Modeling the Loss Distribution

Sudheer Chava, Catalina Stefanescu and Stuart Turnbull *

April 21, 2008

Abstract

This paper focuses on modeling and predicting the loss distribution for credit risky assets such as bonds or loans. We directly model the two components of loss — the default probabilities and the recovery rates given default, and capture the dependence between them through shared covariates. Using an extensive default and recovery data set, we demonstrate the limitations of standard metrics of prediction performance which are based on the relative ordinal rankings of default probabilities. We use different approaches for assessing model performance, including a measure based on the actual magnitude of default probabilities that is more suitable for validating the loss distribution. We show that these approaches allow differentiation of default and recovery models which have virtually identical performance under standard metrics. We elucidate the impact of the choice of default and recovery models on the loss distribution through extensive out-of-sample testing. We document that the specification of the default model has a major impact on the predicted loss distribution, while the specification of the recovery model is less important. Further, we analyze the dependence between the default probabilities and recovery rates predicted out-of-sample. We show that they are negatively correlated, and that the magnitude of the correlation varies with the seniority class, the industry and the credit cycle.

*Sudheer Chava: Mays School of Business at Texas A&M University. Email: schava@mays.tamu.edu. Catalina Stefanescu: London Business School. Email: cstefanescu@london.edu. Stuart Turnbull: Bauer College of Business at University of Houston. Email: sturnbull@uh.edu. We are grateful to Alexander McNeil, Amiyatosh Purnanandam, Tyler Shumway, Matthew Spiegel (the Editor), Raman Uppal, seminar participants at the Bank of England, ETH Zürich, the Federal Reserve Board (Washington), McGill University, Rice University, York University, conference participants at BMBF Münich, Derivatives, Securities and Risk Management Conference (FDIC), and the INFORMS 2007 Annual Meeting, as well as to two anonymous referees for helpful comments which greatly improved the paper. All remaining errors are our own.

1 Introduction

This paper focuses on modeling and validating the loss distribution for credit risky assets such as bonds or loans, an important and difficult issue for financial intermediaries and regulators. The main components of the loss distribution are probabilities of default and recovery rates given default. There is a large literature on predicting defaults, and a separate emergent literature on modeling recovery rates. In this paper we generate the loss distribution by jointly modeling the default probabilities and the recovery rates with shared covariates, and we analyze their interdependence. To the best of our knowledge, this is the first study that explicitly models the loss distribution by jointly modeling defaults and recoveries, and that investigates the impact of the choice of default and recovery models on the loss distribution.

This is particularly relevant in light of the advanced internal ratings based approach in the Basel II accord, which allows banks to develop their own estimates of default probabilities and of recovery rates, so that these reflect the nature of their portfolios. Banks now have an incentive to use their own estimates to model the loss distributions, however they face challenges raised by the plethora of existing default and recovery models. First, to what extent does the predicted loss distribution depend on the choice of default and recovery models? Second, what are the criteria for choosing appropriate default and recovery models for generating the loss distribution? Basel II stipulates that banks' estimates of default probabilities and recovery rates are subject to supervisory review, but it doesn't explicitly indicate how financial institutions can show that their quantitative models are reasonable in order to gain regulatory approval for the resulting estimates.[1] This paper addresses these issues and makes three contributions.

Our main contribution is to elucidate the impact that the choice of default and recovery models has on the predicted loss distribution. Using an extensive default and recovery data set,[2] we investigate four default models and three recovery models inspired by extant finance literature, and use them to predict out-of-sample the loss distributions in different portfolios of bonds. We first find that based on the standard performance metrics for default and recovery prediction used in the literature, there is virtually no performance difference both between the four default models and between the three recovery models. This may suggest that the choice of any particular combination of default and recovery models should have little impact on the predicted loss distribution. We show, however, that the specification

[1] In fact, regulators themselves are also unsure how to assess whether the models that an institution uses are reasonable (BCBS, 2005b).

[2] The sample is drawn from firms in the intersection of CRSP, Compustat, and Moody's Default Recovery Database during 1980–2004. Section 3 has more details on the sample and variable construction.

of default model has a crucial impact on the predicted loss distribution. We find that the Shumway (2001) default model predicts loss distributions where the expected loss is closest to the actual realized loss in the portfolio, whereas the Duffie, Saita and Wang (2007) default model leads to consistent overestimation of the realized loss. We show that the choice of recovery model has a smaller impact on the predicted loss distribution than the choice of default model. This impact does not depend on the inclusion of macroeconomic variables, and it varies only marginally with the choice of obligor-specific variables in the recovery model.

Our second contribution is to propose a framework for validating models based on new approaches for assessing out-of-sample prediction performance. Existing performance measures for out-of-sample predictions, such as the power curve and the ability to rank defaulted firms in top risk deciles, can assess whether a given model correctly identifies some firms as being relatively riskier than others, but they do not indicate whether a model generates default probabilities that are adequate in absolute terms. This is, however, a crucial property required from any default model used for loss estimation, since the predicted loss distribution depends on the actual magnitudes of default probabilities rather than on the relative firm riskiness ranking. For our data set, all four default models have virtually identical prediction performance under standard metrics. We find that a more complex picture emerges, however, when other performance metrics are taken into account. First, we use the common ratio introduced by Mamaysky, Spiegel and Zhang (2007), a metric based on ordinal rankings which measures the extent to which models rank the same firm in the same decile. For our data set, this metric highlights differences in the predictions from the four default models. Further, we present an alternative approach for assessing default prediction accuracy, which consists in comparing the actual realized number of defaults in a given portfolio with the total number of defaults in that portfolio predicted out-of-sample. Using this alternative validation approach, we find that there are substantial differences between the out-of-sample predictions of the four default models. For a specific portfolio of telecommunication firms, the expected number of defaults closest to the realized benchmark is predicted by the Shumway (2001) default model, while the expected number of defaults farther away from the benchmark is predicted by the Duffie et al. (2007) model.

In the default model specifications we explicitly account for the effects of unobservable measurement errors, missing variables and incomplete information[3] on default probabilities.

[3]In practice investors usually have only incomplete information about the true state of a firm. There are differences between firms that affect their default probabilities but are not directly observable, such as variations in managerial styles, in the skill sets of workers, and in firm culture. Even differences in such areas as production skills, resource usage, cost control, and risk management are only partially revealed

To this end we use frailty models[4] which extend the standard hazard function approach; they incorporate unobservable firm heterogeneity by multiplying the hazard rate by a latent random variable common to all firms in a given industry group. We find that accounting for the unobserved heterogeneity does not significantly increase the prediction performance of default models as assessed by standard performance metrics, and we provide an explanation as to why this is not surprising. We further show that frailty models, however, predict different distributions for the total number of defaults in a portfolio than the corresponding default models without frailty.

Our third contribution is to analyze the dependence of predicted default probabilities and predicted recovery rates in our joint modeling framework based on shared covariates, and to generate new insights related to this dependence. The Basel II framework recognizes that changes in the probability of default and in the loss given default are generally related for most asset classes, and it requires financial institutions using the advanced internal ratings based approach to recognize this dependence (BCBS, 2005a). We demonstrate that the default probability and recovery rate predicted out-of-sample are negatively correlated, consistent with regulatory requirements. Furthermore, the magnitude of the correlation varies with industry and seniority class. We find that the inclusion of macroeconomic variables in the recovery models has virtually no impact on the variation over time of the correlation between predicted recovery rates and default probabilities, while the particular choice of obligor specific variables included in recovery models has only a marginal impact on the annual variation of the correlation. The dependence between the default probabilities and the recovery rates also implies that the loss distribution for a portfolio of loans will vary substantially from that estimated using the foundation internal ratings based approach with its assumed loss given default.

This paper is related to several different strands of previous research. There is a large and growing literature devoted to the modeling of the probability of default — see Shumway (2001), Chava and Jarrow (2004), Campbell, Hilscher and Szilagyi (2006), Duffie et al. (2007), and Bharath and Shumway (2007). An extensive survey of methodologies is given in Altman and Hotchkiss (2005). There is also an emerging literature addressing the modeling of the determinants of the recovery rate given default. A survey of empirical evidence regarding the properties of recovery rates is given in Schuermann (2004) — see also Acharya, Bharath and Srinivasan (2003). Several studies model the dependence between the proba-

in accounting statements. The uncertainty surrounding investors' projections will, in general, depend on the state of the economy, the state of the particular sector in which the company operates, and the unique characteristics of the firm.

[4] Gagliardini and Gourieroux (2003) and Schönbucher (2003) introduced the concept of frailty in the finance literature. For recent empirical work see Duffie, Eckner, Horel and Saita (2006).

bility of default and the recovery rate given default, by assuming there is a common latent factor affecting both (Frye, 2000; Pykhtin, 2003; Dullmann and Trapp, 2004). However, to the best of our knowledge, there are no empirical studies that attempt to explicitly model the covariates affecting the probability of default, the recovery rate given default, their dependence and the impact on the loss distribution.

The paper is structured as follows. In Section 2 we develop our modeling methodology, and in Section 3 we describe the data set used in this study. The empirical results for the estimation of the probability of default and of the recovery rate are given in Section 4. In Section 5 we investigate the modeling of the loss distribution, and Section 6 concludes the paper with a summary of our findings.

2 The Default and Recovery Models

In this section we first describe the specification of default models with unobservable heterogeneity and develop the estimation methodology. Next, we discuss several specifications of recovery rate models.

2.1 The Default Models

2.1.1 Model specification

The sample data contains firms grouped in G groups or industries. Let n_i be the number of firms in the ith group, and $n = \sum_{i=1}^{G} n_i$ be the total number of firms in the sample. During the observation period $[0, T]$, any particular firm may experience a default, may leave the sample before time T for reasons other than default (for example a merger, an acquisition, or a liquidation), or may survive in the sample until time T. A firm's lifetime is said to be censored if either default does not occur by the end of the observation period, or if the firm leaves the sample because of a non–default event. Let T_{ij} denote the observed (possibly censored) lifetime of the jth firm in the ith group, and let $\perp_{ij}$ be the censoring indicator, where $\perp_{ij}= 1$ if T_{ij} is a default time and $\perp_{ij}= 0$ if T_{ij} is a censoring time. The total number of failures in group i is given by $\perp_{i.}= \sum_{j=1}^{n_i} \perp_{ij}$. For every $s = 1, \ldots, d$, let $\delta_{ij}(s) = 1$ if the jth firm in the ith group is in the sample at time t_s, and zero otherwise. For example, if the firm is in the sample at the beginning of the observation period and censoring only occurs at time T, then $\delta_{ij}(s) = 1$, for $s = 1, ..., T$.

Let $X_{ij}(t)$ be a $1 \times K$ vector of covariates at time t. The vector $X_{ij}(t)$ usually includes a constant component representing an intercept term, and it is composed of both firm-specific variables and macroeconomic variables. Information about the firm-specific variables terminates at time T_{ij}, and information about the macroeconomic variables is available at all times. We observe the covariates at discrete time intervals $0 < t_1 < t_2 \cdots < t_d \leq T$, and assume that $X_{ij}(t)$ is constant during the period between two consecutive observations.

Let $\lambda_{ij}(t)$ be the default intensity function (the hazard function) for the jth firm in the ith group. In order to model the correlation between defaults of firms in the same group, we assume that the unobservable heterogeneity can be represented by a latent non–negative random variable Y_i common to all firms in the same industry, which we shall refer to as frailty[5] and which represents the effects of the unobservable measurement errors and missing variables.[6] The shared frailty Y_i acts multiplicatively on the intensity functions $\lambda_{ij}(t)$, so that the hazard rates are specified by

$$\lambda_{ij}(t) = Y_i \exp(X_{ij}(t)\beta), \tag{1}$$

where β denotes the $K \times 1$ vector of regression parameters. Conditional on the unobserved Y_i, the lifetimes of firms in the ith group are independent. When the unknown Y_i is integrated out, the lifetimes become dependent; the dependence is induced by the common value of Y_i.

The shared frailty model specified by (1) is a natural approach for modeling dependence and taking into account unobservable heterogeneity. The model can be easily extended to the case where the frailties are time-varying, multivariate rather than univariate, or obligor

[5]An introduction to frailty models is given in Kiefer (1988), Klein and Moeschberger (1997, chapter 13), and Hougaard (2000, chapter 7). There is a large biostatistical and demographic literature on frailty modelling, but to–date there have been only a small number of applications in the credit risk area. Gagliardini and Gourieroux (2003) and Schönbucher (2003a) introduce the notion of unobservable heterogeneity or frailty to model information driven contagion.

[6]Let $X^T(t)$ represent the true value of the vector of covariates and $X(t)$ be the observed covariates, where we assume that

$$X_k^T(t) = X_k(t) + e_k(t).$$

Here $e_k(t)$ is the measurement error of the kth covariate for the firm. Hence $X^T(t)\beta = X(t)\beta + y(t)$, where $y(t)$ represents the effects of the measurement errors and β is a vector of parameters giving the dependence of the default intensity on the covariate vector. We shall assume that the baseline default intensity is $\lambda_0(t) = \exp(X(t)\beta)$.

If there are missing variables, let $m(t)$ denote the vector of missing variables and β^M the corresponding vector of parameters. The intensity is now given by

$$\lambda(t) = \exp(X(t)\beta + m(t)\beta^M + y(t)),$$

which we can rewrite as

$$\lambda(t) = Y(t)\exp(X(t)\beta).$$

specific rather than shared by all obligors in the same sector. Such extensions allow modeling of more flexible patterns of default dependence. For example, the shared frailty model (1) implies positive correlation of defaults within an industry; in practice, however, some degree of negative correlation may be conceivable, for example due to competition. The multivariate lognormal frailty model (Stefanescu and Turnbull, 2006) can accommodate negative default dependence as well.

