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

Edward Altman* Egon Kalotay†

September 30, 2010

*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: egon.kalotay@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.

Abstract

We present an intuitive Bayesian approach to modeling the distribution of recoveries on defaulted debt by accommodating the distinguishing features of ultimate recovery outcomes using mixtures of distributions. Our empirical findings show that expectations of industry-level default conditions at the time of default and a measure of debt subordination (the Debt Cushion) capture economically important variation in the shape of recovery distributions. We demonstrate how to adapt estimation results to target portfolios whose characteristics do not match those of the estimation sample. We show that simple mixture-based estimates of recovery distributions, adapted to reflect the Debt Cushion of defaulted exposures and industry distress conditions at the time of default, substantially outperform popular alternatives in forecasting recoveries on portfolios of defaulted debt.

JEL Classification: G10, G20, G21, & G29

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

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 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 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 effective 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, although popularly applied, estimation techniques tend to both systematically 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.

In the next section we we first provide a brief overview the empirical 'facts' about losses in the event of default gleaned from prior empirical studies of such outcomes, and we describe the data in section 3. In Section 4 we describe the rationale of our approach to modeling defaulted debt recovery outcomes as mixtures of distributions. We discuss specific econometric issues in Section 5. In Section 6 we present our empirical findings including a comparison of portfolio loss estimates obtained using mixture models to those derived from popular alternatives.

2 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, known as the Debt Cushion, also seems to matter. The analysis ofKeisman and Van de Castle (1999) suggests that all else equal, the larger the Debt Cushion, the higher the recovery.

3. Recoveries tend to be lower in recessions or when the rate of aggregate defaults is high. 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. Further,

Frye (2000b) shows that a 10% realized default rate resulting in a 25% reduction in recoveries relative to its normal year average.

4. Industry matters. Acharya, Bharath, and Srinivasan (2007) suggest that macroeconomic conditions do not appear to be significant determinants of individual bond recoveries after accounting 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 apparent. The 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, or probability masses near zero and 1 – important features, as we will show, of the empirical distribution of recoveries.

Motivated by these considerations, we present a novel approach to modeling the distribution of recoveries on defaulted loans and bonds based on mixtures of normal distributions. As we will demonstrate shortly, not only is our approach flexible enough to accommodate the empirical features of recovery distributions, it is also adaptable in terms of the manner in which the 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.

3 Ultimate Recoveries on Defaulted Loans and Bonds: Data Description

We use discounted ultimate recoveries from Moody's Ultimate Recovery Database. 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 calculate 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, which is 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. We focus on features of the data relevant to understanding whether our sample is consistent with some of the stylized facts listed earlier, asEmery, Cantor, Keisman, and Ou (2007) provides a very detailed description of the database.

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), which was 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.

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, thus greatly boosting the relative proportion of high recoveries.

The issue of variation in 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 4 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 4, it is apparent that the recoveries associated with the highest ranking collateralized claims, loans and collateralized bonds, are highly skewed. The distributions exhibit high variability but little bi-modality. At the other extreme, the distributions of recoveries on junior bonds and subordinated bonds exhibit strong right skewness. There is however some suggestion of bi-modality in the subordinated bond distribution.

Panels (b), (d), (f) and (h) of Figure 4 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 in that recovery distributions associated with the highest ranking, collateralized exposures, and the lowest ranked bond exposures tend to be highly skewed and closer to unimodal. However, 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 4. Loans and collateralized bond exposures at one extreme, and junior debt at the other are skewed in opposing directions. Loans and collateralized bond exposures are 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 is *a-priori* consistent with the data, as well the need for caution in interpreting the concept of average recovery.

4 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.

Specifically, our modeling approach rests on the assumption that recovery outcomes y can be thought of as draws from a 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 with a high degree of accuracy. The tractability of the mixture components implies that properties of the distribution that are relevant to inference obtain quite easily.

