*	0.124*	0.142*	0.139*	0.156**	0.155**	0.140*	0.130*	0.137*
	(0.071)	(0.071)	(0.071)	(0.072)	(0.073)	(0.073)	(0.074)	(0.074)	(0.075)	(0.076)	(0.077)	(0.077)
$\mathrm{CASH/TA}_{\mathrm{trend}}$	-0.569***	-0.562***	-0.476**	-0.242	-0.263	-0.178	0.110	0.118	-0.002	-0.065	-0.131	-0.285
	(0.198)	(0.197)	(0.202)	(0.204)	(0.208)	(0.212)	(0.216)	(0.220)	(0.226)	(0.222)	(0.224)	(0.227)
$\mathrm{NI/TA}_{\mathrm{level}}$	-2.328***	-2.304***	-2.268***	-2.148***	-2.184***	-2.123***	-2.074***	-2.064***	-2.060***	-1.969***	-2.070***	-2.024***
	(0.223)	(0.224)	(0.229)	(0.233)	(0.236)	(0.238)	(0.241)	(0.243)	(0.245)	(0.249)	(0.249)	(0.253)
$\mathrm{NI/TA}_{\mathrm{trend}}$	-0.696***	-0.601***	-0.444*	-0.117	-0.328	-0.508**	-0.564**	-0.494**	-0.321	-0.304	-0.386	-0.336
	(0.216)	(0.222)	(0.229)	(0.234)	(0.230)	(0.231)	(0.232)	(0.238)	(0.248)	(0.250)	(0.246)	(0.251)
$\mathrm{SIZE}_{\mathrm{level}}$	-0.159***	-0.157***	-0.156***	-0.156***	-0.155***	-0.152***	-0.149***	-0.146***	-0.145***	-0.146***	-0.146***	-0.146***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
$\mathrm{SIZE}_{\mathrm{trend}}$	-0.287***	-0.293***	-0.275***	-0.271***	-0.245***	-0.212***	-0.194***	-0.178***	-0.176***	-0.167***	-0.139***	-0.125**
	(0.048)	(0.048)	(0.049)	(0.049)	(0.050)	(0.050)	(0.050)	(0.051)	(0.052)	(0.052)	(0.052)	(0.053)
M/B	-0.051***	-0.048***	-0.048***	-0.041***	-0.038***	-0.037***	-0.032***	-0.036***	-0.032***	-0.024**	-0.026***	-0.027***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
SIGMA	0.637***	0.660***	0.624***	0.554***	0.512***	0.529***	0.523***	0.541***	0.477***	0.434***	0.393**	0.434***
	(0.146)	(0.147)	(0.149)	(0.152)	(0.151)	(0.152)	(0.154)	(0.156)	(0.158)	(0.160)	(0.162)	(0.165)

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 (1991-2010). Panel B presents the out-of-sample (cross-section) results for the entire sample period (1991-2010) 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 (2001-2010). We re-estimate the model at each month-end starting from the first month of 2001 and use only the data available at the time for estimation. For each panel, we report the results for the full sample, the non-financial sub-sample and the financial sub-sample. The full-sample accuracy ratios based on the smoothed parameter values are also reported, where smoothing is performed by applying the Nelson-Siegel (1987) method.

Panel A: In-sample result						
	1 month	3 months	6 months	12 months	24 months	36 months
Full sample	92.81%	90.85%	88.06%	82.70%	73.44%	65.69%
Full sample (smoothed)	92.88%	90.91%	88.11%	82.67%	73.42%	65.67%
Non-financial	92.52%	90.33%	87.34%	81.78%	73.00%	65.92%
Financial	95.17%	94.49%	92.92%	89.07%	76.07%	58.32%
Panel B: Out-of-sample (cross-section) result						
	1 month	3 months	6 months	12 months	24 months	36 months
Full sample	92.42%	89.89%	86.42%	80.69%	71.48%	64.00%
Full sample (smoothed)	92.12%	89.75%	86.46%	80.73%	71.45%	64.00%
Non-financial	92.05%	89.35%	85.57%	79.66%	71.23%	64.49%
Financial	98.02%	96.64%	95.63%	90.51%	75.85%	55.62%
Panel C: Out-of-sample (over time) result						
	1 month	3 months	6 months	12 months	24 months	36 months
Full sample	91.74%	90.06%	87.24%	82.39%	75.90%	69.80%
Full sample (smoothed)	91.71%	90.10%	87.18%	82.26%	76.00%	69.78%
Non-financial	91.70%	89.94%	87.02%	82.03%	75.99%	70.18%
Financial	94.61%	93.81%	91.90%	88.79%	81.57%	71.62%

Table 7. Comparing the forward intensity model with the DSW model

This table reports the accuracy ratios for the forward intensity model and the DSW model developed in Duffie, *et al* (2007). We randomly select 3,000 non-financial firms to conduct the tests. The covariates as well as the autoregressive model are same as Duffie, *et al* (2007). Panel A reports the in-sample results for all firms and the entire sample period (1991-2010). Panel B still employs the in-sample estimates but reports the results for the period 2001-2010 which coincides with the period for the out-of-sample tests. Panel C reports the out-of-sample (over time) results for the sample period (2001-2010). For the out-of-sample comparison, we re-estimate both models at the end of January 2001 and use only the data available at that time of estimation.

Panel A: In-sample result (1991-2010)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	90.56%	89.29%	87.68%	85.17%	81.62%	79.60%
Forward Intensity	90.56%	88.65%	85.71%	80.65%	71.85%	65.91%
Panel B: In-sample result (2001-2010)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	90.96%	90.28%	88.62%	87.67%	83.46%	79.42%
Forward Intensity	90.96%	89.96%	87.29%	85.01%	78.73%	73.84%
Panel C: Out-of-sample (over time) result (2001-2010)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	91.21%	90.12%	87.08%	81.80%	69.05%	62.19%
Forward Intensity	91.21%	90.36%	87.81%	85.51%	78.63%	74.25%