The frailty has an assumed prior distribution which is updated as the default information set evolves over time. For example, if no firms within a particular sector default, this might help to increase confidence in the credit worthiness of the firms in this sector. Conversely, if there is a failure in a particular sector or the aggregate number of defaults in the economy increases, this might adversely affect the assessment of credit worthiness. There is a range of choices for the distribution of the frailties — due to mathematical convenience, the most popular is the gamma distribution[7] $G(r, \alpha)$. With gamma frailties, the scale parameter needs to be restricted for identifiability reasons, and the standard restriction is $r = \alpha$ as this implies a mean of one for Y. We complete the specification of model (1) by assuming that the sector frailties Y_i are independent and identically distributed with a gamma distribution $G(1/\theta, 1/\theta)$, with $\theta > 0$. The unconditional frailty means are thus equal to one, while the conditional means vary across sectors.

2.1.2 Model estimation: maximum likelihood

In this subsection we discuss the estimation of the parameters of the frailty default model through maximum likelihood. The parameters to be estimated are the regression coefficients β and the frailty variance θ. Let $L(\theta, \beta)$ denote the likelihood which is given by the product of the likelihood of the frailties and of the survival likelihood conditional on the frailties.[8]

[7]The gamma density function of Y is given by $f(y) = \alpha^r y^{r-1} \exp(-\alpha y) \cdot \frac{1}{\Gamma(r)}$, where $\Gamma(r)$ is the gamma function. The expected value is $E[Y] = r/\alpha$ and the variance is $\mathrm{Var}(Y) = r/\alpha^2$. The parameter α is the scale parameter and r is the shape parameter.

[8]It is also possible to assume that the covariates $\{X_{ij}(t)\}$ follow a stochastic process with parameter vector γ, for example, an autoregressive time series process. Then the sample likelihood would also include the likelihood function $L_X(\gamma)$ of the covariates; the maximization program separates, implying that γ is estimated separately from β and θ. In general, the estimation of γ is the standard numerical procedure of fitting a multivariate time series process to the covariate vectors $\{X(t)\}$.

Since lifetimes in each group are independent conditional on the group frailty, we obtain that

$$\begin{aligned} L(\theta,\beta) &= L_Y(\theta)\cdot\prod_{i=1}^{G}\prod_{j=1}^{n_i} L(\theta,\beta\,|\,T_{ij},\perp_{ij},X_{ij},y_i), \\ &= L_Y(\theta)\cdot\prod_{i=1}^{G}\prod_{j=1}^{n_i}[y_i\exp(X_{ij}(T_{ij})\beta)]^{\perp_{ij}}\cdot\exp\left(-\int_0^{T_{ij}}\lambda_{ij}(t)dt\right), \end{aligned} \tag{2}$$

where the likelihood function for the frailties is given by

$$L_Y(\theta)=\prod_{i=1}^{G} f(y_i)=\prod_{i=1}^{G}\frac{1}{\theta^{1/\theta}\Gamma(1/\theta)}y_i^{1/\theta-1}\cdot\exp(-y_i/\theta), \tag{3}$$

and the integrated hazard is

$$\int_0^{T_{ij}}\lambda_{ij}(t)dt = Y_i\sum_{s=1}^{d}\delta_{ij}(s)\exp(X_{ij}(t_s)\beta)\equiv Y_i\Lambda_{ij}. \tag{4}$$

Replacing expressions (3) and (4) in (2) and taking logarithms, it follows that the log–likelihood function is

$$\begin{aligned} \log L(\theta,\beta) &= \log L_Y(\theta)+\sum_{i=1}^{G}\sum_{j=1}^{n_i}[\perp_{ij}[\log(y_i)+X_{ij}(T_{ij})\beta]-y_i\Lambda_{ij}] \\ &= \sum_{i=1}^{G}[(1/\theta-1+\perp_{i.})\log(y_i)-y_i/\theta]-G[\log\Gamma(1/\theta)+(1/\theta)\log(\theta)] \\ &+\sum_{i=1}^{G}\sum_{j=1}^{n_i}[\perp_{ij}X_{ij}(T_{ij})\beta-y_i\Lambda_{ij}]. \end{aligned} \tag{5}$$

In order to maximize the likelihood, we use the Expectation–Maximization (EM) algorithm (Dempster, Laird and Rubin, 1977), which is the classic tool for obtaining maximum likelihood estimates from incomplete or missing data. The complete data consists of the realized values of the frailties $Y_1,\ldots,Y_G$ and the uncensored lifetimes. The observed but incomplete data consists of the observed lifetimes $\{T_{ij}\}$ and the censoring indicators $\{\perp_{ij}\}$. The EM algorithm starts with some initial estimates; for the β coefficients these can be computed by ignoring the frailty terms, and the initial estimate for the frailty variance θ can be set equal to one. Then the algorithm iterates between two steps: the expectation (E) step computes expected values of the sufficient statistics for the complete data, conditional on the observed data and current values of the parameters. In the maximization (M) step, new estimates of the unknown parameters are obtained by numerically maximizing the likelihood computed with the expected values of the sufficient statistics from the previous E–step. These two steps are repeated until convergence is achieved, and it can be shown that, under mild conditions, the EM algorithm converges to the maximum likelihood estimates.

Conditional on the observed data $\{T_{ij}, \perp_{ij}, X_{ij}\}$ and on the current values of parameters θ and β, the frailty Y_i has a gamma distribution $G(A_i, C_i)$ with scale parameter $C_i = 1/\theta + \sum_{j=1}^{n_{ij}} \Lambda_{ij}$ and shape parameter $A_i = \perp_{i.} + 1/\theta$. The conditional means are therefore

$$\begin{aligned} E[Y_i] &= A_i/C_i, \\ E[\log(Y_i)] &= \psi(A_i) - \log(C_i), \end{aligned} \quad (6)$$

where $\psi(\cdot)$ is the digamma function. From (5) and (6) it follows that the expected value of the log–likelihood function which is maximized in the M-step is given by

$$\begin{aligned} E[\log L(\theta, \beta)] = & \sum_{i=1}^{G} (1/\theta - 1 + \perp_{i.})[\psi(A_i) - \log(C_i)] - [A_i/C_i]/\theta \\ & -G[\log \Gamma(1/\theta) + (1/\theta)\log(\theta)] + \sum_{i=1}^{G} \sum_{j=1}^{n_i} \perp_{ij} X_{ij}(T_{ij})\beta - [A_i/C_i]\Lambda_{ij}. \end{aligned}$$

After convergence of the EM algorithm, the standard errors of the estimates of θ and β can be computed from the inverse of the observed information matrix. Using these estimates, we can also calculate the expected value of the frailty for each group.

This methodology can be easily extended to the case of competing risks.[9] Firms may exit the sample for reasons other than default, such as a merger or an acquisition, and these non–default events are all competing risks that may cause censoring of a firm's lifetime. With multiple causes for exit, we may consider a multivariate frailty model with one frailty component for each cause of exit. The likelihood function is separable, under the assumption that the frailty components are independent with gamma distributions. Maximum likelihood estimates of the frailty variances and of the covariate effects parameters can then be computed using an extension of the EM algorithm as outlined previously.

[9] An introduction to competing risk models is given in Crowder (2001). See also Hougaard (2000), Lawless (2003), and Duffie et al. (2007).

2.1.3 Default prediction with frailty models

Frailty survival models are particularly suitable for default modeling, as an approach that is able to account for both unobservable heterogeneity and default contagion. In our extensive empirical experience with default data,[10] frailty models have proved very flexible for capturing default correlation and have shown a substantial improvement in in-sample goodness-of-fit measures over survival models without frailty. For out-of-sample performance measures, however, the results are mixed. Empirical studies found that frailty models do not perform significantly better out-of-sample than the similar hazard models without frailty, when the performance measure is defined as the percentage of defaulted firms correctly ranked in the top risk deciles (Duffie et al., 2006). Across different data sets, time horizons, covariate specifications and clustering patterns, the improvements in out-of-sample prediction by taking frailty into account are at best marginal. In this subsection we give a theoretical justification for these empirical insights, and argue that the benefits from using frailty default models are only apparent when different measures of out-of-sample prediction performance are used.

Without loss of generality, in this subsection we focus on out-of-sample prediction performance for firms in one given industry group during one given year in the out-of-sample horizon, as the results will hold aggregating over all industry groups and all years in the horizon. To simplify notation, we therefore drop the year and industry group subscripts for the remaining of this subsection.

Let n be the number of firms alive (and at risk) at the beginning of the year, and let X_i be the covariate vector for firm i, with $i = 1, \ldots, n$. We consider two hazard rate models with exponential intensities, where the hazard functions for firm i are given by

$$\lambda_{f,i}(t) = Y \exp(\beta X_i), \tag{7}$$

$$\lambda_{nf,i}(t) = \exp(\beta X_i), \tag{8}$$

where β is the vector of regression parameters. Model (7) is a frailty model where the frailty Y is shared by all firms in the group. Model (8) is the corresponding hazard rate model without frailties. Denote $\alpha_i = \exp(\beta X_i)$ with $\alpha_i > 0$, and let L be the Laplace transform of the distribution of Y, defined as $L(s) = E_Y[\exp(-sY)] = \int \exp(-sy) f(y)\, dy$, where E_Y is the expectation taken with respect to Y and $f(y)$ is the density function of Y.

The one-year ahead probability of default for obligor i predicted by the hazard rate

[10]The details and results are available upon request from the authors — see also Chava, Stefanescu and Turnbull (2006).

model (8) without frailty is given by

$$PD_{nf,i} = 1 - \exp[-\exp(\beta X_i)] = 1 - \exp(-a_i). \tag{9}$$

The one-year ahead probability of default for obligor i conditional on the frailty Y and predicted by frailty model (7) is given by

$$PD_{f,i|Y} = 1 - \exp[-Y \exp(\beta X_i)] = 1 - \exp(-Ya_i), \tag{10}$$

hence the unconditional probability of default is

$$PD_{f,i} = E_Y[PD_{f,i|Y}] = E_Y[1 - \exp(-Ya_i)] = 1 - L(a_i). \tag{11}$$

Let us denote by $\perp$ the default indicator defined for each obligor i by

$$\perp_i = \begin{cases} 1, & \text{if obligor } i \text{ defaults} \\ 0, & \text{otherwise} \end{cases},$$

and let $D = \sum_{i=1}^{n} \perp_i$ be the number of obligors defaulting during the year. The following proposition gives the expected numbers of defaults under the two models (7) and (8). The proof is included in Appendix A.

Proposition 1 *The expected number of defaults predicted by the hazard rate model (7) with frailty is given by*

$$E_f[D] = \sum_{i=1}^{n} [1 - L(\alpha_i)], \tag{12}$$

and the expected number of defaults predicted by the hazard rate model (8) without frailty is

$$E_{nf}[D] = \sum_{i=1}^{n} [1 - \exp(-\alpha_i)]. \tag{13}$$

The following proposition describes the effect of accounting for frailty on the predicted number of defaults. The proof is also included in Appendix A.

Proposition 2 *If the mean of the frailty distribution $E[Y]$ is not greater than one, then the expected number of defaults predicted by the frailty model is less than that predicted by the hazard rate model without frailty.*

In particular, if the frailty has a gamma distribution $Y \sim G(1/\theta, 1/\theta)$, then $E(Y) = 1$ and Proposition 2 implies that $E_f[D] \leq E_{nf}[D]$.

The difference between the default probabilities estimated by the hazard rate models with and without frailties is given by

$$PD_{nf,i} - PD_{f,i} = L(\alpha_i) - \exp(-\alpha_i).$$

In practice, α_i is very small. Since the Laplace transform for the gamma frailty is $L(s) = (1 + s\theta)^{-1/\theta}$, we can show using a Taylor's series expansion that the two default probabilities are approximately equal, that is $PD_{nf,i} - PD_{f,i} = O(\alpha_i)$. This provides an explanation for the results discussed later in Section 4.2.3, which show that there are only minor differences in prediction performance for hazard rate default models with and without frailty.

2.2 The Recovery Rate Models

In this subsection we briefly describe several specifications of recovery rate models. Let $R_i(t)$ be the recovery rate of firm i at time t. We assume that $R_i(t)$ depends on a set of covariates $X_i(t)$ through a function of the linear form $X_i(t)\beta_r$, where β_r is a vector of regression coefficients. Note that the covariate vector $X_i(t)$ may include macroeconomic, industry, firm, and bond specific variables.

Many of the extant studies assume that recovery rates depend linearly on the available covariates (Acharya et al., 2003; Varma and Cantor, 2005), so that

$$R_i(t) = X_i(t)\beta_r.$$

Note, however, that in practice recovery rates are always non-negative and usually less than one.[11] Since the linear specification implies that the recovery rates are unconstrained, it may lead to predicted recovery rates that are negative or greater than one and it is thus not appropriate for modeling recoveries. We investigate instead two other specifications.

The probit transformation gives

$$R_i(t) = \Phi(X_i(t)\beta_r),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, and

[11] It is sometimes possible for recovery rates to be greater than one, especially if bond prices within one month of default are used. This, however, is an anomaly that only rarely happens.

it implies that $X_i(t) = \Phi^{-1}(R_i(t))$ (Andersen and Sidenius, 2005).

The logit transformation gives

$$R_i(t) = \frac{1}{1 + \exp(X_i(t)\beta_r)},$$

implying that $X_i(t) = \log(L_i(t)/R_i(t))$, where $L_i(t) = 1 - R_i(t)$ is the loss given default (Schönbucher, 2003b).

In practice, the logit and probit models lead to very similar insights. Indeed, for our empirical analysis of recovery rates reported in Section 4 we obtained very similar goodness-of-fit measures and predicted recovery rates under the assumptions of a logit or probit specification.

3 Data Description

In this section we first describe the data sources and then discuss the covariates used at different stages of the analysis. Summary statistics for all covariates used in the study are available in Table 1.

3.1 Data Sources

Our primary data source for the empirical analysis is Moody's Default Recovery Database (Moody's DRS) that contains detailed information on all bonds that are rated by Moody's during our sample period 1980–2004. In particular, Moody's DRS has information on default history of the bonds and recovery rates in the event of default. See Covitz and Han (2004), Varma and Cantor (2005) and Duffie et al. (2007) for more details on Moody's DRS. We restrict our attention to only those firms that are in the intersection of Moody's DRS, CRSP, and COMPUSTAT databases during 1980–2004. The COMPUSTAT (active and research) files for this period provide the firm level balance sheet data, and CRSP provides the market data.

