A Flexible Approach to Modeling Ultimate Recoveries on Defaulted Loans and Bonds*

Edward Altman[†]& Egon Kalotay[‡]

May 10, 2010

Abstract

We present in this paper an intuitive Bayesian approach to modeling the distribution of discounted ultimate recoveries on defaulted debt using mixtures of distributions. We show that the technique is flexible enough to accommodate important idiosyncratic features of recovery distributions and how to adapt the results to target portfolios whose characteristics do not match that of the estimation sample. These applications do not require strong assumptions about the form of the relation between the characteristics of the borrower, the defaulted debt and recovery outcomes. Our empirical results provide insights to the attributes of debt and conditions at the time of default that are associated with economically important variation in the shape of recovery distributions. Expectations of industry level default conditions at the time of default, in conjunction with the Debt Cushion (an economic measure of debt subordination), prove to be of particular value in modeling portfolio-level recoveries. We benchmark the performance of our model against commonly used alternatives in estimating recoveries on portfolios of defaulted debt. Our simple "mixture-based estimates", adapted to reflect the Debt Cushion and industry-specific distress conditions at the time of default, substantially out-perform competing alternatives.

JEL Classification: G10, G20, G21, & G29

Keywords: Ultimate Recovery, Credit Portfolio Losses, Mixture of Distributions

*This paper is a draft: comments welcome.

[†]Stern School of Business, New York University. E-mail: ealtman@stern.nyu.edu, Ph: (212) 998-0709, Fax: (212) 995-4220, Postal Address: New York University, Salomon Center, 44 West 4th Street, Suite 9-61, New York, NY 10012, USA.

[‡]Faculty of Business and Economics, Macquarie University. E-mail: ekalotay@efs.mq.edu.au, Ph: 61-2-9850 8490, Fax: 61-2-9850 8497. Postal Address: Faculty of Business and Economics, Macquarie University, North Ryde, NSW 2109, Australia.

1 Introduction

The economic value of debt in the event of default is a key determinant of the default risk premium charged by a lender and the credit capital charged to limit exposure to losses. The pricing of default risk insurance (CDS contracts) and the emergence of defaulted debt as an investment class add further incentive to better understand the structure of payoffs in the event of default.[1]

Building on insights from prior empirical research, we adopt in this paper a Bayesian perspective and model the distribution of recoveries using mixtures of (normal) distributions. First, we demonstrate that the overall (unconditional) distribution of recoveries, appropriately transformed, is well approximated by a mixture of three normal distributions. We show that the mixture of normals accommodates the unusual empirical features of the sample that may compromise inference based on OLS regression, or the calibration of Beta distributions to recovery outcomes which is often done in practice.

Second, the numerical estimation of the normal mixture is based in part on simulation of latent data using the technique of Gibbs sampling.[2] We make novel use of such data to extract information about how facility characteristics and market conditions at the time of default affect recoveries. In this way we are able to infer information about the conditional distribution of recoveries from a flexible, unconditional estimation procedure. For example, using such an approach we can show that industry-level default expectations have a substantial impact on mixture probabilities, particularly the probability of high recovery outcomes.

Third, we show by way of a re-sampling experiment that our mixture-based estimates of losses very closely approximate the true empirical distribution of losses on portfolios of defaulted debt instruments. Using the latent data generated in the course of estimation, we re-weight the mixture probabilities to reflect the characteristics of the portfolio of interest. Our results suggest that such re-weighting is a convenient way of combining the mixture components to reflect the characteristics of the portfolio of interest. In particular, we find that mixture probabilities tailored to reflect the debt cushion of the exposures within a portfolio and industry-level distress expectations at the time of each default capture the distribution of actual portfolio losses with remarkable accuracy. Further, we show that more restrictive (but popularly applied) estimation techniques tend to systematically

[1] Altman and Karlin (2010) estimate the face and market values of distressed and defaulted debt in the U.S. over time with the most recent estimates at the end of 2009 being $1.6 trillion face value and $1.0 trillion market value. Over 200 institutions invest in these securities.

[2] Gibbs sampling is a well-established Monte Carlo approach to generating draws from analytically intractable multivariate distributions using known marginal (conditional) distributions. In Appendix I.3 we provide a simple example of the technique and in Section 4 we detail its role in our particular application. Refer to Casella and George (1992) for a comprehensive general exposition.

and substantially over or under-estimate outcomes in the left-hand tail of the recovery distribution.

Throughout this study we utilized Moody's estimates of ultimate recovery, discounted to the date of default, as our measure of the economic value accruing to the holder of a particular security or facility at the time of default. Although generally not observable at the time of default, our objective is to model the variable that best reflects economic value at the time of default.

We provide a detailed description of the data in Section 2, but we first provide a brief overview the empirical 'facts' about losses in the event of default gleaned from prior empirical studies of such outcomes. In Section 3 we describe the rationale of our approach to modeling defaulted debt recovery outcomes as mixtures of distributions, and discuss specific econometric issues in Section 4. In Section 5 we present our empirical findings including a comparison of portfolio loss estimates obtained using mixture models to those derived from popular alternatives.

Empirical Characteristics of Recoveries on Defaulted Loans and Bonds

It is widely recognized that the economic payoffs to debt-holders in the event of a particular default depend on the complex interplay of a host of factors (often idiosyncratic) surrounding the default and the attendant processes. Perhaps for this reason, much modeling effort has been unashamedly empirical – seeking to establish the nature of empirical regularities tying debt characteristics or measures of economic conditions to aggregate-level recoveries, or distributions of recoveries. While such models have yielded a distinct set of insights about the behavior of recoveries on defaulted loans and bonds, they have also revealed features of the data that flag the need for caution in interpreting, generalizing or indeed acting upon results obtained from popular tools of inference.

The key empirical features of recoveries can be summarized as follows:

1. Recovery distributions tend to be bimodal, with recoveries either very high or low, implying as Schuermann (2004)[3] observes, that the concept of 'average recovery' is potentially very misleading.

2. Collateralization and degree of subordination are the key determinants of recovery on defaulted debt. The value of claimants subordinate to the debt at a given seniority (or the 'debt cushion') also seems to matter. All else equal, the larger the debt cushion, the higher the recovery – refer to Keisman and Van de Castle (1999).

[3]Schuermann's work provides an excellent review of the empirical features of recoveries while Altman, Brady, Resti, and Sironi (2005) combine a theoretical review as well as important aggregate-level findings.

3. Recoveries tend to be lower in recessions or when the rate of aggregate defaults is high. Specifically, Altman, Brady, Resti, and Sironi (2005) demonstrate an association between default rates and the mean rate of recovery, whereby up to 63% of the variation in average annual recovery can be explained by the coincident annual default rate.

4. Industry matters. Acharya, Bharath, and Srinivasan (2007) suggest that macroeconomic conditions do not appear to be significant determinants of individual bond recoveries once one accounts for industry effects.

5. Variability of recoveries is high – even intra-class variability, after categorization into sub-groups. For example, Schuermann (2004) notes that senior secured investments have a flat distribution, indicating that recoveries are relatively evenly distributed from 30% to 80%.

While the empirical characteristics of the uncertainty associated with recoveries on defaulted debt are well understood, the potential shortcomings of popular modeling approaches are also apparent. Use of OLS regression models and calibrated Beta distributions are two important examples. While regression models provide simple, intuitive summaries of data relationships, they focus attention on variation in the mean. On the other hand, Beta distributions calibrated to historical data on loss given default (LGD) are used in many commercial models of portfolio risk to characterize the distribution of recovery outcomes.[4] However, although the Beta distribution is convenient, Servigny and Renault (2004) observe that it cannot accommodate bi-modality, nor probability masses near zero and 1 – important features, as we will show, of the empirical distribution of recoveries.

Motivated by such considerations, we present in this paper a novel approach to modeling the distribution of recoveries on defaulted loans and bonds based on mixtures of normal distributions. As we demonstrate shortly, not only is our approach flexible enough to accommodate the empirical features of recovery distributions, it is very adaptable in terms of the manner in which data is utilized for estimation and/or inference. Conditioning information can be used to impose restrictions on estimation, or, for inferring the importance of hypothesized recovery determinants.

Our empirical findings complement recent work based on extracting information about recoveries from CDS spreads. For example, Das and Hanouna (2009) find a negative association between recoveries and hazard rates *inferred* from CDS spreads whereas we provide evidence of the impact of default expectations on the distribution of recovery *realizations*, as measured by ultimate recovery.

[4]Portfolio Manager (Moody's KMV), Portfolio Risk Tracker (Standard and Poor's) and CreditMetrics (J.P. Morgan) are all based on the assumption that LGD is described by a Beta distribution.

2 Ultimate Recoveries on Defaulted Loans and Bonds: Data Description

We use in this study discounted ultimate recoveries from Moody's Ultimate Recovery Database (URD). Moody's ultimate recovery database provides several measures of the value received by creditors at the resolution of default, usually upon emergence from Chapter 11 proceedings. Moody's obtain discounted ultimate recoveries by discounting nominal recoveries back to the last time interest was paid using the instrument's pre-petition coupon rate. The database, kindly provided by Moody's of New York, includes US non-financial corporations with over $50m debt at the time of default. The sample period covers obligor defaults from April 1987 to August 2006, covering 3492 debt instruments, of which approximately 60% are bonds. Emery, Cantor, Keisman, and Ou (2007) provide a very detailed description of the database. We focus on features of the data relevant to understanding whether our sample is consistent with some of the stylized facts listed in Section 1.

The histogram of discounted ultimate recoveries for loans and bonds in panel (a) of Figure 1 exemplifies four important attributes of recovery distributions: bi-modality; probability masses at the extremes; truncation and variability. A notable difference between the ultimate recoveries in Figure 1 and those reported by Schuermann (2004) (based on the 2003 year-end Moody's Default Risk Service database) is the substantially higher proportion of high recoveries in our data. Two differences between the samples are apparent however. First, although the current sample contains only three more years of data, the total number of observations utilized in the current study is much larger (3492 vs. 2025). Second, the proportion of loans in the current sample is almost twice as high as that of Schuermann's – greatly boosting the relative proportion of high recoveries in the current sample.

The issue of sample composition raises the question of whether the stylized facts about recoveries hold true if one looks at more homogeneous pools of exposures. The results of Schuermann (2004) suggest that bi-modality and variability do change according to the seniority of exposures. So, in Figure 2 we present a range of discounted ultimate recovery histograms by instrument type, seniority and collateralization to explore whether the same is true of the current sample. From panels (a), (c), (e) and (g) of Figure 2, it is apparent that the recoveries associated with the highest ranking, collateralized claims (loans and collateralized bonds) are highly skewed, with little evidence of bi-modality, but high variability. At the other extreme, the distribution of recoveries on junior bonds and subordinated bonds also exhibit strong right skewness. However there is some suggestion of bi-modality in the subordinated bond distribution.