As shown in Figure 2, the simplest 2-normal mixture, which is obtained by drawing with equal probability from two normal distributions of differing mean but equal standard deviation, can be used to approximate a symmetrical bimodal 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 simple because we know the number of normals required to simulate the density of interest, as well as the parameters of the simulation. 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, for instance, the Markov Chain Monte Carlo (MCMC) technique of Gibbs sampling.

Before describing our approach in more detail, we note that fully non-parametric procedures can also accommodate the empirical features of recoveries data, as shown by Renault and Scaillet (2004). In particular, Renault and Scaillet (2004) estimate recovery distributions using the beta kernel density estimator of Chen (1999). However, our mixture-based approach has a number of advantages, including: relative simplicity, ease of inference and the economic interpretability of model outputs. Further, while we use the entire sample in estimation, we can easily adapt our findings to portfolios or economic conditions that differ from the estimation sample.

5 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, which are 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:

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

where α_j is the mean of mixture component j conditional on a linear combination of x, and its variance h_j. 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. We estimate the joint posterior of all parameters using the Markov chain Monte Carlo (MCMC) technique of Gibbs Sampling.[6] However, in order to accomplish this, two problems must be addressed.

First, there are no directly observable data 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 identification problem is circumvented by way of a labeling restriction. 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 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 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 in more detail shortly, it enables us to compute the posterior probability with which observations, or 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.[12] Throughout this paper we use the inverse CDF of the Student-T with $v = 20$. The rationale for this modeling choice is detailed in Section 4 of the Technical Appendix.

6 Empirical Estimates of Recovery Distributions

Modeling the pooled distribution of transformed recoveries as a mixture of normals, we first summarize how well our estimates capture the empirical features of the sample. Comparing Panels (a) and (b) of Figure 3 with the corresponding panels of Figure 1 provides a visual summary of overall success. We plot in Figure 3 the distribution of 3,500 recoveries simulated from the 3-component mixture. Both the raw recoveries and the corresponding distribution of transformed recoveries appear closely matched to their sample counterparts.[13]

The degree of correspondence between the properties of the observed sample and the posterior simulation can be quantified by comparing the sample characteristics summarized in Table 1 to those of simulated sample reported in Table 2. Focusing on the pooled sample in the final column of table, 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 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.

Both graphical and numerical summaries of the sample obtained through posterior simulation suggest the mixture of three normals captures the important empirical features of the observed pooled sample. Before considering the results at a more granular level, accounting for important conditioning information, we clarify the economic interpretation of the mixture components.

6.1 Economic Interpretation of the Mixture Components

The posterior distribution of the mixture parameters have a distinct economic interpretation. Our results imply that the pooled distribution of ultimate recoveries can be thought of as a draw from one of three distributions. The parameters are presented in the middle panel of Table 3. 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.

Note that the distributions documented in the second panel of Table 3 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 non-linear impact of deviations from the mean 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 – indicating 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.

6.2 Inferring the Effects of Conditioning Information

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 3. 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 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%. At the same time, 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 that are 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. The 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 these effects using the output of the Gibbs Sampler, we make use of the latent data discussed in Section 5 and generated as per step 5 of the algorithm described in Section 3 of the Technical Appendix.

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 $i = 1 \ldots N$ 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. Further, suppose that the debt exposures $i \in Q$ denote the sub portfolio of interest – recoveries on subordinated debt. Then, p_{Qj}, the mixing weight for portfolio Q associated with component j, can be estimated from the Gibbs output using

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

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 to the unconditional weights reported in Table 3.

6.2.1 Facility Characteristics

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 2 summarizes the properties of recoveries using 10,000 draws from the 3-component mixture. Comparing the properties of the modeled distributions in Table 2 with those of the corresponding empirical samples in Table 1, we find that the mixture weights differ across sub-classes in accordance with economic 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 effectively 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 and 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 from others then we could model them separately, or specify a variant of model (2) incorporating a set of indicators for these exposures prior to estimation.