The default data contains one record for each year of each firm's existence, from the year of listing to the year when the firm has defaulted.[12] The variable indicating default is coded

[12]We use Moody's definition of default in our analysis. Moody's defines default as the event that one or more of the following happen: (a) There is a missed or delayed disbursement of interest and/or principal, including delayed payments made within a grace period. (b) The company files for bankruptcy, administration, legal receivership, or other legal blocks to the timely payment of interest or principal. (c) A distressed

as one in the year of default and zero otherwise. Some firms get delisted from the exchange and may default at a later date. Since there is no market information available for delisted firms, we follow Shumway (2001) and code the year of delisting as the year of default in cases where the firm defaults within five years after delisting. We group firms into industry groups based on four digit DNUM codes. Our data contains 518 groups ranging in size from 1 to 151 firms, with a mean size of 7.9 firms and a median size of 4 firms. The number of defaults in each group ranges from 0 to 24.

The recovery data contains one record for each defaulted bond, where recovery rate on a bond is measured as the bond price within a month after default as given by Moody's DRS.

3.2 Covariates

3.2.1 Macroeconomic variables

In this study we report the results of our investigation on the effects of five macroeconomic variables. These include the *term spread* computed as the difference between the ten year Treasury yield and the one year Treasury yield, the *credit spread* computed as the difference between AAA and BAA yields, and the *three month Treasury yield*, all taken from Federal Reserve's H.15 statistical release. We also use the *S&P 500 index trailing one year return* computed from CRSP, and the *logarithm of the amount of total defaulted debt (in billions of USD)* taken from Moody's 2006 default study.

One would expect that default probabilities increase and recovery rates decrease when the credit spread, the Treasury bill rate, or the logarithm of total defaulted debt increase, since these signal a weaker economy. Conversely, a high value of the return on the S&P 500 index signals a stronger economy, hence we would expect default probabilities to decrease and recovery rates to increase with increasing S&P 500 return.[13]

exchange takes place. This happens either when the exchange has the apparent purpose of helping the borrower avoid default, or when the issuer offers bondholders a new security or a package of securities that represent a diminished financial obligation (such as preferred or common stock, or debt with a lower coupon or par amount, lower seniority, or longer maturity).

[13]In our extensive analysis we also investigated several other macroeconomic covariates that have been previously used in the default and recovery literatures. These covariates include the global speculative grade default rate, the growth in real GDP, the growth in personal income, and the Chicago Fed National Activities Index. We did not find consistently significant effects of these macroeconomic variables in the default and recovery models, and they did not lead to any additional specific insights. Therefore, we do not mention these covariates in our subsequent discussions of the analysis results.

3.2.2 Industry factors

For the frailty default models we use industry groups based on four digit DNUM codes. For the recovery rate models we identify four broad industry classes: transportation, utilities, industrials and financials. We take the transportation class as a baseline and construct dummy variables for the other three industry classes.

3.2.3 Firm level variables

We follow Shumway (2001) in constructing the following firm level variables: the *relative size of the firm* defined as the logarithm of each firm's equity value divided by the total NYSE/AMEX/NASDAQ market capitalization, the *excess return* defined as the return on the firm minus the value-weighted CRSP NYSE/AMEX/NASDAQ index return, the *ratio of net income to total assets of the firm* extracted from COMPUSTAT, the *ratio of total liabilities to total assets of the firm* also extracted from COMPUSTAT, and the *volatility* defined as the idiosyncratic standard deviation of the firm's monthly stock returns computed from CRSP.

Additionally, we also use the firm's *trailing one year stock return* computed by cumulating the firm's monthly return from CRSP, and the *distance-to-default*, essentially a volatility corrected measure of leverage based on Merton (1974) and constructed as in Bharath and Shumway (2007). More details on the construction of this variable are given in Appendix B.

We also construct the following firm level variables from COMPUSTAT for use in the recovery models: the *logarithm of the total assets of the firm*, the *market to book ratio* (a proxy for the firm's growth prospects), and the *ratio of property plant and equipment to total assets* (a measure of the firm's tangible assets). To avoid any outlier effects, all variables are winsorized at the 1% and 99% of the cross–sectional distributions.

For default models, we would expect that default probabilities increase as a result of an increase in the total liabilities to total assets ratio, or in the standard deviation of the firm's returns. Conversely, an increase in relative size, excess return, ratio of net income to total assets, or distance to default, signals a decrease in the risk of a firm, hence we would expect default probabilities to decrease. For recovery rates models, we would expect the market to book ratio to have a positive impact on recoveries.

3.2.4 Bond level variables

Our recovery rate models include the *coupon rate*[14] and *seniority* as bond level variables. We identify five classes of seniority in ascending order of claim priority: junior, subordinated, senior subordinated, senior unsecured and senior secured. We take junior bonds as baseline and construct dummy variables for the other four seniority classes.

In addition to coupon rate and seniority, we also considered the impact on recovery rates of the initial amount issued and of the time to maturity. These covariates, however, did not have consistently significant effects on recovery rates in the wide range of model specifications that we investigated. Therefore, we do not mention them in our later discussion of the analysis insights.

4 Default and Recovery — Empirical Results

In this section we discuss the estimation results for the default and recovery models described in Section 2.

4.1 Default Models: In-Sample Estimates

We report here the results for four default models, including the models from Shumway (2001) and Duffie et al. (2007). The first model MD1 includes excess return, volatility, relative size, net income to total assets and total liabilities to total assets. These are the same covariates as in Shumway (2001), except for firm age which Shumway found statistically insignificant. The second model MD2 is a reduced form specification obtained by dropping the volatility and the accounting variables net income to total assets and total liabilities to total assets from model MD1, and including distance-to-default. The third model MD3 is obtained by adding two macroeconomic variables, the term spread and the credit spread to model MD1. The fourth model MD4 includes the same covariates as in Duffie et al. (2007) – the stock return, the distance-to-default, the S&P 500 return and the three month Treasury yield. These four models are a subset of the many default models that we investigated.[15]

[14]Acharya et al. (2003) argue that if a bond is issued at a discount or premium, then the coupon on the bond will affect the accelerated amount payable to bondholders in bankruptcy, as will the remaining maturity of the issue.

[15]In particular, we also investigated a private firm model that does not utilize any equity market based variables, and a "saturated" model that included all the firm-specific and macroeconomic covariates available in our data set. These models did not provide additional insights over those derived from models MD1–MD4, hence for the sake of conciseness we do not report these results here.

We group firms into industry groups based on four digit and three digit DNUM codes. Our data contains 518 sectors defined by the four digit DNUM codes, and 299 sectors defined by the three digit DNUM codes. The two classifications gave very similar results, the only substantial difference being a higher estimated frailty variance when firms are classified based on four digit DNUM codes than on three digit DNUM codes. This is to be expected, as the frailty variance is a measure of within sector homogeneity, and sectors are more homogeneous when they are defined according to a more refined classification.[16] Consequently we only report the coefficients obtained from the classification using four digit DNUM codes.

Table 2 reports the estimation results.[17] The frailty variance is statistically significant for all models, indicating evidence of default clustering. Almost all the covariate effects parameters are significant and of the expected sign, with the exception of the coefficient for net income to total assets[18] in models MD1 and MD3. The effects of credit spread and term spread variables in model MD3 are only marginally significant; in fact, their inclusion in the model has only a small effect on the magnitudes of the other covariate coefficients, compared with the estimated values for model MD1. Note also that, as expected, the coefficient for the S&P 500 return is negative in model MD4 — the higher the return, the healthier is the economy and the smaller the default probabilities. This contrasts with the results from Duffie et al. (2007) who found that the S&P 500 return had a positive effect in the default model.

For purposes of comparison we also fitted the four default models with the same covariate specifications but without a frailty term in the hazard rate.[19] For models without frailty, the signs of all estimated coefficients generally remain unchanged, with only small changes in the magnitudes of the coefficients. For all four specifications there is a significant deterioration in the log-likelihood function compared to that of the corresponding default model with frailty. A χ^2 test confirms that for the same sets of covariates a frailty default model achieves a statistically significant improvement in fit over a default model without frailty.

[16] As the classification of the sectors becomes more refined, firms within a sector become more homogeneous and the dispersion across sectors increases. Hence the variance θ of the sector specific frailty Y increases.

[17] The log-likelihoods for the four models are not directly comparable, since they are based on slightly different data sets (the samples vary due to missing data for some variables). The parameter estimates, however, are comparable.

[18] The coefficient of this variable is also not significant in Shumway (2001).

[19] The estimation results for these models are available upon request from the authors.

4.2 Default Models: Out-of-Sample Performance

In this subsection we investigate the out-of-sample forecasting performance of the four default models, using a one year horizon as suggested by regulatory requirements. We take several approaches to assessing forecasting performance, and study the relative ability of these approaches to differentiate between models.

4.2.1 Individual firm defaults

The first approach that we consider for assessing prediction performance focuses on predicting individual firm defaults, and it has been extensively employed in extant literature. We define the out-of-sample horizon to be the period 1996–2004. For each year t during 1996–2004, we compute the parameter estimates for all models from data between 1980 and $t-1$, then we forecast default probabilities during year t for each firm that is alive at the beginning of that year. The firms are then ranked into deciles in descending order of their forecasted default probabilities. We record the annual number of actual defaults in each decile, and we also compute the percentages and counts of defaults in each decile aggregated over the 1996–2004 period.

Table 3 summarizes the percentages and counts in the top two deciles for all four models. The first part of the table presents the annual percentages and counts of defaults classified in the top two deciles (out of the total number of defaults) for each model and for each year between 1996–2004. For example, in 1996 model MD1 correctly ranked seven out of the eight defaulted firms (or 87%) in the top two deciles. Note that in some years there are small differences in the total number of defaults between models — this happens because slightly different data sets were used for model calibration and forecasting, due to missing data in some of the variables. The second part of the table presents for each model the percentages and counts of defaults classified in the top two deciles (out of the total number of defaults), aggregated over the entire horizon 1996–2004. Both the annual and the aggregate results show that the predictive performance of all four default models is very similar. Models MD1 and MD3 have the best aggregate performance, correctly identifying 93.62% of the defaulting firms in the first two deciles. Model MD4 correctly identifies 92% of the defaulting firms, consistent with the results in Duffie et al. (2007).

These insights are complemented by the power curves for each default model, reported in Figure 1. The out-of-sample power curves may be a better measure of predictive performance, as they account for both Type I and Type II classification errors. They are obtained by plotting on the vertical axis the cumulative percentage of defaulted firms, versus the rank percentile of firms on the horizontal axis. For example, the plot in Figure 1 shows that around 85% of the defaulted firms were classified in the top risk decile during the out-of-sample period. The plot also shows that the power curves for all four default models are very similar.

Beyond the issue of just predicting firm defaults, we study the extent to which models rank the same firm in the same decile. To this purpose we use the common ratio measure introduced by Mamaysky, Spiegel and Zhang (2007). The common ratio is computed for a given decile and pair of default models, and it is constructed as follows. First, firms are ranked by each model into deciles in descending order of their forecasted default probabilities. Next, for each decile and each pair of models MD_i and MD_j, we construct the union and the intersection of firms ranked by MD_i and MD_j in that particular decile. The common ratio $\eta_{i,j}$ for models MD_i and MD_j and the specific decile is defined as the number of firms in the intersection divided by the number of firms in the union. Table 4 reports the common ratios for each pair of default models and all ten deciles, as well as for the ranking into the joint top two deciles. The common ratios are aggregated over the out-of-sample horizon 1996–2004. For example, models MD1 and MD2 have a common ratio of $\eta_{1,2} = 0.65$ for the first decile, implying that among all the firms that were classified as high risk by *either* of the two models, 65% of the firms were classified as high risk by *both* models. The common ratio is thus a measure of ranking agreement among the models. In general, the maximum value for the common ratio in any given decile is one (indicating complete agreement between models), and the minimum value is zero (indicating complete disagreement between models).

The values of common ratios in Table 4 show that for the top two deciles the rankings from models MD1 and MD3 are almost identical, and those from models MD2 and MD4 are similar. This is also true for the last decile. There is far less agreement, however, for other pairs of models and other deciles. We concluded from Table 3 that models had similar performance in classifying defaulted obligors in the first two risk deciles. Table 4, however, shows that models still differ in their overall relative risk ranking of individual obligors.

4.2.2 Portfolio defaults

The predicted number of defaults in a portfolio is a critical input for the generation of the loss distribution. Consequently, the second approach that we consider for assessing model prediction performance focuses on predicting the total number of defaults in a portfolio. This approach has received less attention in literature, yet it is of major importance for risk and portfolio managers. Unlike the traditional power curve discussed in the previous subsection, predicting the total number of defaults involves the actual magnitudes of predicted default probabilities, rather than just the ordinal riskiness ranking of firms.

As an example, we focus on the portfolio of all firms with DNUM codes between 4800 and 4899 (these are all telecommunication firms), and we predict out-of-sample the total number of defaults in this portfolio during 2003. We focus on telecommunication firms because there were 15 defaults in this industry group in the period 2000–2002 immediately preceding our out-of-sample horizon; this large number of defaults leads us to expect that there is scope for predictions for this portfolio to give interesting insights on the performance differences between the four default models. Figure 2 gives the distributions of the number of defaults during 2003 predicted by all four models. The actual realized number of defaults in the portfolio during 2003 is three.

The four default models predict quite different distributions. A visual inspection of Figure 2 suggests that the expectation of the distribution generated by model MD1 is closest to the realized number of defaults equal to three. The distributions generated by the other three models are shifted to the right relative to model MD1, and the shift is substantial for the distribution generated by model MD4 whose expectation is much larger than the observed number of defaults. This substantial increase in both the mean and the dispersion of the distribution generated by MD4 relative to that generated by MD1, could be explained by the varying impact that omitting relevant covariates from the model has on the estimated default probabilities. It is well-known (Bretagnolle and Huber, 1988) that in the case of survival regression models with censoring, if covariates relevant to survival are omitted from the model then the other parameter estimates are asymptotically biased towards zero, hence the effects of covariates included in the model are underestimated. What impact does the bias of the estimated covariate effects has on the predicted default probabilities? There are two cases to consider.