Panels (b), (d), (f) and (h) of Figure 2 demonstrate that bi-modality is a feature of middle ranking bond exposures, and uncollateralized bonds. Senior secured exposures appear to exhibit three distinct modes, providing a somewhat different picture of the distribution than the density estimate for the corresponding class of bonds provided by Schuermann (2004). However, the message is largely consistent: recovery distributions associated with the highest ranking, collateralized exposures, and the lowest ranked bond exposures tend to be highly skewed and closer to uni-modal, but recoveries on middle ranking and uncollateralized bonds clearly exhibit probability masses at the extremes.

Table 1 summarizes the features of the data presented in Figures 1 and 2. Loans and collateralized bond exposures at one extreme, and junior debt at the other, are skewed in opposing directions and somewhat less volatile than lower ranking or uncollateralized exposures. The idiosyncratic distributional features of the recoveries data, and the variation of the distributions by characteristics such as seniority and collateralization serve to illustrate the challenge of specifying a model of recoveries that's *a-priori* consistent with the data, as well the need for caution in interpreting the concept of average recovery.

3 Defaulted Debt Recoveries as a Mixture of Distributions

Consistent with the stylized facts gleaned from previous empirical studies, our current observations underscore the importance of modeling the overall distribution of recoveries to properly understand the uncertainty associated with defaulted debt exposures. However, instead of trying to force-fit a parametric distribution, we take a Bayesian perspective and characterize the distribution of recoveries using a mixture of distributions approach. By taking the appropriate probability weighted average of normals, we are able to accommodate the unusual (defining) features of such distributions.

To be more specific, our modeling approach rests on the assumption that recovery outcomes y can be thought of as draws from some distribution $g(y)$ of unknown functional form. While the form of $g(y)$ is not known, we set out to approximate it using a weighted combination of standard densities $f(y|\theta_j)$ such that:

$$g(y) \approx \hat{g}(y) = \sum_{j=1}^{m} p_j f(y|\theta_j) \tag{1}$$

where $p_1 + \ldots + p_m = 1$, and the standard densities $f(y|\theta_1), \ldots, f(y|\theta_m)$ form the functional basis for approximating $g(y)$. In our application, the m densities $f(y|\theta_j)$ are chosen to be normal with parameters θ_j. Robert (1996) observes that such mixtures can model quite exotic distributions with few parameters and a high degree of accuracy. The tractability of the mixture components implies that properties of the distribution relevant to inference obtain quite easily.

As shown in Figure II, the simplest 2-normal mixture (obtained by drawing with equal probability from two normal distributions of differing mean but equal standard deviation) can be used to approximate a symmetrical bi-modal distribution. The shape of the distribution can be altered dramatically by varying the parameters of the component normals, as well as the probability of drawing from each. This illustrative example is rather simple because we know the number of normals required to simulate the density of interest, as well as the parameters of the simulation (the properties of each component normal, and the probability of drawing from each). In practice, the modeling parameters need to be estimated based on the available data. Further, given that the mixture parameters are interdependent, all the modeling inputs need to be determined simultaneously. Fortunately, there are well established techniques to solve such problems in a Bayesian framework – the Markov Chain Monte Carlo (MCMC) technique of Gibbs sampling in particular.

4 Econometric Framework: Some Elaboration

As noted earlier, we commence by assuming a normal form for the approximating densities $f(.)$ in (1). We also generalize the notation to account for the possibility of explicit conditioning on other variables x. That is, conditional on x, observable characteristics or factors that are known determinants of recovery, we model ultimate recoveries y using a probability p_j weighted mixture of m normal likelihoods:

$$P(y|x, \beta, h, \alpha, p) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left\{ \sum_{j=1}^{m} p_j \sqrt{h_j} \exp\left[-\frac{h_j}{2}(y_i - \alpha_j - \beta' x_i)^2 \right] \right\}, \tag{2}$$

where α_j is the mean of mixture component j conditional on a linear combination of x, and h_j its variance. If x is excluded from the analysis, then α_j is simply the mean of mixture component j and h_j its variance. The vector β is a $1 \times k$ collection of slope coefficients when each x_i is of dimension $k \times 1$.[5] The sample size is N.

Confronted with the likelihood (2), following the specification in Koop (2003), we adopt proper but minimally informative conjugate priors, on the parameters α, β, h and p, with a view to estimating the joint posterior of all parameters using the Markov chain Monte Carlo (MCMC) technique of Gibbs Sampling.[6] However, to accomplish this, two problems must be addressed.

First, there is no directly observable data with which to estimate the probability weights p_j. Second, there exists an identification problem in that multiple sets of parameter values are consistent with the same likelihood function.[7] Fortunately, there are established solutions to both problems. The second (identification) problem is circumvented by way of a *labeling restriction.* Specifically, we follow Koop (2003) in imposing the restriction that $\alpha_{j-1} < \alpha_j$ for $j = 2 \ldots m$. While there is nothing special about this particular restriction (in the sense that restrictions on other parameters can equivalently solve the identification problem), it facilitates interpretation of the Gibbs output.

The solution to the problem of not observing data with which to estimate p_j involves a well-established technique called *data augmentation.* If one were to observe an indicator variable e_{ij} taking on a value of 1 when observation i is an outcome drawn from mixture component j, and zero otherwise, then the likelihood (2) could be written in terms of it as:

$$P(y|x, \beta, h, \alpha, p, e) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left\{ \sum_{j=1}^{m} e_{ij} \sqrt{h_j} \exp\left[-\frac{h_j}{2} (y_i - \alpha_j - \beta' x_i)^2 \right] \right\}, \tag{3}$$

and estimation would follow easily.[8] However, since we do not observe indicator flags associating observations with mixture components, we rely on the decomposition described in Robert (1996) to generate them as part of the sampling scheme.[9] Specifically, the latent data is generated based on draws from a Multinomial distribution. Conditional on the data and parameters of the mixture components (α_j, β_j, h_j), the latent data draw associated with each observation is an $m-$vector of indicator variables wherein one of the indicators is non-zero. In particular, a value of 1 in position j associates the observation with mixture component j. The probability of an observation being so

[5] For the sake of clarity we suppress wherever possible time and facility/firm subscripts. Unless stated otherwise, all analysis is on data pooled in time series and cross section, and all data contained in x are observable prior to data on the corresponding row of y.

[6] Readers unfamiliar with the technique of Gibbs sampling and its application in Bayesian estimation should refer at this point to Sections I.1 and I.2 of the Appendix.

[7] Refer to page 255 of Koop (2003) for elaboration and an example.

[8] Robert (1996) notes that this re-expression is possible when the likelihood is from an exponential family.

[9] Refer to equation 24.7 in Robert (1996).

assigned to mixture component j on any particular draw of the sampling scheme depends on the relative likelihood of it being observed as an outcome of the particular mixture component.[10]

While the latent data generated as part of the sampling scheme is motivated by computational considerations, it turns out to have an extremely useful economic interpretation in the context of our application. As will be explained shortly, it enables us to compute the posterior probability with which observations (or more importantly, economically interesting sets of observations) were generated with particular mixture components.

A final *a-priori* issue that arises in the modeling of loan recoveries as a mixture of normals is the possibility of observing outcomes of less than zero or greater than unity.[11] As can be seen in Figure 1, it is possible to occasionally observe ultimate recoveries significantly greater than unity. However, we rule out expectations of such outcomes by modeling transformed recoveries, such that they are mapped to the real number line. To do this we constrain observations to the interval $(0, 1)$ and map the resultant series to the real number line using the inverse CDF of a distribution with unbounded support – such as the standard Gaussian.[12]

5 Empirical Findings

We commence by discussing in Section 5.1 how we determine the number of mixture components with reference to both information criteria and the degree of correspondence between the observed sample and simulated predictive distributions. In Section 5.2 we discuss the economic interpretation of the mixture components and introduce our measure of mixing probabilities associated with sub-portfolios of exposures, conditional on characteristics or economic conditions. We apply our approach to modeling the impact of industry-level default expectations (gleaned from equity prices) on discounted ultimate recovery outcomes. In Section 5.3 we present results based on a mixture model that makes no *a-priori* assumptions about the form of the relation between the recovery outcomes and their determinants, and we contrast these findings with those based on a flexible linear regression model. We conclude our results by quantifying economic performance of alternative modeling approaches by way of a Monte Carlo study. In particular, we consider in Section 5.5 the problem of the modeling the distribution of recoveries associated with portfolios of defaulted loans and bonds.

[10] See step 5 of the sampling scheme detailed in Appendix I.3.

[11] See Carey and Gordy (2004) for a discussion of such outcomes and the related discussion of apparent violations of the absolute priority rule.

[12] Hu and Perraudin (2002) use the inverse CDF of the Gaussian to map recovery outcomes to the real number line.

5.1 Determining the Number of Mixture Components

The pooled distribution of recoveries on loans and bonds, presented in Panel (a) of Figure 1, serves as our starting point. As noted in Section 4, we map observations of recovery, lying between 0 and 1, to the real number line using the inverse CDF of a distribution with unbounded support on the real number line. There are many candidate distributions that we could use to perform the transformation. With no clear *a-priori* reason to prefer any of them, it would be simplest if we could state that this modeling choice is innocuous. However, it does have an impact on the shape of the transformed distribution and the subsequent choice of mixture components.