6.2.2 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 4 the analogue of the results in Table 2, conditional on industry distress at the time of default. That is, we construct the mixtures in Table 4 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 4 and the corresponding estimates in Table 2 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 of the hypothesis, that is, the probability of high recoveries drops substantially in years where expectations of default risks are elevated at the time of default.

Overall, in the event of industry distress at the time of default, the probability of drawing a recovery outcome from the third mixture component (implying full recovery) declines from 35% in Table 2 to 26% in Table 4, but the probability of drawing from the first mixture component (implying total loss) increases only slightly from 6 to 7%. Similar effects are observable in sub-categories of exposures, such as non-collateralized bonds where 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 4 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 where 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 where 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.

6.3 Assuming Linearity: Regression Estimates

Altman, Brady, Resti, and Sironi (2005) and others have used 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, these 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 5 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 5 reports the posterior mean of the regression model estimates. While the sign or 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 in light of the discussion in Section 6.1. 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 5 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 $m = 3$ without explicit 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 6 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.

7 Modeling Out-of-Sample 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 within-sample performance. In this final section of our results we present an out-of-sample measure of model fit in terms of a stylized application of the alternative modeling approaches. Specifically, we benchmark the ability of each approach to approximate the left-hand tail of the recovery distribution. To do this we conduct an out-of-sample simulation study using the following procedure.

1. We split the sample. All observations up to and including 2001 are set aside as an estimation sample. The remaining 5 years of the sample comprise the test period. In this way, 1477 observations are used for estimation and the remaining 2015 are used for testing.

2. We draw a random sample of 150 recoveries on defaulted loans and bonds from the test sample and compute the ultimate recovery on an equally-weighted portfolio of the selected exposures. This value is stored as an outcome of the empirical loss distribution.

3. We then draw and store a portfolio loss outcome from each of the following models, estimated using the estimation sample.

    (a) *3-Mix Comp:* is based on the components of the 3-mixture unconditional estimation. However, in this case the mixture components are re-weighted based on the Debt Cushion associated with the sampled exposures as well as the expectations of default in the borrowers' industries at the time of default. This is done in accordance with the probabilities summarized in Table 7.

(b) *3-Mix Char:* is a draw from a mixture wherein the components are re-weighted to reflect the Debt Cushion and Collateralization characteristics of the portfolio sampled from the test period. The Debt Cushion categories used for the purposes of this exercise correspond to the column categories in Table 7. This distribution does not incorporate any prior knowledge about industry-level default expectations at the time of default.

(c) *3-Mix Base:* is a draw from the posterior recovery distribution based on the 3-mixture specification. This distribution is estimated using the pooled sample recovery outcomes observed prior to 2002. The distribution reflects estimation risk in parameter estimates but does not incorporate any prior knowledge about the exposure, facility or industry-level default expectations.

(d) *Regression:* is a draw from the posterior distribution of recoveries implied by a single mixture regression model incorporating all the the conditioning variables reported in Table 5. 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 variant of the regression specification reported in Table 5. 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 subset of outcomes observed during the estimation period that match the sampled observations in terms of their Debt cushion category and industry default expectations at the time of default. The parameters of the Beta distribution corresponding to each of the categories in Table 7 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* insofar as the same conditioning information is used for modeling. Given the popularity of the Beta distribution in modeling default losses, it serves as an interesting benchmark of practical significance.

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

Before considering the results, we summarize the three-fold purpose of the simulation study as follows. First, we use it to benchmark the out-of-sample performance of our mixture-based approach relative to widely used alternatives: regression and calibrated Beta distributions in particular. Second, in evaluating the ability of each modeling approach to approximate the lower tail of portfolio recovery distributions we provide an economic rather than purely statistical measure of modeling success. Third, the simulation study serves to demonstrate our approach to adapting mixture estimates to the modeling of target portfolios whose characteristics differ from those of the sample used for estimation.