In the first case, assume that the model includes only firm-specific covariates. The predicted default probabilities will then be either underestimated or overestimated, depending on the sign of the value of each firm covariate and on the sign of the corresponding covariate effect. For example, if a covariate has an estimated negative effect $\beta < 0$ which is biased towards zero, and if that covariate has a positive value for the prediction year, then this will lead to an overestimated predicted default probability. Conversely, a covariate with a positive biased estimated effect and a positive value would lead to an underestimated predicted default probability. The overall impact of the biased estimates for all firm-specific covariates on the predicted default probability of the firm can thus be either positive or negative, and it will differ from firm to firm.

In the second case, assume that the model includes also macroeconomic covariates, in addition to firm specific covariates. Using the same argument as in the previous paragraph, a biased estimate of the effect of a macroeconomic covariate will essentially lead to either underestimated or overestimated default probabilities, depending on the signs of the covariate value and of its estimated effect parameter. However, unlike for firm-specific covariates, this impact will be *systematic across all firms*, since the macroeconomic covariate values are common to all obligors. As a consequence, *all* predicted default probabilities will be either underestimated or overestimated.

In the case of our analysis, default model MD3 contains both firm specific and macroeconomic covariates, and the results in Table 2 show that all the firm specific covariates are highly statistical significant, with the exception of net income to total assets. Omitting the two macroeconomic covariates term spread and credit spread from this specification, we obtain default model MD1. Since macroeconomic effects are significant for default, we expect that coefficient estimates for model MD1 are asymptotically biased towards zero. Indeed, Table 2 shows that all estimated covariate effects are smaller in absolute value for model MD1 than for model MD3, with the exception of the volatility effect. The impact that these changes have on predicted default probabilities, however, varies from firm to firm, since model MD1 only includes firm specific covariates and their signs can be both positive and negative.

Default model MD2 is obtained from model MD1 by replacing volatility and the two accounting variables with distance-to-default. Panel A in Table 5 gives the Pearson correlations between all firm specific covariates. Although distance-to-default is moderately correlated with the volatility and the two accounting variables, it can be argued that these three omitted covariates contain information relevant to default that is not entirely captured in the distance-to-default variable. Therefore, we expect that the estimated effects for the remaining common covariates are smaller in absolute value for model MD2 than for model MD1. This is indeed the case for both common covariates, relative size and excess return. Again, since these are firm level covariates, the impact that these changes have on predicted default probabilities varies from firm to firm.

Finally, default model MD4 contains both macroeconomic and firm specific covariates. Note that none of the firm specific covariates that are highly significant in models MD1 and MD3 are included in model MD4. They are instead replaced with the stock return and the distance-to-default, and Panel A in Table 5 shows that these two variables are moderately correlated with the other five firm specific covariates. It is possible that the stock return and the distance-to-default are able to capture the entire spectrum of firm specific information relevant for default, and that no further firm specific covariates are needed. To investigate if this is the case, we perform an exploratory factor analysis on all seven firm specific covariates. We find that there are three significant factors corresponding to the eigenvalues greater than one of the correlation matrix, and these factors together explain 77% of the total variance. Panel B in Table 5 gives the factor loadings for the seven variables; the first factor loads highly on the excess return and on the stock return, the second factor is strongly correlated with volatility, relative size, and distance-to-default, while the third factor loads highly on the accounting variable total liabilities to total assets.

These results suggest that two dimensions are not sufficient for capturing the range of firm-specific information relevant for default, and that at least three dimensions of firm-specific covariates may be necessary in a default model. In our case, default model MD4 relies on only two firm covariates. Following the arguments earlier in this subsection, this implies that the estimated covariate effects in MD4 will be biased towards zero. This bias does not lead to a lack of power since the coefficient estimates in MD4 are still highly statistically significant, but it does, however, have a substantial impact on the predicted default probabilities. In particular, the estimated effects of the T-bill yield (-0.046) and of the S&P 500 return (-1.075) in model MD4 are biased towards zero. Note that the values of the T-bill yield are always positive, and the values of the S&P 500 return are positive in six out of the nine years between 1996–2004 (the average for the S&P 500 return during the out-of-sample horizon is 0.114). Therefore, the bias towards zero in the estimated negative effects of these two covariates leads to an increase in the predicted default probabilities for *all* firms. As a result, the distribution of the number of defaults in the telecoms portfolio predicted by model MD4 in Figure 2 is substantially shifted to the right with respect to the distributions predicted by the other three models.

Recall that, as discussed in the previous subsection, all four default models have almost identical performance for our data set according to the power curve. They predict, however, very different distributions for the total number of defaults in the portfolio. When interest centers exclusively in predicting default as a dichotomous variable for each individual firm, the traditional metric of the power curve may be an adequate measure of model performance. When interest, however, lies in predicting quantities that involve the actual magnitudes of default probabilities rather than just their ranking (such as, for example, predicting the number of defaults in a portfolio), the power curve is no longer an adequate measure of model performance. This issue will be particularly relevant in our discussion of loss distributions in Section 5, since the magnitudes of default probabilities are a crucial component in any loss calculations.

4.2.3 Comparison of models with and without frailty

In this subsection we investigate whether the inclusion of a frailty term has any impact on the default models predictions. To this purpose we consider two measures of predictive performance discussed in the previous subsections, the power curve and the total number of portfolio defaults, and compare the predictions from pairs of models that have the same covariate specifications but differ in their inclusion of the frailty term.

For predictions of individual firm defaults, the results in Table 3 show that models with and without frailty are almost identical in their prediction performance, both at annual level and at aggregate level over the entire horizon. In addition, Figure 3 shows that the power curves for models with and without frailties are also virtually identical, for all four covariate specifications considered. Overall, the inclusion of a frailty term does not improve prediction performance at individual firm level, which is not surprising given the results in Section 2.1.3.

For predictions of the total number of defaults in a specific portfolio, however, some differentiation of models with and without frailty is possible. Figure 4 gives the distribution of the total number of defaults predicted out-of-sample during 2003 in the portfolio of telecommunication firms. The distributions are computed from each of the four default models, both with and without frailty. Figure 4 shows that in all cases the distributions generated by frailty models have a larger mean and in most cases also larger dispersion than those generated by models without frailty. This is to be expected: recall that this portfolio of telecommunication firms had a considerable number of defaults in the period just preceding the out-of-sample horizon. The frailty models account for this and estimate high frailty values for this industry group. The four estimates for the conditional expected frailty for the telecoms group during 2003 computed with expression (6) for each model, ranged between 1.20 and 2.03, compared with an unconditional mean of the frailty distribution of one (and an estimated variance θ of around 0.17). The models without frailty, on the other hand, do not account for the recent default history of this specific industry group, and as a consequence their predicted default probabilities are smaller than those computed from the frailty models. This explains the shift in the mean for all four cases shown in Figure 4. It is also to be expected that the distributions computed from frailty models have longer tails and thus larger variance than those computed from models without frailty. This is consistent with the frailty interpretation of accounting for the impact of unobservable heterogeneity and missing covariates at industry group level; the larger variance reflects the increased prediction uncertainty due to the effect of missing potentially relevant information. The difference between the distributions generated by models with and without frailty is largest for default model MD2, and smallest for default model MD4.

4.3 Recovery Rate Models

In this subsection we discuss three models for recovery rates that we subsequently use for the modelling of expected loss. The goal of this study is not to derive new predictive models for recovery rates or to investigate exhaustively the determinants of recoveries, since extant literature has already focused on this topic (Acharya et al, 2003 and 2007). The modelling of recovery rates in this paper is only a preliminary stage towards the ultimate objective of assessing the impact of default and recovery models on predicted loss.

Extant literature showed that contract characteristics, firm specific variables and macroeconomic variables are important factors that affect recovery rates. We experimented with many model specifications using the variables described in Section 3.2. The models that we retained during the model selection process have little or no redundant information in the form of covariates that are not statistically significant, since the presence of such covariates simply increases the model complexity without providing more explanatory or predictive power. We next briefly discuss the general insights from our analysis of many different covariate specifications, then we describe the three models that we retain for our subsequent analysis of predicted loss.

Among the contract characteristics only the coupon rate had a statistically significant coefficient in almost all models that we considered. In particular, the logarithm of the issue size and the logarithm of the maturity outstanding were consistently not significant, similar to results from Acharya et al. (2003). To model the seniority class of a bond, we use four seniority dummies and take the junior class as baseline. All four dummies (subordinate, senior subordinate, senior secured and senior unsecured) were generally statistically significant. Although the differences between their estimated coefficients are small, the ordering of the coefficients is as expected implying that senior secured bonds earn on average higher recoveries than senior unsecured bonds, which in turn have larger recoveries than senior subordinated bonds.

Among the firm characteristics, the relative size, the logarithm of total assets, the market to book ratio, and the ratio of tangible assets to total assets were statistically significant in most models that we considered. The effect of distance-to-default is generally significant as well, but it loses significance when other firm characteristics are also included in the model. To account for the industry of a firm, we use four industry dummies following Moody's classification: transportation, utilities, financials and industrials.[20] We take transportation

[20]The sample sizes in these industry groups are unequal — bonds of industrial firms account for around 82% of the recoveries in our sample.

as baseline, and find that the coefficient of the utilities dummy was generally significant and positive in most models that we considered, while the coefficients for financials and industrials dummies were not. Most industry dummies lose significance when macroeconomic variables are also included in the model.

Among the macroeconomic variables, the Treasury 3-month yield, the return on the S&P 500 index, and the logarithm of the amount of all defaulted debt are statistically significant in most models. The credit spread is never significant in any model that we considered, and the term spread is only marginally significant in a few models.

These insights are summarized in Table 6 which reports the estimates from fitting three selected recovery models among the multitude of models that we considered. For each model we estimate both the logit and probit specifications described in Section 2.2. The estimation results for the logit model are similar to the ones for the probit model, and the out-of-sample performance of the logit model is slightly superior to the performance of the probit model. Consequently we only report in Table 6 the results for the logit models.

Models MR1 and MR2 are based on all classes of variables including contract, firm, and macroeconomic factors, whereas model MR3 contains only contract and firm variables and lacks macroeconomic factors. Since models MR1 and MR2 share the contract characteristics and macroeconomic variables, the difference between them lies in the way in which firm specific information is taken into account; model MR1 uses distance to default as a single firm-specific covariate, while model MR2 contains the relative size, the logarithm of total assets, the market to book ratio, and the ratio of tangible assets to total assets. In the next section, the comparison between the impact of models MR1 and MR2 on predicted loss will enable us to test whether the way in which firm specific information is taken into account has an effect on predicted loss distributions.

Models MR2 and MR3 share the same contract and firm specific information, and the only difference between them lies in the fact that model MR2 also includes macroeconomic variables while model MR3 does not. In the next section, the comparison between the impact of models MR2 and MR3 on predicted loss will enable us to test whether the inclusion of macroeconomic information has an effect on predicted loss distributions. Note also from Table 6 that parameter estimates for shared covariates are similar in terms of statistical significance for all three models, and the estimated coefficients have the expected sign.

We are not aware of any studies that examine the out-of-sample prediction performance of recovery rate models. We assess the out-of-sample performance of models MR1, MR2 and MR3 by using a rolling horizon calibration method. Similar to the methodology for default prediction in Section 4.2, we define the out-of-sample horizon to be 1996–2004. For each year t during this horizon, we compute the parameter estimates for all recovery models from data between 1980 and $t - 1$, then we forecast recovery rates during year t for all outstanding bonds of each firm that is alive at the beginning of that year. These forecasts are then compared with the actual realized recovery rates for bonds defaulted during year t. The average root mean square error of prediction (RMSE) for all three recovery models is reported in the last row of Table 6. The out-of-sample RMSE values are very similar, virtually identical, across all three recovery rate models.

5 Loss Distributions

In this section we first investigate the correlation between the default probabilities and the recovery rates predicted out-of-sample by our models. Next, we describe the methodology for computing the expected loss and we investigate the out-of-sample loss distributions obtained under different default and recovery models.

5.1 Default and Recovery Correlation

Empirical evidence shows that ex-post the frequency of default and the recovery rate given default are negatively correlated (Altman et al., 2005). In this subsection we investigate whether this empirical relationship between actual *realized* default frequencies and average recoveries is also apparent in the *predicted* default probabilities and recovery rates computed out-of-sample with our models. Note that the empirical relationship is based on aggregate data, since it involves frequencies of multiple firm defaults. Our study, however, investigates the correlation at individual firm level, as our methodology allows the computation of individual firm default probabilities.

Table 7 summarizes the aggregate correlations between predicted out-of-sample default probabilities and recovery rates, for each pair of frailty default and recovery models. Similar to the methodology for out-of-sample testing, at the beginning of each year t between 1996–2004, we first estimate the coefficients of all recovery and default models using the data between year 1980 and year $t - 1$. Next, we predict out-of-sample the default probability during year t for each firm alive and with bonds outstanding at the beginning of year t. Then we predict out-of-sample the recovery rate during year t for each bond; for firms with multiple bonds of the same seniority, we take the average of the predicted recovery rates within the same seniority class. Finally, we stratify the firms by industry and the bonds by seniority, and within each industry and seniority class we compute the bivariate Pearson correlation coefficient between the default probability and the recovery rate across all firms and all years. These coefficients are reported in Table 7.

All the aggregate correlations are negative and highly statistically significant, and they vary with industry group and seniority level. Across almost all seniority levels, the correlations are larger in absolute value for firms in the utility sector than for firms in the industrial sector, which are in turn higher than for firms in the financial sector. The aggregate correlations also vary with the choice of default and recovery models. For any given choice of recovery model, the correlations obtained with default models MD1 and MD3 are similar, generally smaller in absolute value than those obtained with default model MD2, which in turn are usually smaller than those obtained with default model MD4.