In Panel (b) of Figure 1, we show the distribution obtained using the inverse of the Student-T CDF, with $v = 20$, to transform the data. Clearly, the transformed data remains multi modal. Using widely-applied information criteria to guide our choice of the optimal number of mixture components, it appears that the data transformed using the inverse of Student-T with $v = 20$ is optimally modeled with a 3-component mixture. As can be seen in the lower panel of Table 2, the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the Hannan-Quinn criterion (HQ) are simultaneously maximized when the number of mixture components $m = 3$. If, for example, we transform the data using the inverse of the Gaussian CDF, then the results in the upper panel of Table 2 obtain. Not only do the information criteria suggest a larger number of mixture components, they provide inconsistent signals. That is, BIC and HQ imply the optimality of 5 mixture components, while AIC suggests 6.[13]

While information criteria help guide the choice of m, the success of any specification is best evaluated in terms of how well the mixture specification captures the empirical features of the data. Comparing Panels (a) and (b) of Figure 2 with the corresponding panels of Figure 1 provides a visual summary of overall success. In particular, we plot in Figure 4 the distribution of discounted ultimate recoveries based on a mixture of 3 normals. The distribution of 3,500 recoveries simulated from the 3-component mixture and the corresponding distribution of transformed recoveries appear closely matched to their sample counterparts. The correspondence between the properties of the observed sample and the sample generated from the using the Gibbs output are borne out in the final column of summary statistics reported in Table 3. The mean of the simulated sample (0.59) is close to that of the observed sample (0.56), and both the standard deviation and interquartile range of the simulated data and observed data are almost identically matched. The median of the

[13]Given that the information criteria are best regarded as a rough guide to optimal model selection, we examine the impact of varying m on reported results. Further, in the interests of full disclosure, we note that we settled on using the transformation based on the inverse CDF of the Student-T with $v = 20$ after after estimating the results using the inverse of the Gaussian, and three variants of the Student-T: $v = 10$, $v = 20$ and $v = 50$. The information criteria obtained using $v = 20$ were consistent in suggesting a parsimonious specification with $m = 3$.

simulated data (0.65) is higher than the 0.58 median of the sample data, but otherwise, the overall distributional properties of the pooled sample of discounted ultimate recoveries appear closely matched by the simulated sample of similar size based on a 3-component mixture.

5.2 Economic Interpretation

The mixture parameters, or (more precisely) the posterior distribution of the mixture parameters, have a distinct economic interpretation. The unconditional distribution of ultimate recoveries on defaulted debt can be thought of as a draw from one of three distributions, the parameters of which are presented in the middle panel of Table 4. The first component of the mixture implies a mean recovery of zero, the second mixture component is a distribution with a mean of 35% and standard deviation of 27.5%, and the third implies a recovery of 100% with no economically meaningful variation.

A Brief Aside: Interpretation of Mixture Components

Note that the distributions documented in the second panel of Table 4 are derived by mapping the draws from each mixture component back to a measure of recoveries using the CDF of the Student-T with $v = 20$. The values of $E(\alpha_i|y)$ and $E(h_i|y)$ reflect the mean and variability of the mixture components associated with recoveries transformed using the Student-T CDF with $v = 20$. These parameters must be interpreted with caution. For example, the posterior mean $E(\alpha_2|y) = -0.55$ does NOT imply that the mean recovery obtains by direct substitution of -0.55 to the CDF of the relevant Student-T because the impact of deviations from the mean of the transformed data is non-linear and asymmetric.

The latter is most easily observed in Figure 5 wherein the deviations from each mixture component (sampled as transformed data) are mapped to the [0,1] recovery scale using the Student-T. For example, from the solid line associated with the second mixture component, variation of a half standard deviation above the mean implies a 20% increase in recovery, while the corresponding deviation below the mean implies a decrease of 15%. It is also worth noting that a draw from the second component of the mixture delivers a recovery of 30% of less with 50% probability, and a recovery of less than 10% is only 0.75 standard deviations below the mean – thus occurring with a 40% (normal) probability. By way of contrast, there is a much lower probability of a recovery in the right hand tail of the distribution: a recovery of 90% or more is 1.75 standard deviations above the mean realized with a probability of less than 4%.

Figure 5 also illustrates the economic insignificance of variation in draws from the first and third mixture components delineated by circles and triangles respectively. In neither case does variation from the mean of the respective mixture component imply variation in the quantum of recovery when transformed to an economically meaningful scale using the Student-T cdf.

End of Aside

In the absence of any information about the characteristics of the defaulting firm or the nature of the facility or prevailing macroeconomic conditions, the posterior probability of the recovery outcome being drawn from each mixture component can be gleaned from the second last row of Table 4. Given that draws from the first and third components of the mixture imply zero ultimate recovery and full recovery respectively, the probability weightings on these mixture components provide an intuitive basis for summarizing risk exposures and making comparisons of relative risk exposures. To illustrate the latter point, we observe for the moment that in the absence of any other information the posterior probability of drawing from the first mixture component and thus realizing zero ultimate recovery is 6%, while the probability of drawing from the third mixture component and thus realizing 100% ultimate recovery is 35%.

In assessing the potential payoffs of defaulted debt, an analyst familiar with the stylized facts about recoveries on defaulted debt will consider the characteristics of the debt security relevant to recovery, such as its seniority and degree of collateralization, as well as the extent to which the macroeconomy or particular industry of the defaulting firm is stressed. Such micro or macro-level conditioning information is likely to alter the mixing probabilities associated with different mixture components. For example, all else being equal, the probability of a recovery outcome being a draw from the first (third) mixture component should be lower (higher) for loans than it is for bonds. To quantify such effects using the output of the Gibbs Sampler, we make use of the latent data, discussed in Section 4, and generated as per step 5 of the algorithm described in Appendix I.3.

Recall from the likelihood in equation (3) that each observation i is associated with a mixture component j by way of the indicator variable e_{ij}. Each iteration of the Gibbs sampler involves drawing from the conditional posterior of the indicator variables e_{ij} for each $i = 1 \dots N$ exposure, thus providing the information required to compute the probability (mixing) weights associating particular portfolios of exposures with each mixture component. Suppose for example that we are interested in modeling this distribution of recoveries on subordinated debt. Suppose the debt exposures $i \in Q$ denote the sub portfolio of interest − recoveries on subordinated debt in our example. Then, p_{Qj}, the mixing weight for portfolio Q associated with component j, can be estimated from the Gibbs output using

$$\hat{p}_{Qj} = \sum_{g=1}^{G} \frac{e_{Qj}^{[g]}}{n(Q)G} \tag{4}$$

where e_{Qj} denotes all e_{ij} such that $i \in Q$, G is the total number of post burn-in iterates from the Gibbs sampler, and $n(Q)$ is the number of observations in Q. Using equation (4) we can compute the mixing probabilities for a portfolio of subordinated debt exposures as the proportion of non-zero indicator variables sampled for each component j. Based on the Gibbs output we find mixture probabilities of $\hat{p}_{Q1} = 23\%$, $\hat{p}_{Q2} = 67\%$ and $\hat{p}_{Q3} = 10\%$ for subordinated debt exposures – findings that accord with expectations relative the unconditional weights reported in Table 4.

Computing the mixture probabilities for different classes of debt exposures using the Gibbs output enables us to generate the predictive distributions specific to each class. Table 3 summarizes the properties of recoveries using 10,000 draws from the 3-component mixture. Comparing the properties of the the modeled distributions in Table 3 with those of the corresponding empirical samples in Table 1, we find that the mixture weights differ across sub-classes in accordance with intuition. All else being equal, unconditional recovery distributions associated with more senior and collateralized claims exhibit a higher weight on mixture component 3, and a lower weight on mixture component 1. Most encouragingly, the mean, median and variability of the sub-classes observed in the data is well approximated by the modeled distributions – in terms of both absolute and relative comparisons.

The mixture of distributions appears to work least well for Junior and Subordinated debt claims. As noted earlier, the right-skewed empirical distributions of recovery on Junior and Subordinated debt depart most markedly from those of other defaulted claims, exhibiting a heavy concentration of very low recoveries, with relatively few high recoveries. However, in making these observations we note that the two categories in question comprise just over 10% of the sample, and are almost non-existent in recent years. As such, the empirical properties of the samples associated with these categories (recoveries on Junior debt in particular) may not extend more generally. Further, if we are convinced that Junior debt and Subordinated debt recoveries are sufficiently 'different' to others, we could model them separately, or specify a variant of model (2) incorporating a set of indicators for such exposures prior to estimation.

5.3 Inferring the Impact of Industry-Level Default Expectations

Macro or systematic factors that seem to influence recovery include the economy-wide distress rate, and measures of relative industry performance prior to the time of default. Specifically, Altman, Brady, Resti, and Sironi (2005) demonstrate a significant inverse association between default rates and the mean rate of recovery, whereby up to 63% of the variation in average annual recovery can be explained by the coincident annual default rate. More recently, the findings of Acharya, Bharath, and Srinivasan (2007) suggest that macroeconomic conditions do not appear to be significant determinants of recovery once one accounts for industry effects. Earlier work by Carey and Gordy (2004) suggests that the relation between macro-economic conditions and recoveries is driven by the increase in 'bad' (< 60% recoveries) during high default periods.

While economic intuition and empirical evidence suggests a link between the aggregated level of recoveries and recessions (or times of high aggregate default), the impact of such recessions on the distribution of recovery outcomes is less well understood. Given the inherent limitations of aggregate recovery metrics, we construct distributions of recovery conditional on expectations of industry-level default rates. Specifically, we match all defaulting firms to one of the 17 Fama-French industry portfolio groups and use the corresponding *industry default likelihood* as a market-based measure of expectations.[14] Industry default likelihood at time t is the mean of the Merton (1974)-model implied risk-neutral default probabilities for all firms of a given industry at time t. For purposes of modeling states of relatively high industry-level default risk, we identify observations where the industry default likelihood was 1.7 or more standard deviations above its time series mean at the time of firm default.[15]

We present in Table 5 the analogue of the results in Table 3, conditional on industry distress at the time of default. That is, we construct the mixtures in Table 5 using equation (4) as before. We compute, however, the mixing probabilities using only those observations within each class that are associated with defaults at times of industry distress. Differences between the distributional summaries provided in Table 5 and the corresponding estimates in Table 3 suggest a substantial shift in the projected distribution of recovery outcomes. This finding is consistent with a hypothesis suggested by Carey and Gordy (2004), namely, that the distribution of LGDs shifts to the right in good years, relative to bad years. Carey and Gordy (2004) also suggest that a higher proportion of bad LGD firms may be selected into bankruptcy in high default years while less-than-bad LGDs may not be significantly affected. Our findings are consistent with the former but not the latter part

[14] The industry portfolio classifications correspond to the 17 industry portfolio groupings kindly provided Kenneth French in his data library at: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[15] Refer to Altman, Fargher, and Kalotay (2009) for details of how the probabilities are estimated and how the industry-level indexes are applied to default risk estimation.

of the hypothesis: the probability of high recoveries drops substantially in years where expectations of default risks are elevated at the time of default.

Overall, the probability of drawing a recovery outcome from the third mixture component (implying full recovery) declines from 35% in Table 3 to 26% in Table 5, but the probability of drawing from the first mixture component (implying total loss) increases only slightly from 6 to 7%. A similar example is afforded by non-collateralized bonds: the probability of drawing from the third mixture component falls from 16% to 11%, while the probability of drawing from the first mixture component (implying no recovery) increases from 11% to 12% only. While there is a slight increase in the probability of a total loss, and a substantial decline in the probability of a total recovery when default occurs at a time of industry distress, it is interesting to note that the mean recovery changes very little (from 49% to 47%).