With the objectives of the exercise in mind, we turn to the results in Table 8. The tail percentiles of the predictive distribution of portfolio losses generated using *3-Mix Comp* approximate the percentiles of the actual recovery distribution with substantially greater accuracy than other approaches. The absolute error associated with the tail estimates obtained using *3-Mix Comp* range from 3.4% to 4.4%. At the other end of the spectrum, the largest absolute predictive errors, which range from 16.3% to 18%, are those associated with the calibrated Beta distribution *Beta Comp*. The performance of *3-Mix Comp* and *Beta Comp* is directly comparable in the sense that both sets of forecasts are conditioned on the same information: Debt Cushion and expectations of industry distress at the time of default.

The regression-based forecasts of recovery, *Regression* and *2-Mix-Reg*, account for the broadest range of borrower and facility level characteristics by assuming a linear relation between transformed recoveries and the conditioning variables. Assuming that the errors are normally distributed yields the best predictive performance with tail approximation errors ranging from 5.8% to 7.0%.The regression-based errors are up to twice the magnitude of those obtained using *3-Mix Comp* – forecasts obtained using a much narrower set of conditioning information.

Comparing across the three set of mixture-based forecasts *3-Mix Comp* , *3-Mix Char* and *3-Mix Base* demonstrates the importance of re-weighting the mixture components such that both default expectations and facility level characteristics are captured. The surprising aspect of our results is the finding that the combination of only two conditioning variables, Debt Cushion and industry-level default expectations, do remarkably well in both absolute and relative terms. The variables were chosen because they each represent composite sources of information.

Debt cushion as 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 (2000a) 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 (2010) 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 our out-of-sample forecasting experiment suggest an important role for both Debt Cushion and industry default expectations in modeling portfolio credit losses as the combination of the two variables provide in our application a simple basis for adapting the mixture probabilities to the characteristics of the target portfolios. However, it is quite possible that 'better' results are obtainable by expanding the set of conditioning variables.[17]

8 Summary and Conclusions

We have presented 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 using a mixture of three normal distributions. Not only do the mixture components and the associated probability weights have an intuitive economic interpretation, they better accommodate the features of empirical recovery distributions than popularly applied alternatives.

In most practical modeling applications the characteristics of the sample used for modeling do not match those of the portfolio of interest. We address this problem by making use of latent data that is generated in the course of estimation to adapt the estimation results to the target portfolio of interest. In so doing, we make full use of the sample for estimation and we avoid making strong assumptions about the form of the relation between conditioning information and the mean outcome. Using information about the Debt Cushion of defaulted exposures and industry-level expectations of default at the time of default enables us to model the distribution of recoveries with remarkable accuracy – both within and out-of-sample.

It follows that our empirical results are of interest on two levels. First, our findings complement and extend those of earlier studies. For example, 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. Altman, Brady, Resti, and Sironi (2005) and Acharya, Bharath, and Srinivasan (2007) focus on the relation between realized aggregate default rates and mean recoveries, while the findings of Das and Hanouna (2009) are based on the relation between default and recovery expectations inferred from CDS spreads. Second, the potential practical utility of modeling recoveries using a simple mixture of distributions is illustrated by the results of our out-of-sample portfolio modeling study. Our mixture-based estimates, adapted to reflect the Debt Cushion and industry-specific distress conditions at the time of default, substantially out-perform popular competing alternatives.

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 when they are considering investments, including those whose market values are at or near par value, given their probability of default. 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, 2010, "A Simple Empirical Model of Equity Implied Probabilities of Default," working paper, Macquarie University.

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, Salomon Center, New York University.

Carey, M., and M. Gordy, 2004, "Measuring Systematic Risk in Recoveries on Defaulted Debt I: Firm-Level Ultimate LGDs," working paper, Federal Reserve Board, Washington.

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

Chen, S. X., 1999, "Beta kernel estimators for density functions," *Computational Statistics & Data Analysis*, 31(12), 131–145.

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., 2000a, "Collateral Damage," *Risk*, pp. 91–94.

———, 2000b, "Depressing Recoveries," *Risk*, pp. 108–111.

Hu, Y., and W. Perraudin, 2002, "The Dependence of Recovery Rates and Defaults," working paper, Birkbeck College, University of London.