Are the patterns detected in aggregate correlations over the entire horizon 1996–2004 preserved for disaggregate annual correlations? Figure 5 gives the annual correlations between out-of-sample predicted default probabilities and recovery rates for each pair of frailty default models MD1 and MD4 and recovery models MR1, MR2 and MR3. The firms are all industrials, and the recovery rates are all for senior unsecured bonds. We focus on this seniority class because it has the largest number of recoveries in our data set, hence the correlation results are less likely to be affected by small sample size effects.

It is apparent from Figure 5 that the annual variation of correlations over the forecasting horizon is related to the credit cycle. The correlations increase in absolute value as the state of the economy worsens in 2001, then decrease as the economy improves during 2003–2004. This is consistent with insights from Das and Hanouna (2007) who also find that the correlations of default probabilities and recovery rates become increasingly negative with increasing default risk in the economy. The general level of the correlations in Figure 5 is also of the same magnitude with the values found by Das and Hanouna (2007).

The patterns of annual variation of correlations in Figure 5 differ considerably between the correlations computed with model MR1 and those computed with models MR2 and MR3. For any specific default model, the correlations implied by models MR2 and MR3 are very similar and generally smaller in absolute value than those implied by model MR1. The main difference between model MR1 and models MR2 and MR3 is in the way in which firm-specific information is taken into account. Figure 5 thus shows that the modeling of firm-specific information in the recovery models has a significant impact on both the level and the annual variation of the correlation values over the forecasting horizon. The only difference between models MR2 and MR3 lies in the fact that MR2 includes macroeconomic variables, while MR3 does not. Since recovery models MR2 and MR3 lead to similar levels of correlation, this implies that the inclusion of macroeconomic factors in the recovery models does not have an impact on either the level or the annual variation of correlation.

Figure 5 also shows that for any specific recovery model the correlations implied by default model MD4 are always larger in absolute value than those implied by default model MD1, although the difference is generally marginal. These results are consistent with the patterns in aggregate correlations reported in Table 7.

5.2 Predicted loss distributions

In this section we investigate the impact that the choice of default and recovery models has on predicted loss distributions for a portfolio of obligors. We assume that the face value of each bond in the portfolio is one. With the notation from previous sections, let us denote by $\perp_i(t)$ an indicator function that equals 1 if firm i defaults in period t conditional on survival up to period t, and 0 if default does not occur in period t. Let us also denote by $R_i(t)$ the recovery rate for bonds of firm i at time t. The loss $L_i(t)$ from obligor i in period t is then

$$L_i(t) = \begin{cases} 1 - R_i(t) & \text{if } \perp_i(t) = 1 \\ 0 & \text{if } \perp_i(t) = 0. \end{cases}$$

The expected loss over the next one period horizon[21] computed at time t is then

$$E_t[L_i(t+1)] = E_t[\perp_i(t+1) \cdot (1 - R_i(t+1))],$$

where both the default indicator $\perp_i(t+1)$ and the recovery rate $R_i(t+1)$ depend on a set of covariates $X_i(t+1)$.

We exemplify this methodology through an application to predicting loss distributions in specific portfolios from our sample data. We first focus on the portfolio of firms with DNUM codes between 4800 and 4899 (telecommunication firms), for which we predicted out-of-sample the total number of defaults during 2003 in Section 4.2.2. Three firms defaulted in this portfolio during 2003. Several of these firms had multiple bonds outstanding, leading to four defaulted bonds and a loss of 2.74 (under the assumption that the face value of each bond is one). Using the predicted out-of-sample default probabilities and recovery rates for all firms in this portfolio and for all their bonds, we generated the out-of-sample loss

[21] Note that using this methodology it is possible to compute the expected loss over multiple periods horizons as well (Chava et al. 2006). This entails, however, modelling the evolution of the macroeconomic and firm specific covariates, for example, by means of a stochastic process such as an autoregressive time series.

distributions for each combination of default and recovery models, based on 10000 simulated loss scenarios. Figure 6 gives the probability density functions of the loss distributions during 2003 in this portfolio, predicted out-of-sample during 2003 by all four frailty default models and all three recovery models.

For any of the three recovery models, the loss distributions generated with default models MD2 and MD3 are quite similiar. The distribution generated with default model MD1 has thinner tails, but comparable expectation and mode. Default model MD4, however, leads to entirely different results. As we noted in Section 4.2.2, model MD4 predicts higher default probabilities out-of-sample than the other default models. This has a strong impact on the predicted loss distribution, which has both higher expectation and higher variance than the loss distributions predicted with the other three default models. In particular, the actual realized value 2.74 of the loss in this portfolio falls far in the left tail of the distribution generated by MD4.

For any of the four default models, the loss distributions generated with recovery models MR2 and MR3 are very similar, while the distribution generated with MR1 is only slightly different. Since the difference between models MR1 and either MR2 or MR3 lies in the way in which firm specific information is taken into account, it follows that, as long as firm specific information is accounted for, the inclusion of any particular components of this information in the recovery models does not have a major impact on predicted loss. Similarly, since the only difference between models MR2 and MR3 lies in the fact that MR2 also includes macroeconomic variables while MR3 does not, it follows that the inclusion of macroeconomic information *in the recovery models* does not make a large difference on predicted loss. Notice that Figure 6 implies that the choice of default model has much more impact on predicted loss than the choice of recovery model.

Table 8 gives the expectation, median, and three different percentiles of the loss distributions from Figure 6. The percentiles correspond to different thresholds for the value-at-risk. Note that the expected loss values generated by default model MD1 are the closest to the actual realized loss of 2.74. Consistent with the insights from Figure 6, the expected loss and values-at-risk generated with recovery models MR2 and MR3 are virtually identical for all default models. The numbers in Table 8 are also consistent with Figure 2 which shows that MD2 and MD3 generate very similar loss distributions, while MD4 generates a loss distribution with longer tails and higher expectation.

We next investigate the predicted loss distributions for bonds stratified by seniority level. We consider all firms with debt outstanding across all industries and focus on bonds in the subordinate, senior subordinate, and senior unsecured class. We choose these seniority classes because each of them had a sufficiently large sample of defaulted bonds to lead to interesting insights. During the out-of-sample year 2003, the sample contains defaults of 3 subordinate bonds, 9 senior subordinate bonds, and 35 senior unsecured bonds. Assuming a face value for each bond of one, the actual realized loss during this period was 2.69 for subordinate bonds, 7.36 for senior subordinate bonds, and 21.35 for senior unsecured bonds.

Using the predicted out-of-sample default probabilities and recovery rates for all firms and for all their bonds in these three classes, we generated the out-of-sample loss distributions during 2003 for each combination of default models MD1 and MD4 and recovery models MR1 and MR2, based on 10000 simulated loss scenarios.[22] In each simulation scenario, a firm defaults if the firm's predicted default probability is greater than a randomly generated variable from the uniform distribution on $[0, 1]$. When a firm defaults, we assume that all its outstanding bonds default and we compute the loss using the predicted recovery rate for each bond. The overall loss in a simulation scenario is the loss from all outstanding bonds of all firms defaulting in that scenario. Figure 7 gives the predicted out-of-sample probability density functions of the loss distributions for bonds of different seniority levels.

As in the previous example of the portfolio of telecommunication firms, the choice of default model seems to have a crucial impact, as the comparison between the plots from MD1 on the left side and the plots from MD4 on the right side of Figure 7 shows. For all three seniority classes, the loss distributions predicted by MD1 have smaller expectations and variances than the distributions predicted by MD4. The actual realized losses in each seniority class are close to the expectation of the corresponding loss distributions predicted by MD1, but fall quite far in the left tail of the loss distributions predicted by MD4. Consistent with the insights from the analysis of the telecommunications portfolio, there are relatively small differences between the distributions generated by the two recovery models MR1 and MR2 for any given default model. This holds for all seniority classes and it confirms that the choice of recovery model does not have a major impact on predicted loss.

6 Summary

This paper addresses the issue of modeling and validating the loss distribution in the presence of unobservable heterogeneity in firm characteristics. To generate the loss distribution, it

[22]Since the example of the telecommunication portfolio showed that the distributions generated by recovery models MR2 and MR3 are almost identical, for clarity purposes we do not report further the plots generated by model MR3. Similarly, among default models we restrict attention to MD1 and MD4, since models MD2 and MD3 generate similar distributions to MD1.

is necessary to model the probability of default and the recovery rate given default. In this paper we focus on two main issues — how to determine the appropriate combination of default and recovery models to be used for generating the loss distribution, and how to validate the methodology through out-of-sample testing.

For out-of-sample testing, we first focus on the standard performance metrics for default prediction: the ability to correctly identify defaulted firms (corresponding to Type I errors), and the power curve (corresponding to Type I and Type II errors). We show that four default models inspired by extant literature have very similar performance according to the standard metrics, based on the analysis of a large default data set over the horizon 1980–2004. We also test whether the frailty component of the hazard function increases the ability to predict defaults. We find that accounting for frailty significantly increases the log-likelihood function, but that the frailty has virtually no impact on out-of-sample prediction performance as measured by the standard performance metrics. We argue that the standard metrics cannot assess whether a model generates default probabilities that are adequate in absolute terms, because they are only based on the relative ordinal rankings of default probabilities and not on the actual magnitudes of these probabilities. We present a different approach for assessing default prediction accuracy, by comparing the actual realized number of defaults in a given portfolio with the total number of defaults in that portfolio predicted out-of-sample. We find that a more complex picture emerges when using this approach, and that differences between total numbers of defaults predicted by the four default models can be substantial. The appropriate choice of default model will subsequently have a crucial impact on loss prediction, since the loss distribution depends on the actual magnitudes of default probabilities rather than on the relative firm riskiness ranking.

For the purpose of predicting loss in a given portfolio, we also address the issue of modeling recovery rates. We use three different specifications for recovery in the event of default and find that all three recovery models have similar out-of-sample performance. We show that the default probabilities and recovery rates predicted out-of-sample are negatively correlated, and that the magnitude of the correlation is related to the credit cycle and varies with industry and seniority class. The level of the correlation depends crucially on the choice of default model, and to a smaller extent on the choice of recovery model. In particular, the annual variation of the correlation does not depend on whether macroeconomic variables are taken into account in the recovery models, and it is only marginally affected by the particular choice of obligor specific variables included in recovery models.

Finally, we investigate the impact that the choice of default and recovery models has on the predicted loss distribution. We show that the default model specification significantly affects the predicted loss distribution. In particular, the Shumway (2001) default model predicts loss distributions where the expected loss is closest to the actual realized loss in the portfolio, whereas the Duffie et al. (2007) default model leads to consistent overestimation of the realized loss. We find that the recovery model specification has a smaller impact than the default model specification on the predicted loss distribution; this impact does not depend on the inclusion of macroeconomic variables, and it is again only marginally affected by the particular choice of obligor specific variables in the recovery model.

Appendix A: Proof of Propositions

Proof of Proposition 1. Using equation (9), under the hazard rate model without frailty the expected number of defaults is

$$E_{nf}[D] = \sum_{i=1}^{n} E_{nf}(\perp_i) = \sum_{i=1}^{n} PD_{nf,i} = \sum_{i=1}^{n} [1 - \exp(-\alpha_i)].$$

Using the law of iterated expectations and equations (10)–(11), we obtain

$$\begin{aligned} E_f[D] &= \sum_{i=1}^{n} E_f(\perp_i) = \sum_{i=1}^{n} E_Y[E_f(\perp_i \mid Y)] \\ &= \sum_{i=1}^{n} E_Y[PD_{f,i|Y}] = \sum_{i=1}^{n} E_Y[1 - \exp(-Y\alpha_i)] \\ &= \sum_{i=1}^{n} [1 - L(\alpha_i)], \end{aligned}$$

which completes the proof.[23] ∎

Proof of Proposition 2. Since $\alpha_i = \exp(\beta X_i) > 0$ for all i, it follows that $\exp(-Y\alpha_i)$ is a strictly convex function in Y. Then from Jensen's inequality and the fact that $E[Y] \leq 1$, it follows that

$$L(\alpha_i) = E[\exp(-Y\alpha_i)] > \exp[-E[Y]\alpha_i] \geq \exp(-\alpha_i). \tag{14}$$

Replacing (14) in expressions (12) and (13) and aggregating over all firms i, we obtain that $E_f[D] \leq E_{nf}[D]$. ∎

[23] It can also be shown that the variance of the number of defaults predicted by the hazard rate model (7) with frailty is given by

$$Var_f[D] = \sum_{i=1}^{n} L(\alpha_i) + \sum_{i,j=1, i \neq j}^{n} L(\alpha_i + \alpha_j) - \left[\sum_{i=1}^{n} L(\alpha_i) \right]^2,$$

and that the variance of the number of defaults predicted by the hazard rate model (8) without frailty is

$$Var_{nf}[D] = \sum_{i=1}^{n} \exp(-\alpha_i)[1 - \exp(-\alpha_i)].$$

Appendix B: Construction of the Distance-to-Default Variable

We closely follow Bharath and Shumway (2007) for the construction of the Distance to Default measure. This variable is traditionally computed based on the Merton (1974) model, under which the firm value is assumed to follow a geometric brownian motion

$$\frac{dV}{V} = \mu\, dt + \sigma_V\, dW,$$

where V is the total value of the firm, μ is the expected continuously compounded return on V, σ_V is the volatility of firm value and dW is a standard Weiner process. Under the assumptions of the Merton model, the equity of the firm is a call option on the underlying value of the firm with a strike price equal to the face value of the firm's debt and a time-to-maturity of T. Based on the Black-Scholes formula, the market value of the firm's equity denoted by E is given by

$$E = V\mathcal{N}(d_1) - e^{-rT}F\Phi(d_2),$$

where F is the face value of the firm's debt, r is the instantaneous risk-free rate, $\Phi(\cdot)$ is the cumulative standard normal distribution function, and

$$d_1 = \frac{\log(V/F) + (r + \sigma_V^2/2)T}{\sigma_V\sqrt{T}}, \quad d_2 = d_1 - \sigma_V\sqrt{T}.$$

Using an application of Ito's lemma and the fact that $\dfrac{\partial E}{\partial V} = \Phi(d_1)$, it follows that the second equation in this model links the volatility of the firm value and the volatility of the equity:

$$\sigma_E = \frac{V}{E}\Phi(d_1)\sigma_V.$$

The unknowns in these two equations are the firm value V and the asset volatility σ_V. The known quantities are the equity value E, the face value of debt or the default boundary F, the risk-free interest rate r, and the time to maturity T. Since we have two equations and two unknowns, we can solve for V and σ_V directly. The distance-to-default DD is then defined as

$$DD \equiv \frac{\log(V/F) + (\mu - \sigma_V^2/2)T}{\sigma_V\sqrt{T}}.$$

The book debt variable is defined as the sum of short-term debt and long-term debt taken from the COMPUSTAT files. Following Campbell, Hilscher and Szilagy (2006), we substitute the missing or zero values of book debt when computing the distance-to-default. If the value of book debt is missing, we substitute it with the value of (median of book debt /total liabilities)* total liabilities, where the median is taken over the entire data set. If the value of book debt is zero, we also substitute it with the value of (median of book debt /total liabilities)*total liabilities, however this time the median is taken over the values of book debt in the range (0, 0.01). The book debt variable is winsorized at 0.5% and 99.5% in the construction of the distance-to-default measure.