The lower panel of Table 5 provides the characteristics of the mixture distribution conditional on industry default expectations being below the median of such expectations at the time of default.[16] Focusing again on the extreme mixture components, times of low industry default expectations are marked by a substantial increase in the probability of full recovery: the probability weighting on the third mixture component increases substantially while the weighting on the second mixture component tends to decline accordingly – again suggesting that most of the adjustment occurs in the right hand tail of the distribution. Another notable feature of the distribution is the variation in the impact of industry default expectations across exposure categories: bonds are more sensitive than loans, and uncollateralized exposures are more sensitive than collateralized.

Taken together, our findings suggest that the distribution of recoveries exhibits substantial variation in accordance with a market-based measure of the industry-level default outlook. However, our results also highlight several sources of asymmetry in the adjustment to variation in the default outlook. First, most of the variation in recovery outcomes coincident with variation in industry default outlook is captured by variation in the probability of a full recovery. When the industry default outlook is relatively bad (good), the probability of a full recovery diminishes (increases), while the probability of a total loss is relatively invariant. Second, the magnitude of the adjustments in mixing probabilities varies according to the characteristic(s) of the (sub) portfolio.

[16]The median value of standardized industry default expectation at the time of default (for the current sample) is 1.01. This implies that the median default occurs when the expectation of industry level default is approximately one standard deviation above its time-series mean.

5.4 Conditional (Regression) Estimates

Altman, Brady, Resti, and Sironi (2005) (and subsequent studies) use time-series regressions to model the relation between aggregate recoveries and the concurrent aggregate default rate, and the strength of the aggregate-level association documented in their work is argued to show the importance of accounting for such effects in credit risk models. Trück, Harpainter, and Rachev (2005) show that the CBOE market volatility index increases the explanatory power of the Altman, Brady, Resti, and Sironi (2005) model of mean recovery.

Acharya, Bharath, and Srinivasan (2007) estimate pooled facility-level regressions to study the relation between defaulted debt recoveries and their posited determinants at facility, firm, industry and economy-level. In addition to providing readily interpretable summaries of data relationships, such models simplify the modeling of exposures in circumstances where the population or portfolio of interest is comprised of a different mix of characteristics to the historical sample available for estimation.

Since linear models of the relation between the recoveries, characteristics and conditions are of interest both in the context of the literature and real-world practicalities, we re-estimate the posterior distribution of recoveries assuming a linear relation between ultimate recoveries and a set of micro and macro-level recovery determinants. As can be seem from equation (2), if $m = 1$, then this approach is equivalent to a standard regression model. Setting $m \geq 2$ is a generalization of the regression model wherein we model the errors as a mixture of normals. In particular, we estimate the model reported in Table 6 based on $m = 2$. We include dummy variables to capture a range of facility-level information associated with recovery (such as the debt cushion, time in default, instrument rank, collateralization, default type and seniority), together with our measure of industry distress expectations.

Table 6 reports the posterior mean of the regression model estimates. While the sign (direction) of the relation between the conditioning variables and the discounted recovery outcomes accord with economic intuition, the absolute magnitude of the coefficients and the variability of each coefficient's posterior distribution should be interpreted with the discussion in Section 5.2 in mind. Further, the relative magnitude of at least one coefficient may at first seem contrary to intuition. For example, relative to loans (the null case), the posterior mean of the coefficients appears to imply that senior secured bonds have a lower expected recovery than senior unsecured bonds. However, such direct inference is misleading in this case, given that the posterior probability of senior unsecured bonds being drawn from the first mixture component (implying zero recovery) is 5.8% – higher than that of senior secured bonds (≈ 0).

The specification in Table 6 with $m = 2$ was strongly supported relative to a specification with $m = 1$ based on all three information criteria, but soundly rejected in favor of specification with $m = 3$. However, while $m = 3$ seemed optimal based on the information criteria, modeling the errors with 3 or more mixture components results in coefficient estimates that centered on zero. That is, the results of the modeling ended up looking very much like those reported for $m = 3$ with no conditioning variables. While all this suggests that specifying a model wherein recoveries are linearly related to the conditioning variables is rejected by the data, we consider such specifications in the analysis to follow because of their practical appeal and straightforward intuition. Consistent with the signal from the information criteria, however, the simulated distributions reported in Table 7 do a relatively poor job of matching the empirical characteristics of the sample. The standard deviation of the overall distribution and the various sub-portfolios is well approximated, but most other features of the simulated data depart substantially from the empirical observations reported in Table 1.

5.5 Evaluating Portfolio Losses: A Stylized Application

Using information criteria to guide model choice and assessing modeling success in terms of the match between the features of the empirical observations and modeled data enable relative (albeit somewhat subjective) statistical comparisons of performance. In this final section of our results we present an *in-sample* measure of model fit in terms of a simple, stylized application of alternative modeling approaches. In particular, we set out to find the modeling approach that best approximates the left-hand tail of the recoveries distribution on randomized sets of equally weighted defaulted debt portfolios, of a given size. We proceed with the following re-sampling experiment.

1. Draw a random sample of exposures from the pooled sample of discounted ultimate recoveries and evaluate the ultimate recovery on an equally-weighted portfolio of the randomly selected exposures. This value is stored as an outcome of the empirical loss distribution.

2. Based on the same set of exposures we draw and store an outcome from each of the following distributions:

   (a) *3-Mix Base:* is a draw from the unconditional estimate of the (posterior) recovery distribution based on the 3-mixture specification. This distribution utilizes the entire sample of empirical observations. It accounts for estimation risk in parameter estimates but does not incorporate any prior knowledge about the exposure, facility or industry-level default expectations.

(b) *3-Mix Char:* is based on the components of the 3-mixture unconditional estimation. However, in this case the mixing probabilities used to weight the component distributions are re-computed based on particular characteristics of the randomly sampled exposures. The results reported in Table 8 are based on mixture probabilities that are computed based on the Debt Cushion and Collateralization characteristics of the sampled portfolio on each iteration. The Debt Cushion categories used for purposes of this exercise correspond to the column categories in Table 9. This distribution does not incorporate any prior knowledge about industry-level default expectations.

(c) *3-Mix Comp:* is based on the components of the 3-mixture unconditional estimation. However, in this case the mixing probabilities used to weight the component distributions are re-computed based on the Debt Cushion associated with the sampled exposures and the expectations of default in the industry of the defaulting borrower at the time of default. The mixture probabilities associated with each category and the specific modeling choices are reported in Table 9. As such, while this procedure is based on output of the full unconditional 3-mixture distribution, the sampling reflects the Debt Cushion of the exposure and the industry-level default outlook at the time of default.

(d) *Regression:* is a draw from the posterior distribution of recoveries implied by a single mixture regression model using the conditioning variables reported in Table 6. Draws from this distribution reflect the estimation risk associated with parameter estimates.

(e) *2-Mix Reg:* is a draw from the posterior distribution of recoveries implied by the 2-mixture regression specification reported in Table 6. Draws from this distribution reflects the estimation risk associated with parameter estimates.

(f) *Beta Comp:* is a draw from the Beta distribution calibrated to the Debt Cushion category of the facility in question and the level of industry default expectations. The parameters of the Beta distribution corresponding to each of the categories in Table 9 are computed using method of moments estimators and all available sample observations in each category. This technique is the Beta distribution-based counterpart of *3-Mix Comp.* Given the popularity of the Beta distribution in modeling default losses, it serves as an interesting benchmark of practical significance.

3. Steps 1-2 are repeated 10,000 times. The tail percentiles and summary statistics associated with the resultant distributions are computed.

Before considering the results we note that the intention behind this bootstrapping exercise is two-fold. First, to demonstrate the potential practical utility of our approach to using the unconditional mixture distribution when the characteristics of the sample used for estimation does not match the target portfolio (as is generally the case). Second, to provide an economic measure of the ability of alternative approaches to capture key features of portfolio recovery distributions – the tails in particular. The current experiment is not intended as a measure of out-of-sample predictive value. However, if a modeled distribution has poor ability to capture empirically observed outcomes then there is a little basis for expecting it to perform well out-of-sample.

With the objectives of the exercise in mind, we turn to the results reported in Table 8. The distribution of portfolio losses generated using *3-Mix Comp* match the characteristics of the actual (bootstrapped) distribution with far greater accuracy than any of the other candidates considered.[17] Recall that *3-Mix Comp* re-weights the components based on the Debt Cushion of the exposures in the target portfolio, and industry level default expectations at the time of default. Ignoring industry default expectations results in substantial over or underestimation of portfolio level losses (depending on the specification). Considering Debt Cushion and collateralization only results in an approximately 6% underestimation of recoveries throughout the distribution, or equivalently, a 25% overestimation of realized losses based on *3-Mix Char.* Simulating portfolio losses based on the unconditional distribution, as per *3-Mix Base* without reference to any characteristic information or expectations of default, results in a corresponding underestimate of losses.[18]

The remaining estimates reported in Table 8 consistently exhibit larger errors than any based on the unconditional mixture. It is worth noting that the 'standard' regression model tends to underestimate recovery by less than the 2-mixture regression, thus implying (perhaps not surprisingly) that the signal provided by information criteria do not necessarily imply better economic performance.[19] The final column of Table 8, reveals that the Beta distribution, calibrated to the Debt Cushion and industry default outlook observed at the time of default, performs about as poorly as the 2-mixture regression. In the current context, the latter result is the most direct basis on which the gains from using a mixture approximation approximation relative to the Beta distribution can be compared, as both sets of estimates are conditioned on the same set of prior information (Debt Cushion and industry default expectations).

Some Comments on Modeling Choices and Generalizability

The facility and industry-level conditioning variables in our portfolio experiment were chosen with a view to parsimony, for broadly similar reasons.

[17]The conditional mixture probabilities are reported in Table 9.

[18]Investigating the possibility that this is attributable to estimation error in the unconditional mixture weights.

[19]For example, in this instance, the 'better' economic performance would depend on the relative costs associated with errors of over and underestimation.

Debt cushion, suggested by Keisman and Van de Castle (1999), is a facility-level metric that captures not only the rank of debt in capital structure, but the degree of its subordination as a proportion of total claims. Keisman and Van de Castle (1999) present evidence to show that Debt Cushion categories are associated with extreme recoveries. Hence, we have *a-priori* empirical evidence to suggest that Debt Cushion is a potentially useful facility-level characteristic.