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.

Renault, O., and O. Scaillet, 2004, "On the way to recovery: A nonparametric bias free estimation of recovery rate densities," *Journal of Banking & Finance*, 28(12), 2915–2931.

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?," *Credit risk models and management, D. Shimko (ed.), RiskBooks.*

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, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe.

Notes

[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 conditional distributions. We provide a simple example of the technique in a separate Technical Appendix, and in Section 3 of the same document we detail its role in the current application. Refer to Casella and George (1992) for a comprehensive exposition. The Technical Appendix is available for download at the authors' web pages.

[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.

[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.

[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 1 and 2 of the Technical 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).

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

[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.

[13]We detail the procedure we used to identify the particular data transformation and the number of mixture components in Section 4 of the Technical Appendix.

[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`

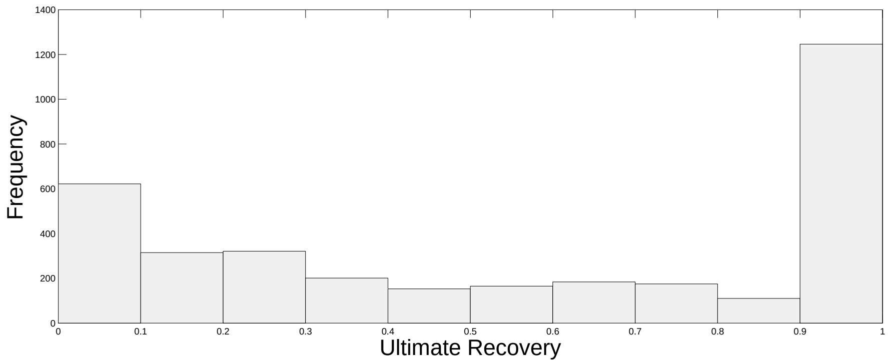
[15]Refer to Altman, Fargher, and Kalotay (2010) for details of how the probabilities are estimated and how the industry-level indexes are applied to default risk estimation.

[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.

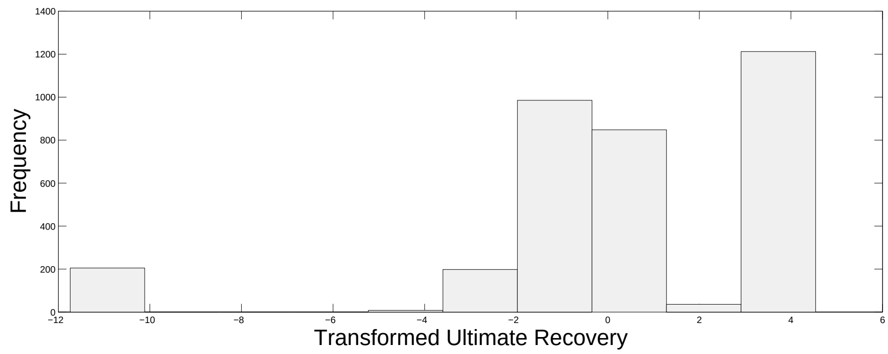
[17]We did not seek to consider a wider set of conditioning variables for two reasons. First, to avoid the trap of data-mining. Second, given the distribution of observations in our data set, conditioning our estimates of mixture weights on many variables for purposes of the simulation study would be problematic.

I 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 3 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 4 for further information.