References

[1] Acharya, V. V., S. T. Bharath, and A. Srinivasan (2003), Understanding the recovery rates on defaulted securities, Working paper, London Business School.

[2] Acharya, V. V., S. T. Bharath, and A. Srinivasan (2007), Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries, *Journal of Financial Economics* 85, 787–821.

[3] Altman, E. I., and E. Hotchkiss (2005), *Corporate Financial Distress and Bankruptcy*, Third Edition, Wiley, New York.

[4] Altman, E. I., B. Brady, A. Resti, and A. Sironi (2005), The link between default and recovery rates: theory, empirical evidence and implications, *Journal of Business* 78, 2203–2227.

[5] Andersen, L., and J. Sidenius (2005), Extensions to the Gaussian copula: random recovery and random factor loadings, *Journal of Credit Risk* 1, 29–70.

[6] Basel Committee on Banking Supervision (2005a), *Guidance on Paragraph 468 of the Framework Document*.

[7] Basel Committee on Banking Supervision (2005b), *Studies on the Validation of the Internal Rating Systems*, Working paper N0. 14.

[8] Bharath, S. T., and T. Shumway (2007), Forecasting default with the Merton distance-to-default model, *Review of Financial Studies*, forthcoming.

[9] Bretagnolle, J., and C. Huber (1988), Effects of omitting covariates in Cox's model for survival data, *Scandinavian Journal of Statistics* 15, 125–138.

[10] Campbell, J. Y., J. Hilscher, and J. Szilagyi (2006), In search of distress risk, NBER Working paper 12362.

[11] Chava, S., and R. A. Jarrow (2004), Bankruptcy prediction with industry effects, *Review of Finance* 8, 537–569.

[12] Chava, S., Stefanescu, C., and S. Turnbull (2006), Modeling expected loss, Working Paper, Bauer College of Business.

[13] Covitz, D., and S. Han (2004), An empirical analysis of bond recovery rates: exploring a structural view of default, Working paper, The Federal Reserve Board of Washington.

[14] Crowder, M. (2001), *Classical Competing Risks*, Chapman & Hall, New York.

[15] Das, S. and P. Hanouna (2007), Implied recovery, Working paper, Santa Clara University.

[16] Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society B* 39, 1–38.

[17] Duffie, D., Eckner, A., Horel, G., and L. Saita (2006), Frailty correlated default, Working paper, Stanford University.

[18] Duffie, D., L. Saita, and K. Wang (2007), Multi–period corporate failure prediction with stochastic covariates, *Journal of Financial Economics* 83, 635–665.

[19] Dullmann, K., and M. Trapp (2004), Systematic risk in recovery rates — an empirical analysis of U.S. corporate credit exposure, Working paper, Deutsche Bundesbank, Frankfurt, Germany.

[20] Frye, J. (2000), Depressing recoveries, *Risk Magazine* 13, 108–111.

[21] Gagliardini, P., and C. Gourieroux (2003), Spread term structure and default correlation, Working paper, University of Toronto.

[22] Hougaard, P. (2000), *Analysis of Multivariate Survival Data*, Springer, New York.

[23] Kiefer, N. M. (1988), Economic duration data and hazard functions, *Journal of Economic Literature* Vol. XXV1, 646–679.

[24] Klein, J. P., and M. Moeschberger (1997), *Survival Analysis*, Springer, New York.

[25] Lawless, J. L. (2003), *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New Jersey.

[26] Mamaysky, H., Spiegel, M., and H. Zhang (2007), Improved forecasting of mutual fund alphas and betas, *Review of Finance* 11, 359-400.

[27] Merton, R. C. (1974), On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

[28] Pykhtin, M. (2003), Unexpected recovery risk, *Risk* 16, 74–78.

[29] Schönbucher, P. J. (2003a), Information driven default contagion, ETH Working paper.

[30] Schönbucher, P. J. (2003b), *Credit Derivatives Pricing Models*, John Wiley & Sons Ltd. New Jersey.

[31] Schuermann, T. (2004), What do we know about loss given default. In *Credit Risk Models and Management*, Risk Books, London.

[32] Shumway, T. (2001), Forecasting bankruptcy more accurately: a simple hazard model, *Journal of Business* 74, 101–124.

[33] Stefanescu, C., and B. W. Turnbull (2006), Multivariate frailty models for exchangeable survival data with covariates, *Technometrics* 48, 411–418.

[34] Varma, P., and R. Cantor (2005), Determinants of recovery rates on defaulted bonds and loans for North American corporate issuers: 1983–2003, *Journal of Fixed Income* 14, 29–44.

Figure 1: Power curves for frailty default models

The plot gives the power curves for all four frailty default models, classifying defaulted firms over the out-of-sample horizon 1996–2004. The power curves illustrate both Type I and Type II classification errors, and are obtained by plotting on the vertical axis the cumulative percentage of defaulted firms versus the rank percentile of firms on the horizontal axis.

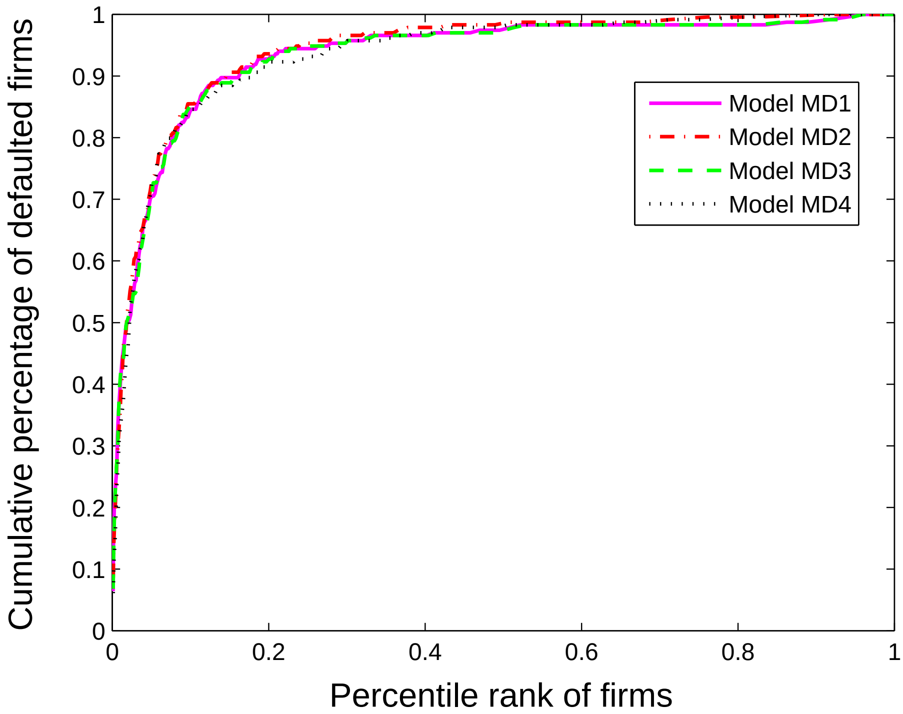

Table 1: **Descriptive Statistics**

The following table presents the descriptive statistics for the variables used in the default and recovery models. The sample contains firms in the intersection of CRSP, COMPUSTAT and Moody's Default Recovery Database during 1980–2004. The unit of measurement is one year for the macroeconomic variables, one firm-year for the firm specific variables, and one bond issue for the bond specific variables. For the firm specific variables used only in the recovery models (logarithm of total assets, market to book ratio, and tangible assets to total assets ratio), the unit of measurement is the year in which recovery was recorded, and the descriptive statistics are thus based on a sample size of around 800 observations. The term spread is the difference of the ten year Treasury yield and the one year Treasury yield, the credit spread is the difference between AAA and BAA yields. The excess return is defined as the return on the firm minus the value-weighted CRSP NYSE/AMEX/NASDAQ index return. The relative size of the firm is defined as the logarithm of each firm's equity value divided by the total NYSE/AMEX/NASDAQ market equity value. The volatility represents the idiosyncratic standard deviation of monthly stock returns of the previous year. The distance-to-default is constructed as in Bharath and Shumway (2007). Tangible assets to total assets denotes the ratio of property plant and equipment to the total assets of the firm. The recovery rate on a bond is measured as the bond price within a month after default, as given by Moody's DRS.

	Mean	25^{th} pctl	50^{th} pctl	75^{th} pctl	Standard deviation
Macroeconomic variables					
Term spread	1.233	0.595	1.220	2.099	1.055
Credit spread	1.094	0.726	1.033	1.356	0.440
T-bill 3-month yield	5.997	3.839	5.490	7.803	3.168
S&P 500 return	0.147	0.033	0.187	0.296	0.163
Logarithm(total defaulted debt)	1.705	0.613	1.837	2.921	1.825
Firm specific variables					
Excess return	0.042	-0.235	-0.009	0.230	0.469
Relative size	-9.089	-10.262	-9.088	-7.845	1.757
Volatility	0.108	0.068	0.095	0.133	0.057
Net income to total assets	0.034	0.009	0.038	0.068	0.071
Total liabilities to total assets	0.630	0.499	0.622	0.762	0.205
Distance-to-default	5.900	2.644	5.106	8.172	4.811
Stock return	0.188	-0.101	0.126	0.390	0.487
Logarithm(total assets)	7.407	6.153	7.464	8.674	1.666
Tangible assets to total assets ratio	0.407	0.228	0.380	0.588	0.240
Market to book ratio	1.333	0.967	1.111	1.372	0.962
Bond specific variables					
Recovery rate	0.337	0.134	0.269	0.488	0.251
Coupon rate	9.732	7.875	9.875	11.875	3.176

Table 2: **Frailty Default Models: Estimation Results**

The following table presents the estimates for the frailty default prediction models with exponential hazards, where models MD1 to MD4 differ in the specification of covariates. Standard errors are given in parantheses, and variable definitions are given in Section 3.2. The last row of the table gives the number N of firms available for estimation for each model.

	MD1	MD2	MD3	MD4
Frailty variance	0.164	0.172	0.161	0.193
	(0.071)	(0.075)	(0.070)	(0.075)
Intercept	-10.654	-6.824	-11.253	-3.235
	(0.441)	(0.500)	(0.492)	(0.159)
Excess return	-2.080	-1.704	-2.165	
	(0.185)	(0.198)	(0.192)	
Relative size	-0.326	-0.259	-0.339	
	(0.044)	(0.043)	(0.044)	
Volatility	2.063		1.948	
	(0.228)		(0.234)	
Net income to total assets	-0.086		-0.115	
	(0.563)		(0.561)	
Total liabilities to total assets	1.655		1.699	
	(0.267)		(0.266)	
Distance-to-default		-0.403		-0.487
		(0.031)		(0.027)
Stock return				-1.712
				(0.193)
Term spread			0.117	
			(0.054)	
Credit spread			0.322	
			(0.160)	
T-bill 3-month yield				-0.046
				(0.022)
S&P 500 return				-1.075
				(0.279)
Log likelihood	-1266	-1157	-1265	-1163
N	2985	3009	3006	3009

Table 3: **Default Forecasting: Percentages of Correctly Classified Defaults**

This table summarizes the out-of-sample forecasting accuracy for the four default model specifications, and reports percentages and counts of correctly classified defaults. The out-of-sample period is 1996–2004, and model calibration is performed on a rolling horizon. For each year t during the 1996–2004 horizon, the parameter estimates are computed from data between year 1980 and year $t-1$, then probabilities of default during year t are forecasted for each firm. The firms are then ranked into deciles each year according to their forecasted probability of default, with the first decile having the largest probabilities of default. The first part of the table presents the annual percentages and counts of defaults classified in the top two deciles (out of the total number of defaults) for each model and for each year between 1996–2004. Note that in some years there are small differences in the total number of defaults between models — this happens because slightly different data sets were used for model calibration due to missing data in some of the variables. The second part of the table presents for each model the percentages and counts of defaults classified in the top two deciles (out of the total number of defaults), aggregated over the entire horizon 1996–2004.