Our use of industry-level default expectations is motivated by a mixture of theory and empirical evidence. The theoretical work of Frye (2000) implies a negative association between the probability of default and recovery outcomes. The empirical estimates of Altman, Brady, Resti, and Sironi (2005) are consistent with Frye's theory to the extent that realizations of default are used in place of expectations. The empirical work of Acharya, Bharath, and Srinivasan (2007) suggests that macroeconomic effects are displaced (or subsumed) by industry effects in models of recovery, and Altman, Fargher, and Kalotay (2009) show that industry-level aggregates of Merton default probabilities play an important role in firm-level models of default risk.

Taken together, the results of the portfolio loss evaluation experiment suggest an important role for both Debt Cushion and industry default expectations in modeling portfolio credit losses: the combination of the two variables provide in our application a very parsimonious basis for adapting the mixture probabilities to the characteristics of the target portfolios. However, we note the following observations and caveats.

1. It is likely that our simple Merton-model based measure of default risk reflects both expectations of default and recovery in the event of default.
2. Our findings appear robust to modeling assumptions specific to the re-sampling experiment.[20]
3. As noted earlier, in their current form our portfolio results represent an alternative measure of model fit, based on a simple stylized application. However, we have reported results based on all the models we've estimated to date. That is, our reported results are not based on any formal or informal process of optimizing fit.

6 Summary

We have presented in this paper a simple Bayesian approach to modeling the distribution of discounted ultimate recoveries on defaulted debt using mixtures of normal distributions. We

[20] Robustness test results based on alternative assumptions about portfolio size and homogeneity of exposure size are available upon request.

show that the technique is flexible enough to accommodate important idiosyncratic features of recovery distributions, and we show how to adapt the results of estimation to target portfolios whose characteristics do not match that of the estimation sample – without making strong assumptions about the form of the relation between the characteristics of the exposures and recovery outcomes. In so doing we make novel use of the latent data that is generated as a by-product of the estimation technique.

Our approach to modeling recoveries using mixtures of distributions enables several interesting empirical insights that complement and extend findings in the literature. First, we demonstrate the importance of industry-level default expectations at the time of firm default as a determinant of the shape of the recovery distribution – complementing the findings of Altman, Brady, Resti, and Sironi (2005) and Acharya, Bharath, and Srinivasan (2007) regarding the relation between realized aggregate default rates and mean recoveries, and the empirical findings of Das and Hanouna (2009) based on the relation between default and recovery expectations inferred from CDS spreads. Second, we show that the Debt Cushion of a facility, combined with information about industry default expectations at the time of default enable us to model the distribution of randomized portfolio losses with remarkable accuracy. Third, our results suggest that common modeling simplifications (such as the modeling of conditional recoveries by way of a calibrated Beta distribution) can be circumvented using mixtures of distributions with encouraging results – as measured by ability to approximate simulated portfolio loss distributions.

In addition to the conceptual contributions implicit in our methodology for estimating discounted ultimate recovery rates, market practitioners can glean important insights at various stages of the investment process. Fixed income analysts and risk managers can better estimate loss-given-default estimates when they are considering investments, including those whose market values are at or near par value, once estimates of default probabilities are established. This has obvious benefits for banks in establishing capital levels under Basel II requirements as well as for investors who simply want to estimate required returns. Further, recovery rate forecasts can be revised over time as aggregate market and industry level estimates of default rates change. Finally, investors who utilize the credit default swap market for hedging and/or more aggressive investment strategies do not need to rely on historical average estimates of recovery rates in their calculation of appropriate spread levels, but can be more precise in the critical ultimate recovery estimate.

References

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(3), 787 – 821.

Altman, E., B. Brady, A. Resti, and A. Sironi, 2005, “The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications,” *The Journal of Business*, 78(6), 2203–2228.

Altman, E., N. Fargher, and E. Kalotay, 2009, “Approximations of Equity-Implied Likelihood of Default using Fundamentals and Industry-Level Expectations,” Working Paper.

Altman, E., and B. Karlin, 2010, “Defaults and Returns in the High-Yield Bond and Distressed Debt Market: The Year 2009 in Review and Outlook,” Special Report.

Carey, M., and M. Gordy, 2004, “Measuring Systematic Risk in Recoveries on Defaulted Debt I: Firm-Level Ultimate LGDs,” Working Paper.

Casella, G., and E. I. George, 1992, “Explaining the Gibbs Sampler,” *The American Statistician*, 46(3), 167 – 174.

Das, S. R., and P. Hanouna, 2009, “Implied recovery,” *Journal of Economic Dynamics and Control*, 33(11), 1837 – 1857.

Emery, K., R. Cantor, D. Keisman, and S. Ou, 2007, “Moody’s Ultimate Recovery Database,” *Moody’s Investor Research: Special Comment*, pp. 2–15.

Frye, J., 2000, “Depressing Recoveries,” *Risk*, pp. 106–111.

Hu, Y., and W. Perraudin, 2002, “The Dependence of Recovery Rates and Defaults,” Working Paper.

Keisman, D., and K. Van de Castle, 1999, “Recovering your money: insights into losses from defaults,” *Standard and Poor’s Credit Week*, pp. 16–34.

Koop, G., 2003, *Bayesian Econometrics*. John Wiley and Sons, England.

Merton, R., 1974, “On the Pricing of Corporate Debt: the Risk Structure of Interest Rates,” *Journal of Finance*, 29, 449–470.

Robert, C., 1996, “Mixtures of distributions: inference and estimation,” *Markov chain Monte Carlo in practice, Walter R. Gilks, D. J. Spiegelhalter, CRC Press*, pp. 441–462.

Schuermann, T., 2004, “What do we know about Loss Given Default?,” Forthcoming in D. Shimko (ed.), Credit risk models and management.

Servigny, A. D., and O. Renault, 2004, *Measuring and Managing Credit Risk*. Mc Graw Hill, New York.

Trück, S., S. Harpainter, and S. Rachev, 2005, “A Note on Forecasting Aggregate Recovery Rates with Macroeconomic Variables,” Working Paper.

I Appendix: Estimation Procedures

Section I.1 of this Appendix provides a brief introduction the Gibbs sampler. Section I.2 clarifies the role of Gibbs sampling in the current application, and Section I.3 provides details of the sampling scheme used in estimation throughout the paper.

I.1 Gibbs Sampling

The Gibbs sampler is an algorithm for sampling from the joint distribution of several random variables, given the distribution of each component variable *conditional on all the other variables*. The Gibbs sampler is commonly used in Bayesian applications to estimate the joint distribution of two or more parameters.

Consider a simple two parameter example.

Suppose we want an estimate of the joint posterior density $P(\mu, \sigma|\mathcal{D})$, $\mathcal{D}$ being the data. The Gibbs sampler provides a means of obtaining a numerical estimate of this density if the the full conditional densities $f(\mu|\sigma, \mathcal{D})$ and $f(\sigma|\mu, \mathcal{D})$ are known.

Under appropriate regularity conditions the following algorithm may be utilized:

1. Draw μ from $f(\mu|\sigma, \mathcal{D})$, conditioning on some initial value of σ.
2. Draw σ from $f(\sigma|\mu, \mathcal{D})$, conditioning on the drawn value of μ.
3. Go back to step 1 and draw μ conditioning on the draw of σ, and continue iterating between step 1 and 2 conditioning the draw at each iteration on the previously drawn value of the other parameter.

As the algorithm proceeds and convergence is achieved, draws from the density $P(\mu, \sigma|\mathcal{D})$ obtain.

I.2 The Current Application

Having specified the form of the (predictive) likelihood $g(y|x, \beta, h, \alpha, p)$ in equation (2), our objective is to characterize the posterior distribution of recoveries $g(y|x)$.

In principle, we could do this directly by integration:

$$g(y|x) = \int\int\int\int g(y|x, \beta, h, \alpha, p)p(d\beta, h, \alpha, p|x)d\beta dh d\alpha dp. \tag{I.1}$$

However, given the analytic intractability of the integration (I.1), we utilize the strategy of Gibbs sampling to obtain draws from $P(\beta, h, \alpha, p|x)$ as detailed in Section I.3. Using the draws from $P(\beta, h, \alpha, p|x)$ to condition random draws from the distribution $g(y|x, \beta, h, \alpha, p)$ enables us to generate our sample from the target posterior distribution $g(y|x)$.

In order to implement the Gibbs sampler, we need to know the exact form of each conditional posterior distribution, namely: $P(\beta|h, \alpha, p, x)$, $P(h|\beta, \alpha, p, x)$, $P(\alpha|\beta, h, p, x)$ and $P(p|\beta, h, \alpha, x)$. By specifying proper, minimally-informative prior distributions each parameter has a well-defined conditional posterior distribution, and the parameters of the posterior are data-determined. With such considerations in mind, we employ the conditional posterior distributions detailed in Section I.3.

I.3 The Gibbs Sampling Scheme

1. Draw from $P(\alpha|\beta, h_j, p_j, e_{ij}, y, x) \sim \mathcal{N}(\bar{\alpha}, \bar{V}_\alpha)I(\alpha_1 < \alpha_2 < \ldots < \alpha_m)$, where $\bar{V}_\alpha = \left[V_{\alpha p}^{-1} + \sum_{i=1}^{N}\left\{\sum_{j=1}^{m} e_{ij}h_j\right\} e_i e_i'\right]^{-1}$ and $\bar{\alpha} = \bar{V}_\alpha\left[V_{\alpha p}^{-1}\alpha_p + \sum_{i=1}^{N}\left\{\sum_{j=1}^{m} e_{ij}h_j\right\} e_i(y_i - x_i'\beta)\right]$, given a Normal conjugate prior on α with mean α_p and precision $V_{\alpha p}$, subject to the labeling restriction reflected in the indicator function $I(\alpha_1 < \alpha_2 < \ldots < \alpha_m)$ that is equal to one when the prior restriction is true, and zero otherwise. The notation e_i denotes an m-vector of indicators wherein one of the elements j is equal to unity, indicating that error i is attributed to the corresponding mixture component. In specifying the prior, α_p is set equal to an m-vector of zeros and $V_{\alpha p}$ is a diagonal matrix with large numbers on the diagonal (10^5), consistent with an absence of non-sample information about the parameters of mixture components.

2. Draw from $P(\beta|\alpha, h_j, p_j, e_{ij}, y, x) \sim \mathcal{N}(\bar{\beta}, \bar{V})$, where $\bar{V} = \left[V_{\beta p}^{-1} + \sum_{i=1}^{N}\sum_{j=1}^{m} e_{ij}h_j x_i x_i'\right]^{-1}$ and $\bar{\beta} = \bar{V}\left[V_{\beta p}^{-1}\beta_p + \sum_{i=1}^{N}\sum_{j=1}^{m} e_{ij}h_j x_i(y_i - \alpha_j)\right]$, given a Normal conjugate prior β with mean β_p and precision $V_{\beta p}$. In specifying the prior, β_p is equal to a k-vector of zeros, and $V_{\beta p}$ is a diagonal matrix with large numbers on the diagonal, consistent with an absence of non-sample information about the parameters of mixture components.