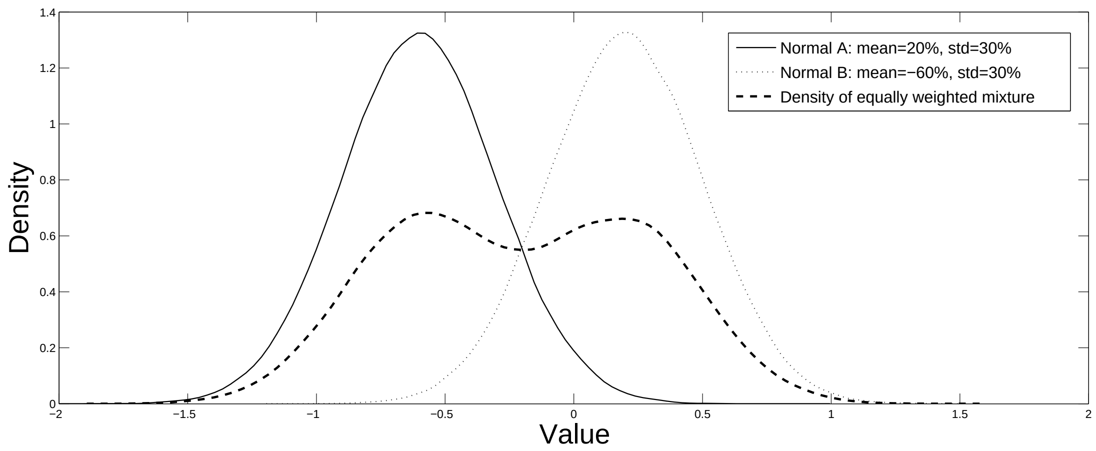
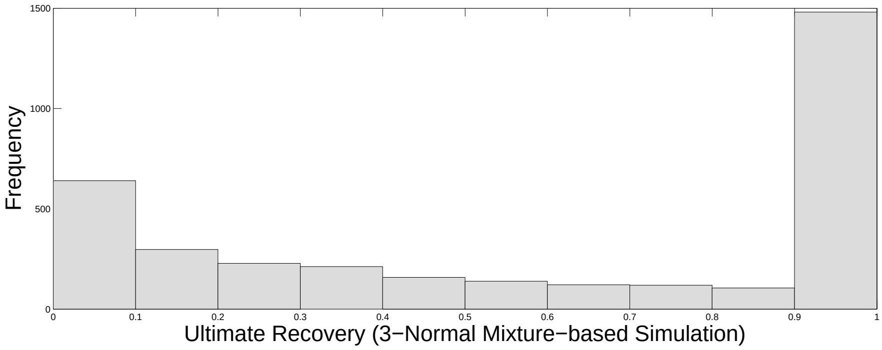
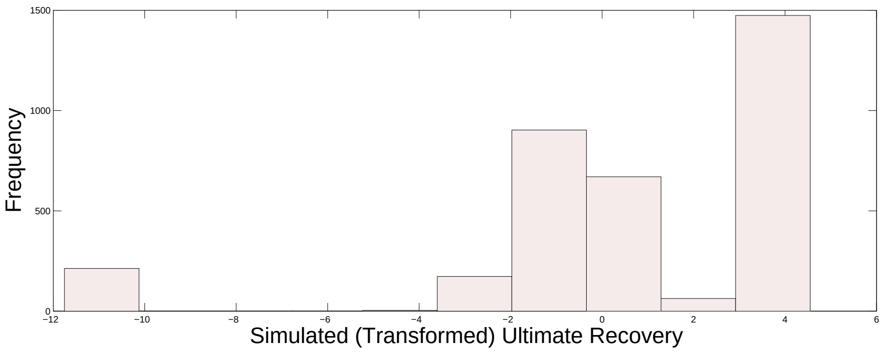


Figure 2: **A 2-Normal Mixture.** The (dashed) mixture density obtained by drawing with equal probability from two normals with differing means.

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

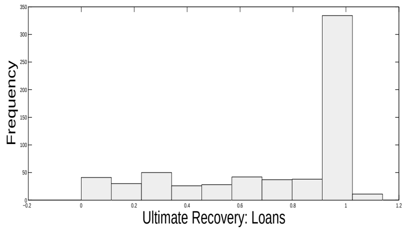


(a) Discounted ultimate recoveries based on simulated sample of 3500

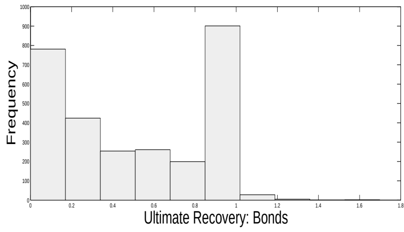


(b) Simulated Inverse T transformed Discounted Ultimate Recoveries, sample of 3500.