	Frailty default models				No-frailty default models			
	MD1	MD2	MD3	MD4	MD1	MD2	MD3	MD4
Annual results								
1996	87%	78%	87%	78%	87%	78%	87%	78%
	7/8	7/9	7/8	7/9	7/8	7/9	7/8	7/9
1997	100%	83%	100%	83%	100%	83%	100%	83%
	7/7	5/6	7/7	5/6	7/7	5/6	7/7	5/6
1998	97%	97%	97%	93%	97%	97%	97%	93%
	28/29	28/29	28/29	28/30	28/29	28/29	28/29	28/30
1999	91%	91%	91%	91%	91%	91%	91%	91%
	30/33	31/34	30/33	31/34	30/33	31/34	30/33	31/34
2000	96%	92%	96%	96%	96%	96%	96%	96%
	25/26	24/26	25/26	25/26	25/26	25/26	25/26	25/26
2001	96%	93%	96%	93%	98%	93%	98%	93%
	51/53	51/55	51/53	51/55	52/53	51/55	52/53	51/55
2002	90%	92%	90%	94%	90%	92%	90%	92%
	44/49	45/49	44/49	46/49	44/49	45/49	44/49	45/49
2003	100%	100%	100%	95%	100%	100%	100%	90%
	20/20	20/20	20/20	19/20	20/20	20/20	20/20	18/20
2004	80%	80%	80%	80%	80%	80%	80%	80%
	8/10	8/10	8/10	8/10	8/10	8/10	8/10	8/10
Aggregated over 1996–2004								
	93.62%	92.02%	93.62%	92.05%	94.04%	92.44%	94.04%	91.21%
	220/235	219/238	220/235	220/239	221/235	220/238	221/235	218/239

Table 4: **Default Forecasting: Common Ratios Across Default Models**

This table illustrates the out-of-sample forecasting agreement between the four default models, using the metric of common ratios. The out-of-sample period is 1996–2004, and model calibration is performed on a rolling horizon. For each year t during the 1996–2004 horizon, the parameter estimates are computed from data between year 1980 and year $t-1$, then probabilities of default during year t are forecasted for each firm. The firms are then ranked into deciles each year according to their forecasted probability of default, with the first decile having the largest probabilities of default. For any two default models MD_i and MD_j and for any given decile, the common ratio $\eta_{i,j}$ is defined as the number of firms classified in that decile by *both* models MD_i and MD_j divided by the number of firms classified in that decile by *either* of the models MD_i and MD_j, over the entire out-of-sample horizon 1996–2004. The common ratio is thus a measure of ranking agreement among the models. In general, the maximum value for the common ratio in any given decile is one (indicating complete agreement between models), and the minimum value is zero (indicating complete disagreement between models).

Decile	$\eta_{1,2}$	$\eta_{1,3}$	$\eta_{1,4}$	$\eta_{2,3}$	$\eta_{2,4}$	$\eta_{3,4}$
1	0.65	0.99	0.58	0.65	0.78	0.58
2	0.35	0.96	0.26	0.35	0.52	0.26
3	0.25	0.94	0.19	0.26	0.42	0.19
4	0.19	0.91	0.14	0.19	0.35	0.14
5	0.17	0.88	0.13	0.17	0.31	0.13
6	0.15	0.83	0.12	0.16	0.32	0.12
7	0.15	0.79	0.13	0.15	0.34	0.12
8	0.18	0.79	0.14	0.18	0.32	0.14
9	0.19	0.78	0.16	0.20	0.37	0.17
10	0.34	0.88	0.29	0.35	0.64	0.31
1 and 2	0.48	0.98	0.40	0.49	0.64	0.41

Figure 2: Distribution of Number of Defaults for Frailty Default Models

The plot gives the probability density functions of the distributions of the number of defaults during 2003 in the portfolio of firms with DNUM codes between 4800 and 4899 (telecommunication firms), predicted out-of-sample by the four frailty default models. The actual realized number of defaults in this portfolio during this period is three.

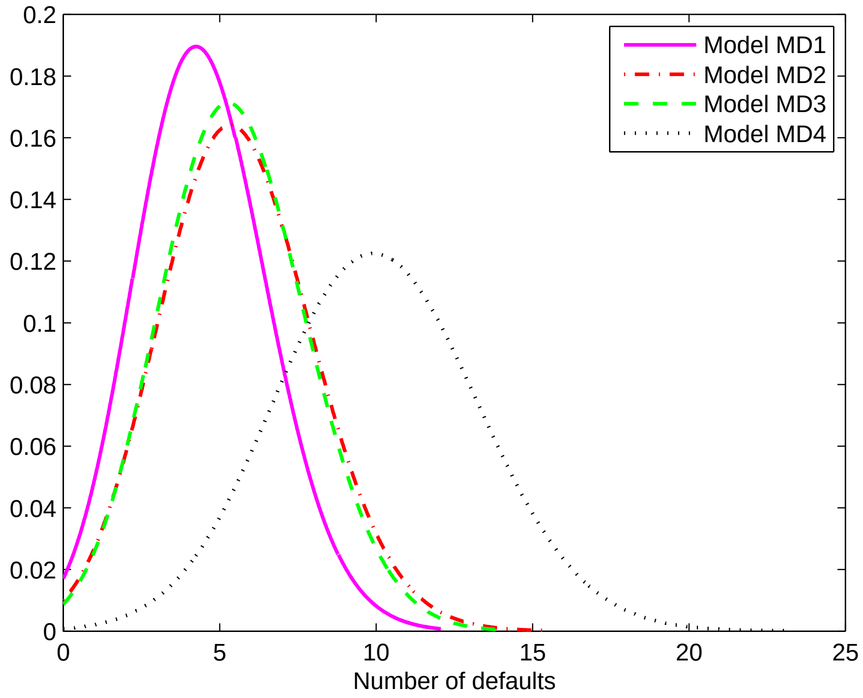

Table 5: **Correlations and Factor Loadings for Firm Specific Variables**

Panel A: Correlations

This panel reports the bivariate Pearson correlation coefficients for the seven firm specific variables.

	Excess return	Volatility	Relative size	Net income to total assets	Total liabilities to total assets	Distance to default	Stock return
Excess return	1.000						
Volatility	.082	1.000					
Relative size	.139	-.417	1.000				
Net income to total assets	.019	-.293	.299	1.000			
Total liabilities to total assets	.016	.035	-.052	-.367	1.000		
Distance-to-default	.226	-.400	.414	.379	-.447	1.000	
Stock Return	.947	.033	.145	.032	.007	.272	1.000

Panel B: Factor Loadings

This panel presents the factor loadings for the seven firm specific variables, resulted from an exploratory factor analysis. The analysis extracted three significant factors corresponding to the eigenvalues greater than one of the correlation matrix, and accounting for 77% of the total variance.

	Factor 1	Factor 2	Factor 3
Excess return	.977	.023	-.007
Volatility	.092	-.694	.009
Relative size	.116	.634	-.039
Net income to total assets	.000	.416	-.361
Total liabilities to total assets	.025	-.042	.938
Distance-to-default	.222	.584	-.447
Stock return	.967	.071	-.017

Figure 3: Power Curves for Default Models With and Without Frailty

The plots give power curves for each of the four default models both with and without frailty. The models classify defaulted firms over the out-of-sample horizon 1996–2004. The out-of-sample power curves account for both Type I and Type II classification errors. They are obtained by plotting the cumulative percentage of defaulted firms on the vertical axis versus the rank percentile of firms on the horizontal axis.

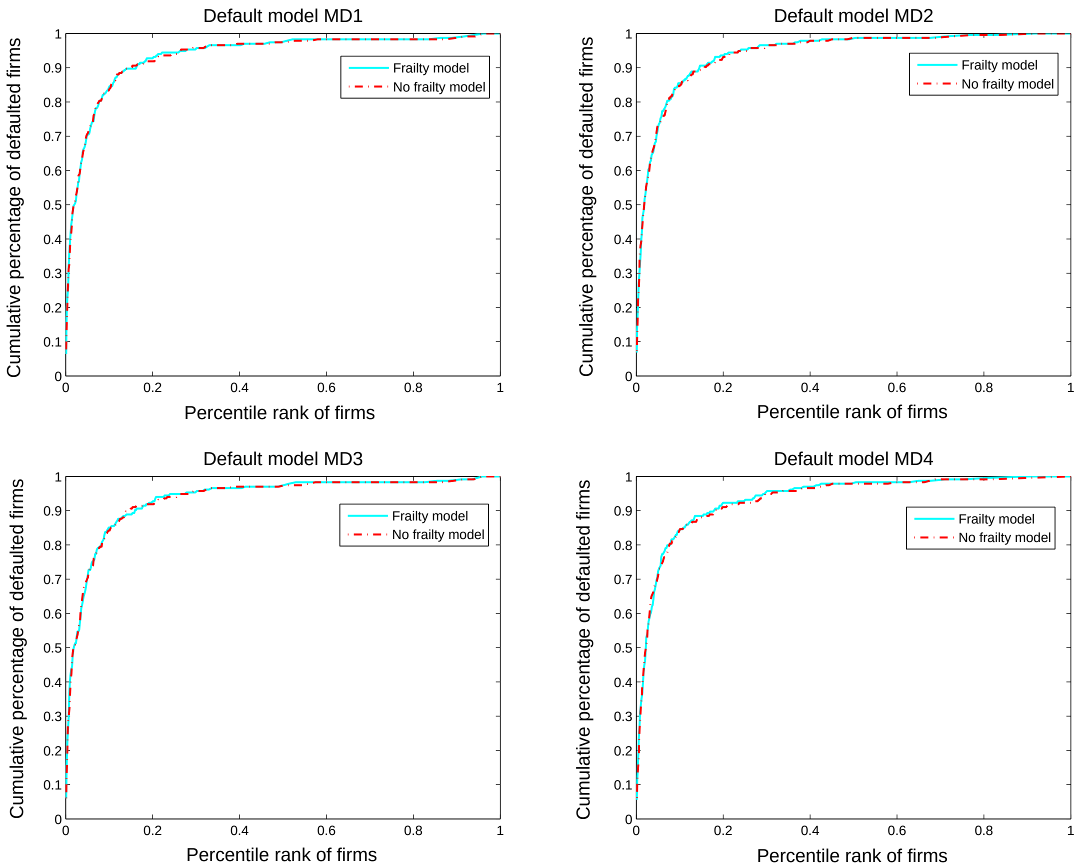

Figure 4: Distribution of Number of Defaults for Models With and Without Frailty

The plots give the probability density functions of the number of defaults during 2003 in the portfolio of firms with DNUM codes between 4800 and 4899 (telecommunication firms), predicted out-of-sample by the four default models, both with and without frailty. The actual realized number of defaults in this portfolio during 2003 is three.

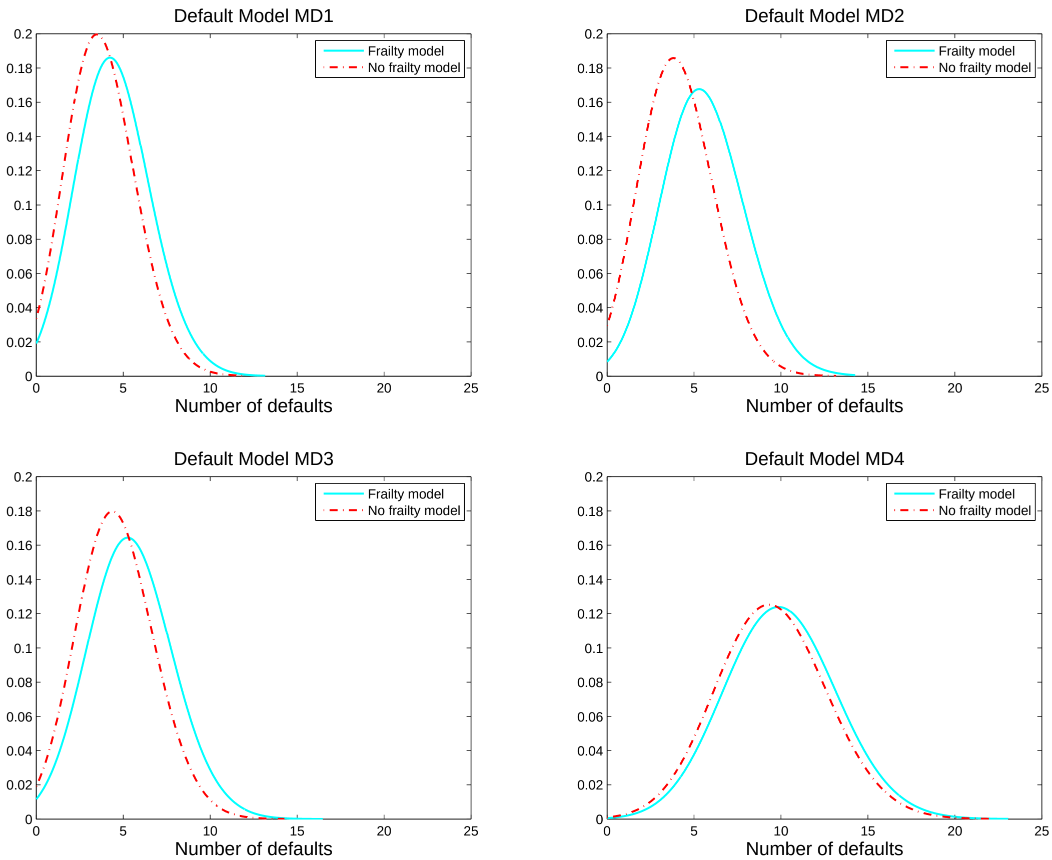

Table 6: **Recovery Rate Models: Estimation Results**

This table reports coefficient estimates from the logit regression relating the recovery rate to the bond, firm, industry and macroeconomic variables during 1980–2004. The dependent variable is the logit of the recovery rate. Other variable definitions are given in Section 3.2 and summary statistics are reported in Table 1. Robust standard errors adjusted for firm level clustering are given in parentheses, and the number of observations (bond recoveries) N is also reported. RMSE is the root mean square error from out-of-sample testing on the horizon 1996–2004. The initial sample period is 1980–1995; rolling one-year ahead predicted recovery rates and the RMSE are calculated for each year during 1996–2004, and the average RMSE is presented.