3. Draw from $P(h_j|\beta, \alpha, p_j, e_{ij}, y, x) \sim \mathcal{G}(\bar{s}_j^2, \bar{v}_j)$, a Gamma distribution, where $\bar{s}_j^2 =$

$\frac{\sum_{i=1}^{N} e_{ij}(y_i-\alpha_j-x_i\beta)^2+v_{jp}s_{jp}^2}{\bar{v}_j}$ and $\bar{v}_j = \sum_{i=1}^{N} e_{ij} + v_{pj}$ and e_{ij} is an indicator taking a value of one if observation i is assigned to mixture component j, and zero otherwise. The prior on h_j is assumed to be Gamma, with $v_{jp} = 0.01$ and $s_{jp} = 1$.

4. Draw from $P(p_j|h_j, \beta, \alpha, e_{ij}, y, x) \sim \mathcal{D}(\bar{\rho})$, a Dirichlet distribution, where $\bar{\rho} = \left[\rho_p + \sum_{i=1}^{N} e_i\right]$ such that e_i is an m-vector of indicators wherein one of the elements j is equal to unity, indicating that error i is attributed to the corresponding mixture component. The corresponding prior vector ρ_p is an m-vector of ones.

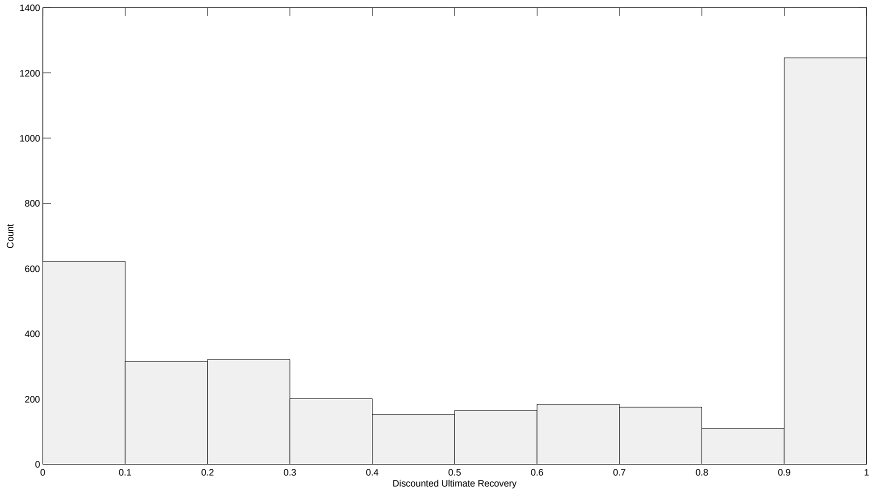
5. Draw e_i from
$P(e_{ij}|h_j, \beta, \alpha, p_j, y, x) \sim \mathcal{M}\left[1, \left(\frac{p_1 f_{\mathcal{N}}(y_i|\alpha_1+x_i'\beta, h_1^{-1})}{\sum_{j=1}^{m} p_j f_{\mathcal{N}}(y_i|\alpha_j+x_i'\beta, h_j^{-1})}, \ldots, \frac{p_m f_{\mathcal{N}}(y_i|\alpha_m+x_i'\beta, h_m^{-1})}{\sum_{j=1}^{m} p_j f_{\mathcal{N}}(y_i|\alpha_j+x_i'\beta, h_j^{-1})}\right)\right]$,
a Multinomial distribution where $f_{\mathcal{N}}(.)$ denotes the normal likelihood value.

The Gibbs sampling scheme involves drawing successively from each of the distributions in steps 1-5, updating at each step the values of the conditioning parameters to be used in the subsequent draws. Looping over steps 1-5 many times results in outcomes from the joint posterior $P(\alpha, \beta, h_j, p_j, e_{ij}|y, x)$. The sampling scheme described in steps 1-5 is estimates the mixture describing y conditional on x. Estimation of the joint posterior based on y alone (without conditioning on x) proceeds analogously, however step 2 is eliminated and the conditioning in each component of the sampling scheme is adjusted accordingly.

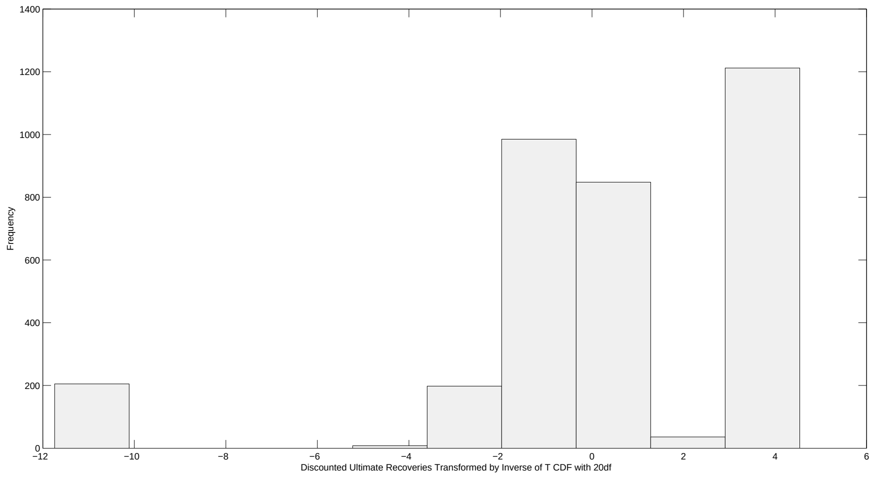
Unless stated otherwise, results reported in this paper are based on 10,000 iterations of the Gibbs sampler, after 100 burn-in draws are discarded.

II Tables and Figures

Figure 1: Sample Distribution of Moody's Discounted Ultimate Recoveries on Loans and Bonds

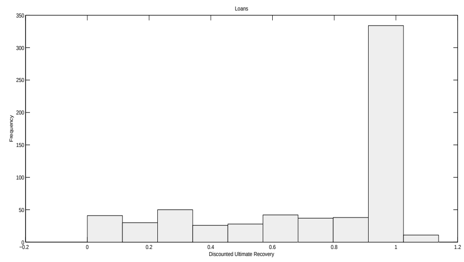


(a) Histogram of Moody's Discounted Ultimate Recovery. Refer to Section 2 for data description. NOTE: The recoveries exceeding 100% are rounded prior to transformation and estimation.

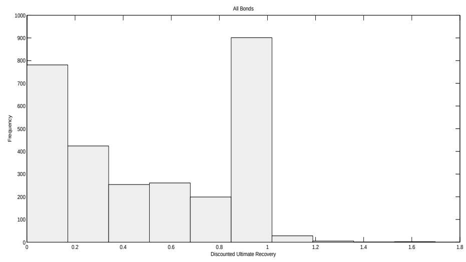


(b) Histogram of Moody's Discounted Ultimate Recovery after transformation through the Inverse T CDF (20df). Refer to Section 3 for further information.

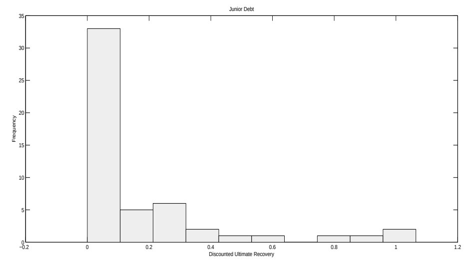
Figure 2: Sample Distribution of Moody's Discounted Ultimate Recoveries: by Type



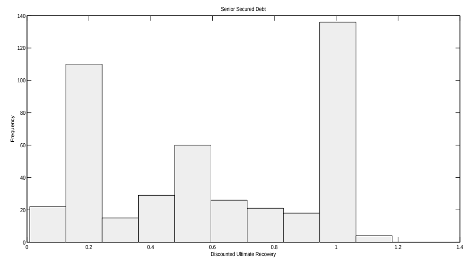
(a) Histogram of Moody's Discounted Ultimate Recovery: Loans Only.



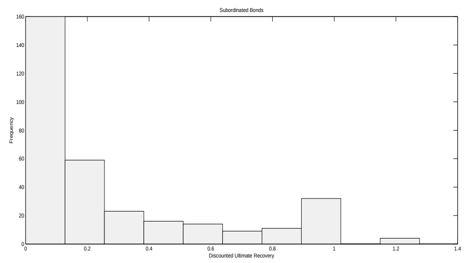
(b) Histogram of Moody's Discounted Ultimate Recovery: Bonds Only.



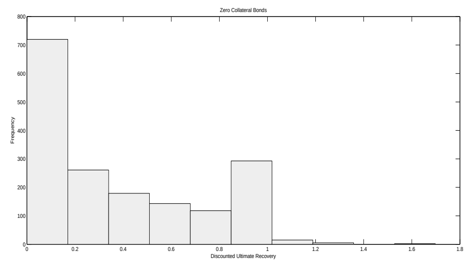
(c) Histogram of Moody's Discounted Ultimate Recovery: Junior Debt Only.



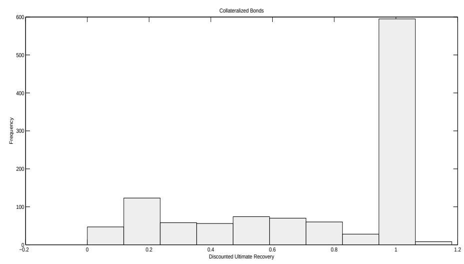
(d) Histogram of Moody's Discounted Ultimate Recovery: Senior Secured Bonds Only.



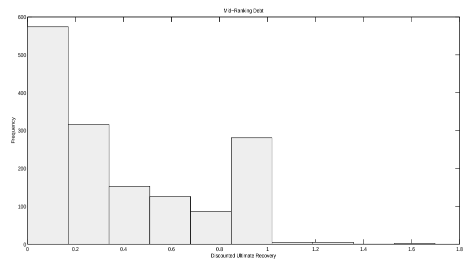
(e) Histogram of Moody's Discounted Ultimate Recovery: Subordinated Bonds Only.



(f) Histogram of Moody's Discounted Ultimate Recovery: Zero Collateral Bonds Only.



(g) Histogram of Moody's Discounted Ultimate Recovery: Collateralized Bonds Only.



(h) Histogram of Moody's Discounted Ultimate Recovery: Debt with Rank 2 or 3 Only.