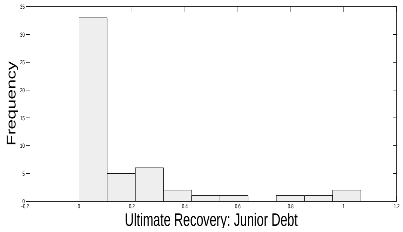
Figure 4: Sample Distribution of Moody's Discounted Ultimate Recoveries: by Type



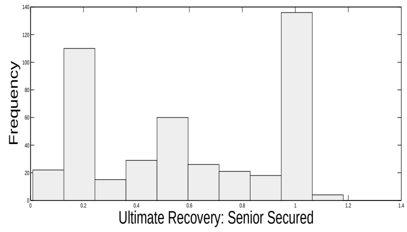
(a) Histogram of Moody's Discounted Ultimate Recovery: Loans Only.



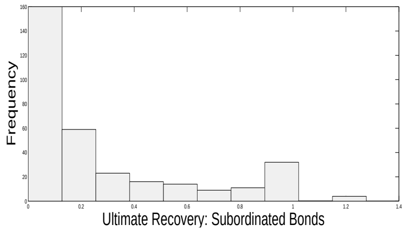
(b) Histogram of Moody's Discounted Ultimate Recovery: Bonds Only.



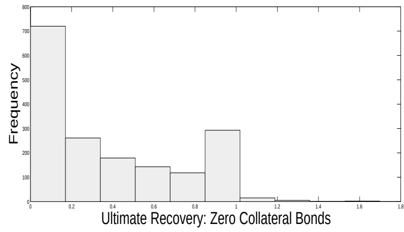
(c) Histogram of Moody's Discounted Ultimate Recovery: Junior Debt Only.



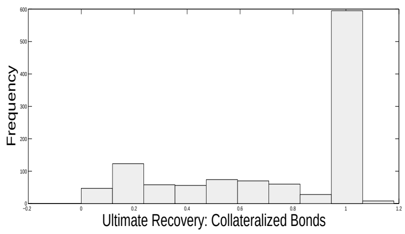
(d) Histogram of Moody's Discounted Ultimate Recovery: Senior Secured Bonds Only.



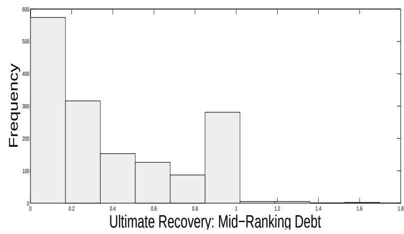
(e) Histogram of Moody's Discounted Ultimate Recovery: Subordinated Bonds Only.



(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.