	MR1	MR2	MR3
Intercept	-1.8223	3.6671	2.6287
	(0.6241)	(1.3202)	(1.0319)
Coupon rate	0.0435	0.0424	0.0378
	(0.0170)	(0.0184)	(0.0194)
Subordinate	1.0581	0.8163	0.9110
	(0.3064)	(0.2605)	(0.2057)
Senior subordinate	1.1885	0.9964	1.2357
	(0.3476)	(0.3261)	(0.2597)
Senior unsecured	1.6084	1.5773	1.7525
	(0.2973)	(0.2593)	(0.1969)
Senior secured	2.0517	1.9367	2.0402
	(0.3774)	(0.3723)	(0.3526)
Utilities	1.0536	0.8674	0.9108
	(0.4297)	(0.4884)	(0.5361)
Industrials	0.4141	0.5818	0.5455
	(0.2782)	(0.3576)	(0.3652)
Financials	0.5118	1.1866	1.0814
	(0.4960)	(0.6208)	(0.5751)
Relative size		0.2962	0.2646
		(0.0636)	(0.0574)
Logarithm(total assets)		-0.3481	-0.3259
		(0.0714)	(0.0657)
Tangible assets to total assets ratio		0.6778	0.7542
		(0.3330)	(0.3432)
Market to book ratio		-0.5948	-0.5510
		(0.1702)	(0.1514)
Distance-to-default	0.1051		
	(0.0384)		
T-bill 3-month yield	-0.0874	-0.1044	
	(0.0427)	(0.0418)	
S&P 500 return	0.7412	1.0590	
	(0.4457)	(0.4370)	
Logarithm(total defaulted debt)	-0.2572		
	(0.0740)		
N	805	736	736
RMSE	0.2337	0.2288	0.2310

Table 7: **Aggregate Correlations of Default Probabilities and Recovery Rates**

This table summarizes the aggregate correlations between default probabilities and recovery rates predicted out-of-sample, for each pair of frailty default and recovery models. Firms are stratified by industry and bonds are stratified by seniority levels. At the beginning of each year between 1996–2004, we first predict out-of-sample the default probability during the year for each firm with bonds outstanding. Then we predict out-of-sample the recovery rate for each bond; for firms with multiple bonds of the same seniority, we take the average of the predicted recovery rates within the same seniority class. Finally, we stratify the firms by industry and the bonds by seniority, and within each industry and seniority class we compute the bivariate Pearson correlation coefficient between the default probability and the recovery rate across all firms and all years.

Seniority	Default models	Utilities Recovery models MR1	MR2	MR3	Industrials Recovery models MR1	MR2	MR3	Financials Recovery models MR1	MR2	MR3
Junior	MD1	-0.50	-0.58	-0.70	-0.29	-0.32	-0.29	-0.19	-0.13	-0.20
	MD2	-0.55	-0.60	-0.70	-0.36	-0.39	-0.37	-0.30	-0.19	-0.23
	MD3	-0.50	-0.58	-0.70	-0.30	-0.33	-0.30	-0.16	-0.10	-0.16
	MD4	-0.57	-0.66	-0.75	-0.47	-0.52	-0.48	-0.40	-0.28	-0.20
Subordinate	MD1	-0.49	-0.43	-0.47	-0.22	-0.16	-0.15	-0.20	-0.14	-0.12
	MD2	-0.59	-0.52	-0.54	-0.28	-0.18	-0.17	-0.25	-0.19	-0.17
	MD3	-0.48	-0.42	-0.46	-0.23	-0.18	-0.17	-0.20	-0.14	-0.12
	MD4	-0.62	-0.56	-0.56	-0.38	-0.28	-0.27	-0.29	-0.21	-0.17
Senior subordinate	MD1	-0.75	-0.75	-0.64	-0.21	-0.29	-0.26	-0.34	-0.44	-0.46
	MD2	-0.85	-0.85	-0.72	-0.28	-0.34	-0.31	-0.36	-0.44	-0.46
	MD3	-0.75	-0.74	-0.64	-0.23	-0.30	-0.28	-0.34	-0.43	-0.46
	MD4	-0.85	-0.87	-0.72	-0.39	-0.42	-0.39	-0.37	-0.44	-0.46
Senior unsecured	MD1	-0.33	-0.39	-0.41	-0.19	-0.13	-0.11	-0.19	-0.17	-0.14
	MD2	-0.39	-0.54	-0.55	-0.23	-0.18	-0.16	-0.21	-0.17	-0.14
	MD3	-0.35	-0.39	-0.42	-0.20	-0.14	-0.12	-0.20	-0.16	-0.14
	MD4	-0.52	-0.55	-0.55	-0.35	-0.27	-0.25	-0.29	-0.18	-0.13
Senior secured	MD1	-0.35	-0.24	-0.25	-0.12	-0.15	-0.15	-0.48	-0.36	-0.15
	MD2	-0.27	-0.45	-0.49	-0.25	-0.21	-0.15	-0.48	-0.35	-0.13
	MD3	-0.36	-0.23	-0.24	-0.14	-0.17	-0.16	-0.48	-0.36	-0.15
	MD4	-0.48	-0.49	-0.49	-0.43	-0.35	-0.28	-0.48	-0.35	-0.14

Figure 5: Annual Correlations of Default Probabilities and Recovery Rates

The plots give the annual correlations between default probabilities and recovery rates predicted out-of-sample, for each pair of frailty default models MD1 and MD4 and recovery models MR1, MR2 and MR3. The firms are all industrials, and the recovery rates are all for senior unsecured bonds. At the beginning of each year between 1996–2004, we first predict out-of-sample the default probability during the year for each industrial firm with senior unsecured bonds outstanding. Then we predict out-of-sample the recovery rate for each bond; for firms with multiple senior unsecured bonds, we take the average of the predicted recovery rates for each bond. Finally, we compute the bivariate Pearson correlation coefficient between the default probability and the recovery rate across all firms.

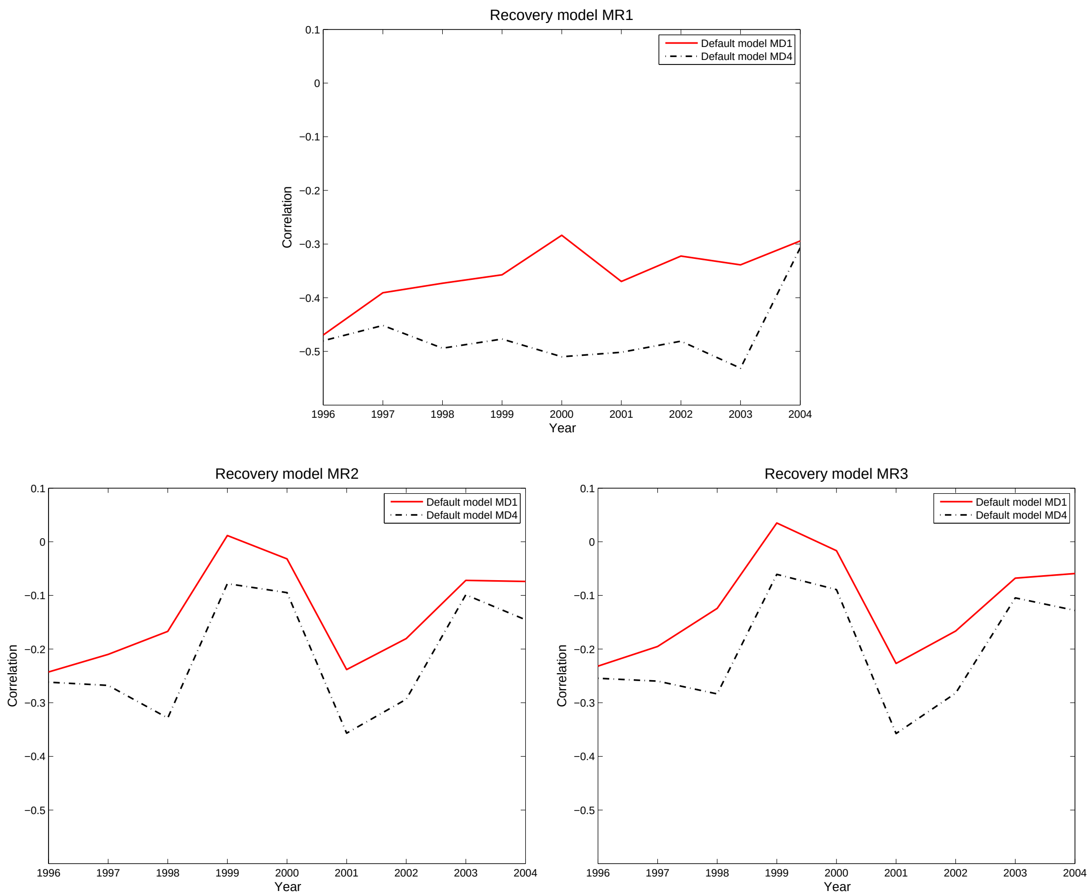

Figure 6: Predicted Loss Distributions

The plots give the probability density functions of the loss distributions during 2003 for the portfolio of firms with DNUM codes between 4800 and 4899 (telecommunication firms), predicted out-of-sample by all four frailty default models and all three recovery models. The actual realized loss in this portfolio during this period is 2.74, with four defaulted bonds arising from three obligor defaults.

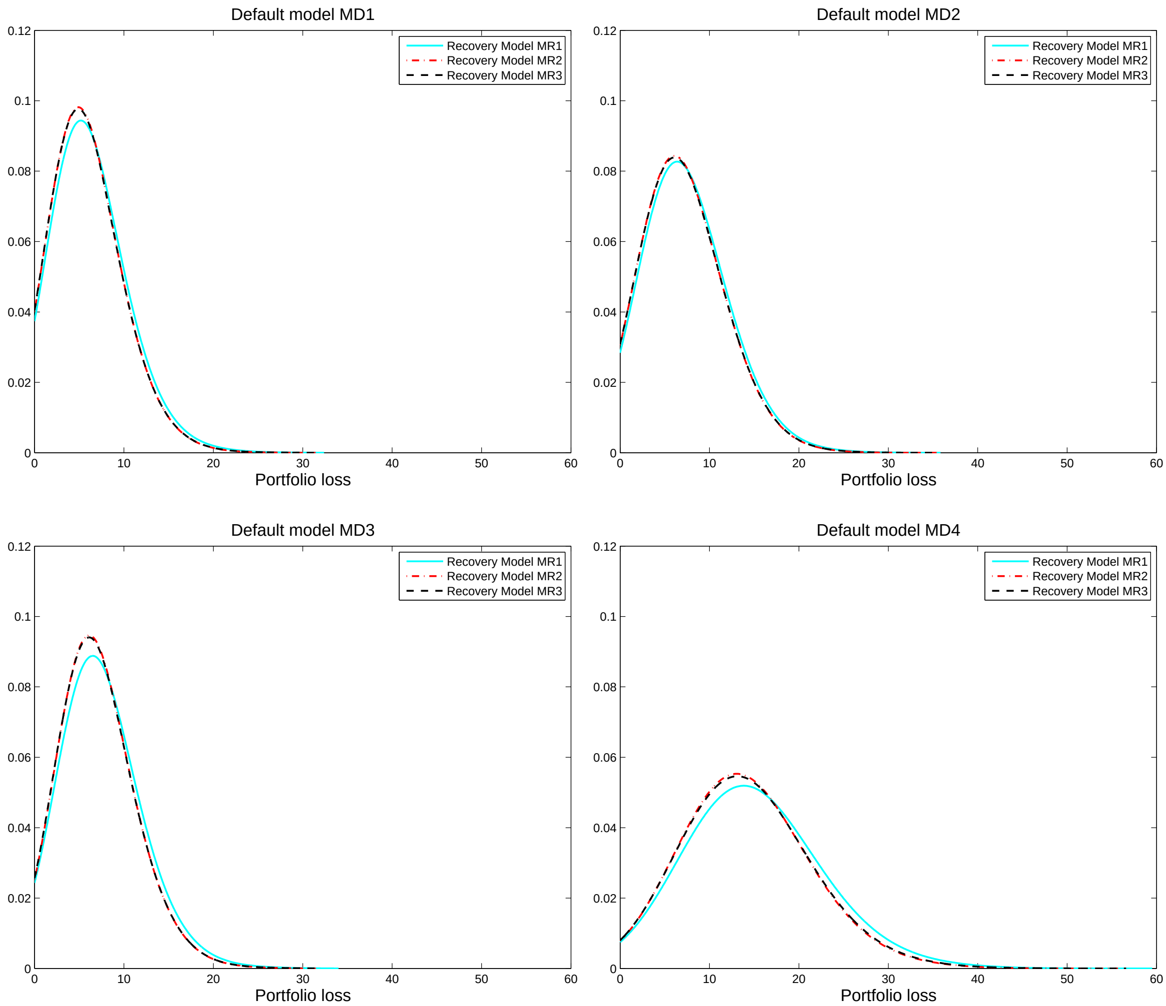

Table 8: **Percentiles of the Loss Distributions**

This table summarizes the expectation and the percentiles of the loss distributions during 2003 for the portfolio of firms with DNUM codes between 4800 and 4899 (telecommunication firms), predicted out-of-sample by all four frailty default models and all three recovery models. The actual realized loss in this portfolio during this period is 2.74, with four defaulted bonds arising from three obligor defaults.

Default model	Recovery model	Mean	Median	95%	99%	99.90%
MD1	MR1	6.06	5.45	12.36	16.72	22.46
	MR2	5.74	5.15	11.80	15.65	21.05
	MR3	5.73	5.13	11.80	15.73	21.23
MD2	MR1	7.27	6.69	14.14	18.49	25.01
	MR2	6.97	6.38	13.74	17.66	23.07
	MR3	6.97	6.37	13.76	17.69	23.37
MD3	MR1	7.38	6.80	14.10	18.37	25.13
	MR2	6.99	6.42	13.40	17.25	23.35
	MR3	6.98	6.40	13.41	17.28	24.58
MD4	MR1	14.95	14.19	25.74	32.68	39.97
	MR2	14.07	13.42	24.08	30.52	37.66
	MR3	14.15	13.47	24.30	30.73	38.17

Figure 7: Predicted Loss Distributions for Different Seniority Levels

The plots give the probability density functions of the loss distributions during 2003 for the portfolios of all bonds of a given seniority level, predicted out-of-sample by the frailty default models MD1 and MD4 and the recovery models MR1 and MR2. The actual realized loss in these portfolios during this period is 2.69 with 3 defaulted bonds in the subordinate class, 7.36 with 9 defaulted bonds in the senior subordinate class, and 21.35 with 35 defaulted bonds in the senior unsecured class.