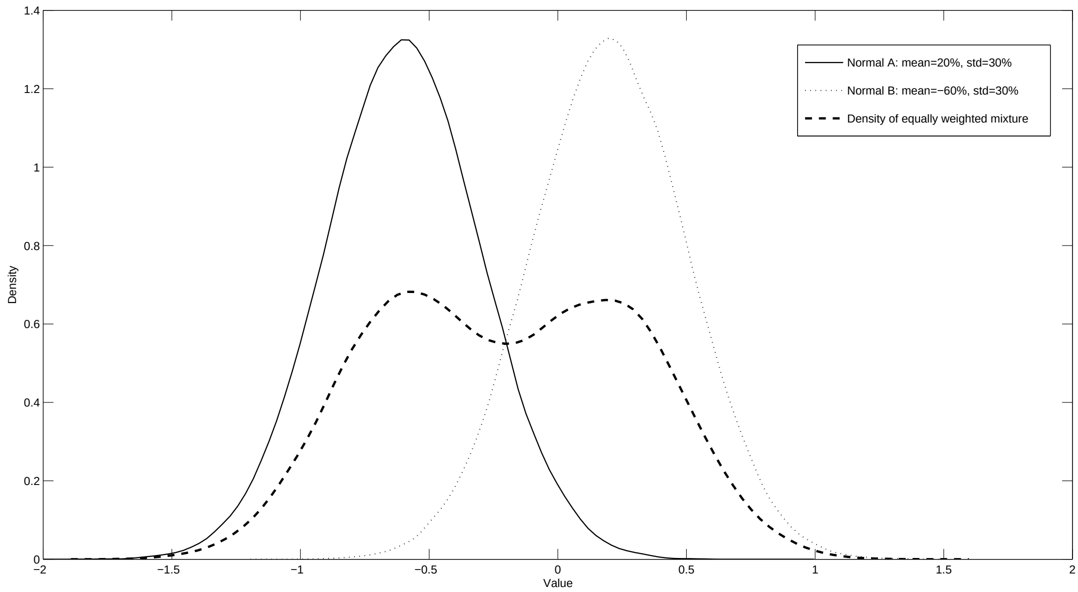
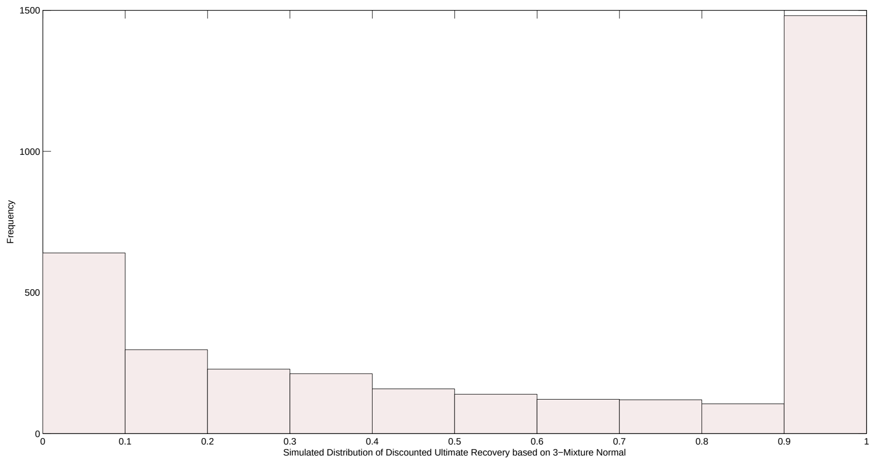
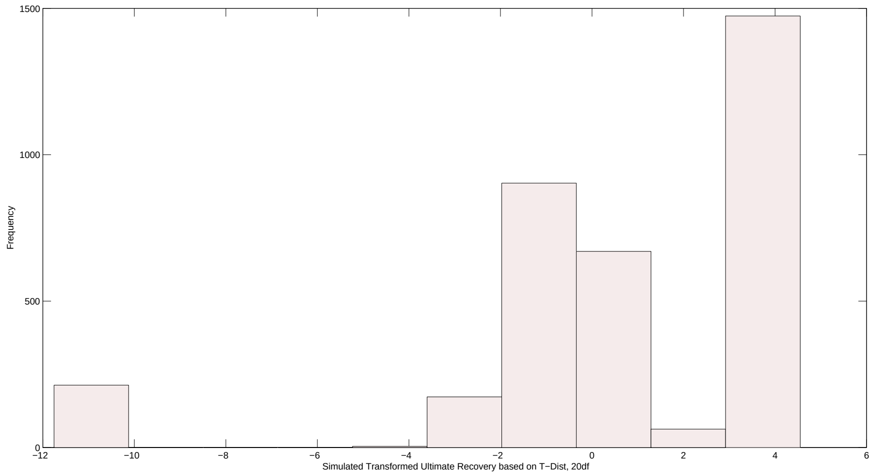


Figure 3: **A 2-Normal Mixture.** The (dashed) mixture density obtained by drawing with equal probability from two normals with differing means.

Figure 4: Approximation of Discounted Ultimate Recoveries on Loans and Bonds using 3-Normal Mixture (Unconditional)



(a) Discounted ultimate recoveries based on simulated sample of 3500



(b) Simulated Inverse T transformed Discounted Ultimate Recoveries, sample of 3500.

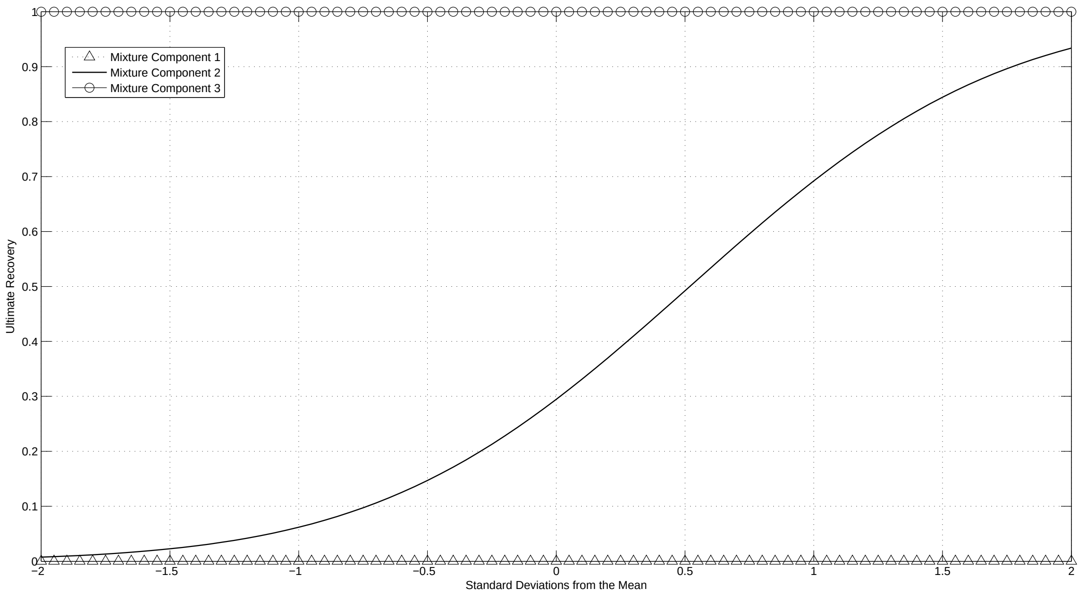


Figure 5: **Mapping Mixture Components to Ultimate Recoveries** This figure maps percentiles of draws from each mixture component to the recovery implied by the CDF of a Student-T with $v = 20$. Triangles denote the first mixture component, a solid line the second, and circles the third. The results are based on 10,000 draws from the sampling scheme described in Appendix I.3.

Table 1: Discounted Ultimate Recoveries: Descriptive Statistics

This table summarizes the characteristics of Moody's discounted ultimate recoveries for loans, bonds, and the categories of exposures presented in Figure 2. "Sub" is subordinated and "Coll" is collateral, "No Coll." is no collateral, "Rank 2 or 3" refers to priority ranking amongst debt claimants. IQR refers to interquartile range, 10% and 90% to the respective percentiles of the distribution.

	Loans	*Bonds*	*Junior*	*Senior Sec*	*Sub*	*Coll. Bonds*	*No Coll.*	*Rank 2 or 3*	*Pooled*
Mean	0.75	0.52	0.16	0.59	0.27	0.74	0.38	0.39	0.56
Median	1.00	0.49	0.04	0.57	0.14	1.00	0.24	0.24	0.58
Std	0.33	0.39	0.26	0.34	0.34	0.33	0.37	0.37	0.39
IQR	0.53	0.85	0.22	0.79	0.42	0.53	0.64	0.61	0.82
10%	0.20	0.01	0.00	0.19	0.00	0.21	0.00	0.00	0.02
90%	1.00	1.00	0.52	1.00	0.98	1.00	1.00	1.00	1.00
N	637	2855	52	441	328	1119	1777	1549	3492

Table 2: Information Criteria for Varying m (Unconditional Case)

The value of each information criterion is evlauated at the posterior mean of the parameters using 10,000 draws from variants of the sampling scheme described in Appendix I.3. AIC is the Akaike Information Criteria, BIC is the Bayesian Information Criteria and HQ is the Hannan-Quinn criteria. Bold type indicates preferred specification. Will add calculation details here.

	# of Mixture Components Transformation: Inverse of Standard Normal CDF				
Criteria	2	3	4	5	6
AIC	-7976	8243	8275	8319	**8325**
BIC	-8013	8187	8202	**8227**	8214
HQ	-8002	8204	8224	**8255**	8248
	# of Mixture Components Transformation: Inverse of Student-t, $v = 20$				
AIC	-9089	**7967**	7965	-	-
BIC	-9126	**7911**	7891	-	-
HQ	-9115	**7928**	7914	-	-

Table 3: Simulated Unconditional Distribution Characteristics based on $m = 3$

This table summarizes the characteristics of the 3-mixture approximation of the discounted ultimate recoveries distribution for loans, bonds, and the categories of exposures presented in Figure 2. "Sub" is subordinated and "Coll" is collateral, "No Coll." is no collateral, "Rank 2 or 3" refers to priority ranking amongst debt claimants. IQR refers to interquartile range, 10% and 90% to the respective percentiles of the distribution. p1, p2 and p3 are the posterior mean probabilities associated with mixture components 1, 2 and 3 respectively. ND is the number of draws from the posterior distribution.