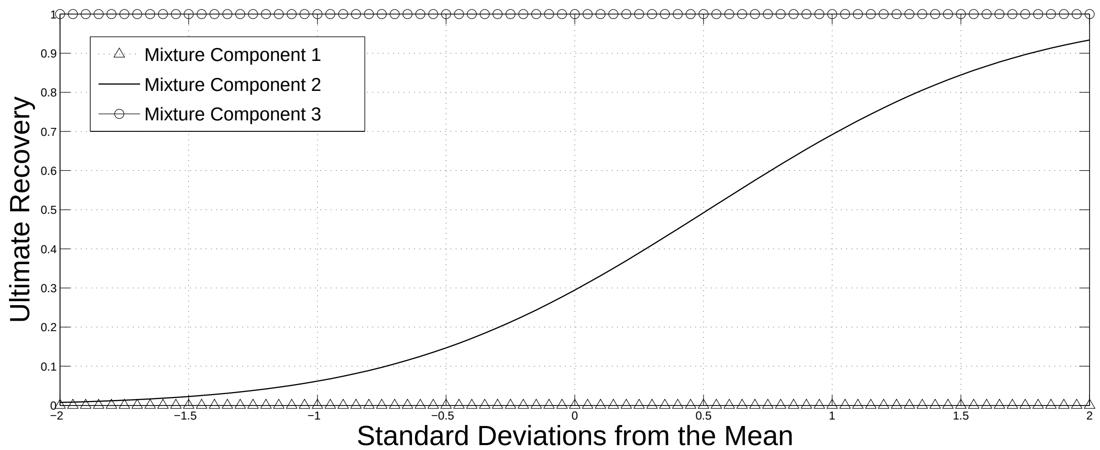


Figure 5: **Mapping Mixture Components to Ultimate Recoveries** This figure maps percentiles of draws from each mixture component to the recovery implied by the CDF of a Student-T with $v = 20$. Triangles denote the first mixture component, a solid line the second, and circles the third. The results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

Table 1: Discounted Ultimate Recoveries: Descriptive Statistics

	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

This table summarizes the characteristics of Moody's discounted ultimate recoveries for loans, bonds, and the categories of exposures presented in Figure 4. "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.

Table 2: Simulated Unconditional Distribution Characteristics based on $m = 3$

	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	**0.59**
Median	1.00	0.62	0.28	0.54	0.43	1.00	0.44	0.49	**0.66**
Std	0.38	0.40	0.43	0.38	0.43	0.37	0.39	0.40	**0.40**
IQR	0.65	0.84	1.00	0.82	0.99	0.64	0.90	0.89	**0.82**
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**

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 4. “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.

Table 3: Mixture Components: Summary Statistics

	Mixture Component		
$i =$	1	2	3
Component Distributions: Normal Parameters			
$E(\alpha_i \mid y)$	-11.74	-0.55	4.54
$E(h_i \mid y)$	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			
$E(p_i \mid y)$	6%	59%	35%
$\sigma(p_i \mid y)$	0.4%	0.80%	1%

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 the Technical Appendix.

Table 4: Simulated Distribution Characteristics based on $m = 3$, Conditional on Industry Distress State

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**

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 4. "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.

Table 5: Simulated Distribution Characteristics based on Linear Regression with $m = 2$

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

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.

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

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

This table is the analogue of Table 2 based on the 2-Mixture conditional specification summarized in Table 5.

Table 7: Mixture Probabilities based on Debt Cushion and Industry Default Conditions

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

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

Table 8: Portfolio Recoveries

	Actual	*3-Mix Comp*	*3-Mix Char*	*3-Mix Base*	*Regression*	*2-Mix Reg*	*Beta Comp*
Mean	83.8	87.3	89.9	90.6	80.5	96.6	71.7
Median	83.8	87.4	89.9	90.6	80.6	96.6	71.7
Std	4.7	4.9	4.8	4.9	5.3	4.8	5.0
IQR	6.4	6.5	6.5	6.7	7.1	6.4	6.7
Lower Tail Percentile Cut-Offs							
5%	76.0	79.2	82.0	82.6	71.6	88.6	63.6
2%	74.0	77.2	79.8	80.5	69.3	86.5	61.5
1%	72.9	76.0	78.4	79.1	67.8	85.0	60.2
0.50%	71.5	74.7	77.0	77.8	66.6	83.8	58.9
0.10%	69.4	71.7	74.5	75.1	64.5	80.8	56.7
% Absolute Error in Estimation: Lower Tail Percentiles							
5%	-	4.2%	7.8%	8.7%	5.8%	16.6%	16.3%
2%	-	4.3%	7.8%	8.7%	6.3%	16.8%	16.9%
1%	-	4.4%	7.7%	8.5%	7.0%	16.6%	17.3%
0.50%	-	4.4%	7.7%	8.8%	6.9%	17.2%	17.7%
0.10%	-	3.4%	7.3%	8.2%	7.0%	16.5%	18.3%

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 7. The posterior percentiles are computed based 10,000 draws and the hypothetical portfolio size is 150 $1.00 exposures.