	Loans	*Bonds*	*Junior*	*Senior Sec*	*Sub*	*Coll. Bonds*	*No Coll.*	*Rank 2 or 3*	**Pooled**
Mean	0.72	0.57	0.44	0.55	0.49	0.71	0.49	0.52	**0.593**
Median	1.00	0.62	0.28	0.54	0.43	1.00	0.44	0.49	**0.657**
Std	0.38	0.40	0.43	0.38	0.43	0.37	0.39	0.40	**0.399**
IQR	0.65	0.84	1.00	0.82	0.99	0.64	0.90	0.89	**0.822**
10%	0.07	0.02	0.00	0.06	0.00	0.08	0.00	0.00	**0.03**
90%	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
p1	0.02	0.07	0.27	0	0.23	0.01	0.11	0.10	**0.06**
p2	0.4	0.63	0.69	0.69	0.67	0.38	0.73	0.73	**0.59**
p3	0.54	0.3	0.04	0.31	0.10	0.62	0.16	0.17	**0.35**
ND	1000	1000	1000	1000	1000	1000	1000	1000	**3500**

Table 4: Mixture Components: Summary Statistics

Summary statistics associated with the mixture components graphed in Figure 5. The posterior mean of each component $E(\alpha_i|y)$, the posterior mean of each respective standard deviation $E(h_i|y)$ and the posterior mean and variability of the mixture weights are computed using 10,000 draws from the sampling scheme described in Appendix I.3.

Mixture Component			
$i =$	1	2	3
Component Distributions: Normal Parameters			
$E(\alpha_i|y)$	-11.74	-0.55	4.54
$E(h_i|y)$	0.007	1.06	0.003
Component Distributions: Mapped to Recoveries			
Mean Recovery	0%	35.18%	100%
Median Recovery	0%	30.20%	100%
IQR	0%	46.40%	0%
Std Dev	0%	27.50%	0%
Unconditional Posterior Probability Weights			
$E(p_i|y)$	6%	59%	35%
$\sigma(p_i|y)$	0.4%	0.80%	1%

Table 5: Simulated Distribution Characteristics based on $m = 3$, Conditional on Industry Distress State

This table summarizes the characteristics of the 3-mixture approximation of the discounted ultimate recoveries distribution for loans, bonds, and the categories of exposures presented in Figure 2. "Sub" is subordinated and "Coll" is collateral, "No Coll." is no collateral, "Rank 2 or 3" refers to priority ranking amongst debt claimants. IQR refers to interquartile range, 10% and 90% to the respective percentiles of the distribution. p1, p2 and p3 are the posterior mean probabilities associated with mixture components 1, 2 and 3 respectively. ND is the number of draws from the posterior distribution. "Industry distress $\geq +1.7\sigma$" refers to draws from the posterior distribution weighted by mixture probabilities conditional on expectations of industry default at the time of default being 1.7 or more standard deviations above its time series mean. "Industry distress < Median" refers to mixture probabilities derived from recoveries on exposures that defaulted when the applicable industry default expectation was below its time series median.

Distribution Characteristics based on $m = 3$, Industry Distress $\geq +1.7\sigma$

	Loans	*Bonds*	*Junior*	*Senior Sec*	*Sub*	*Coll. Bonds*	*No Coll.*	*Rank 2 or 3*	**Pooled**
Mean	0.65	0.53	0.37	0.43	0.48	0.63	0.47	0.47	**0.53**
Median	0.95	0.49	0.00	0.35	0.42	0.77	0.37	0.38	**0.49**
Std	0.39	0.39	0.48	0.34	0.41	0.39	0.39	0.39	**0.39**
IQR	0.76	0.85	1.00	0.59	0.95	0.76	0.88	0.85	**0.87**
10%	0.04	0.02	0.00	0.04	0.00	0.05	0.00	0.00	**0.01**
90%	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
p1	0.06	0.07	0.62	0.00	0.18	0.02	0.12	0.11	**0.07**
p2	0.50	0.71	0.38	0.87	0.70	0.55	0.78	0.75	**0.67**
p3	0.44	0.22	0.00	0.13	0.13	0.44	0.11	0.13	**0.26**
ND	1000	1000	1000	1000	1000	1000	1000	1000	**1000**
N	181	768	13	112	56	445	504	476	**949**

Distribution Characteristics based on $m = 3$, Industry Distress < Median

	Loans	*Bonds*	*Junior*	*Senior Sec*	*Sub*	*Coll. Bonds*	*No Coll.*	*Rank 2 or 3*	**Pooled**
Mean	0.73	0.62	0.43	0.65	0.48	0.74	0.53	0.54	**0.66**
Median	1.00	0.79	0.33	0.84	0.38	1.00	0.50	0.51	**0.99**
Std	0.36	0.40	0.40	0.38	0.42	0.36	0.41	0.40	**0.40**
IQR	0.60	0.82	0.86	0.71	0.98	0.57	0.87	0.87	**0.75**
10%	0.10	0.02	0.00	0.07	0.00	0.10	0.00	0.00	**0.03**
90%	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
p1	0.01	0.06	0.19	0.00	0.20	0.01	0.10	0.09	**0.05**
p2	0.41	0.55	0.77	0.54	0.67	0.39	0.67	0.67	**0.53**
p3	0.58	0.39	0.04	0.46	0.12	0.60	0.23	0.24	**0.42**
ND	1000	1000	1000	1000	1000	1000	1000	1000	**1000**
N	305	1414	26	234	177	889	830	750	**1719**

Table 6: Simulated Distribution Characteristics based on Linear Regression with $m = 2$

This table summarizes the posterior mean and standard deviation of the regression coefficients where errors are captured by a mixture of 2 normals. Note that the dependent variable is Moody's discounted ultimate recovery mapped to the real number line by the inverse of a Student-T CDF with $v = 20$. DC is Debt Cushion, DT is time in default, "IDLI" is mean industry level default likelihood observed at the time of default. The notation $\bar{\beta}$ denotes the posterior mean of the regression coefficient(s) and $\bar{\sigma}$ the posterior standard deviation.

Intercept	$\bar{\beta}$	$\bar{\sigma}$
α_1 $(p = 0.059)$	-7.69	0.81
α_2 $(p = 0.941)$	3.68	1.88
Debt Cushion	$\bar{\beta}$	$\bar{\sigma}$
$DC < 0.25$	-2.01	0.10
$0.25 \leq DC < 0.5$	-1.65	0.10
$0.5 \leq DC < 0.75$	-0.17	0.10
Time in Default	$\bar{\beta}$	$\bar{\sigma}$
1yr $< DT \leq$ 2yr	-0.19	0.07
2yr $< DT \leq$ 3yr	-0.21	0.10
3yr $< DT$	0.79	0.10
Rank & Collateral	$\bar{\beta}$	$\bar{\sigma}$
Rank = 2	-0.44	0.08
Rank = 3	-0.61	0.11
Rank $\geq$ 4	-0.82	0.15
Collateral *(Yes = 0)*	-0.61	0.13
Default Type (Bankruptcy = 0)	$\bar{\beta}$	$\bar{\sigma}$
Default & Cure	3.76	0.36
Other Restructure	1.64	0.10
Distressed Exchange	1.72	0.47
Seniority (Loan = 0)	$\bar{\beta}$	$\bar{\sigma}$
Junior Subordinated	-0.80	0.22
Revolver	0.44	0.09
Senior Secured	-0.68	0.12
Senior Subordinated	-1.00	0.16
Senior Unsecured	-0.27	0.14
Subordinated	-0.76	0.15
Industry Default Likelihood	$\bar{\beta}$	$\bar{\sigma}$
IDLI	-0.39	0.03

Table 7: Simulated Distribution Characteristics based on Linear Regression with $m = 2$

This table is the analogue of Table 3 based on the 2-Mixture conditional specification summarized in Table 6.

Distribution Characteristics based on Linear Regression with $m = 2$

	Loans	*Bonds*	*Junior*	*Senior Sec*	*Sub*	*Coll. Bonds*	*No Coll.*	*Rank 2 or 3*	**Pooled**
Mean	0.80	0.58	0.32	0.67	0.36	0.81	0.46	0.47	0.63
Median	0.98	0.70	0.10	0.85	0.15	0.41	0.42	0.98	0.82
Std	0.31	0.40	0.38	0.36	0.39	0.40	0.40	0.30	0.40
IQR	0.26	0.87	0.71	0.65	0.78	0.89	0.88	0.24	0.82
10%	0.18	0.00	0.00	0.05	0.00	0.00	0.00	0.24	0.00
90%	1.00	1.00	0.97	1.00	0.98	1.00	1.00	1.00	1.00
p1	0.02	0.07	0.27	0.00	0.23	0.11	0.11	0.01	0.06
p2	0.98	0.93	0.73	1.00	0.77	0.89	0.89	0.99	0.94
ND	3000	3000	3000	3000	3000	3000	3000	3000	-
N	637	2855	52	441	328	1777	1549	1714	3492

Table 8: Portfolio Recoveries

This table summarizes percentiles of the simulated portfolio recovery distribution and its various approximations. The methodology used to generate these distributions is detailed in Section 5.5. The posterior percentiles are computed based 10,000 draws and the hypothetical portfolio size is 150 \$1.00 exposures.

	Actual	*3-Mix Comp*	*3-Mix Char*	*3-Mix Base*	*Regression*	*2-Mix Reg*	*Beta Comp*
Mean	84.20	84.83	79.57	89.29	79.59	93.78	75.99
Median	84.25	84.81	79.59	89.36	79.55	93.72	75.94
Std	4.77	4.85	4.79	4.89	5.39	4.90	4.78
IQR	6.38	6.53	6.40	6.68	7.37	6.61	6.44
Lower Tail Percentile Cut-Offs							
5%	76.18	76.88	71.68	81.27	70.72	85.61	68.06
2%	74.39	74.91	69.72	79.26	68.38	83.73	66.22
1%	73.12	73.79	68.25	78.05	66.88	82.24	65.09
0.50%	71.98	72.52	67.16	76.77	65.89	80.89	63.65
0.10%	69.37	69.58	63.94	74.03	63.04	77.66	60.97

Table 9: Mixture Probabilities based on Debt Cushion and Industry Default Conditions

This table summarizes mixture probabilities conditional on category of Debt Cushion DC and mean industry default expectation observed at the time of default.

		Debt Cushion (DC)		
		$DC < 0.25$	$0.25 \leq DC < 0.75$	$0.75 \leq DC$
Industry Distress $\geq +1.7\sigma$	p1	0.10	0.01	0.02
	p2	0.79	0.68	0.33
	p3	0.11	0.31	0.65
Industry Distress $< +1.7\sigma$	p1	0.09	0.01	0.00
	p2	0.71	0.62	0.19
	p3	0.20	0.38	0.81