On Estimating Recovery Rates

Akshit Aggarwal Sida Chen Daniel Debbini Po Hu
Mohammad Ziaemohseni

MS&E 444 Project Report

Abstract

In this report we introduce four different methods for predicting recovery rates based on firm's capital structure and a subset of econometric factors at the time of the default. These methods include linear regression (OLS), nonlinear regressions (CART and MART), and a new method developed on this report based on solving a non-convex optimization problem, which we call it the waterfall model.

1 Introduction

In the event of a corporate default, the recovery rates on different levels of debt depend on the distribution of debt within the company's capital structure. The credit spread that investors should receive also depends on the estimated recovery rate on the debt they hold. In a corporate bankruptcy situation and under the strict legal priority of claims, holders of senior secured debt (bank loans, secured bonds) must be paid in full before senior unsecured debt holders, who must in turn be paid in full before subordinated debt holders. The recovered cash flows down the capital structure until the entire recovered amount has been exhausted. As a result, an upside-down capital structure with a high percentage of senior debt will generate a much lower recovery rate for senior debt holders than a capital structure with a low percentage of senior debt.

A quantitative approach for predicting recovery rates involves modeling an explicit distribution of overall recovery on the bankrupt company's assets and then deriving the expected recoveries using the priority of claims, i.e., the senior secured debt holders are paid back until complete recovery, then cash starts flowing to senior unsecured debt holders and so forth until the pool of recovered assets is exhausted. In practice, settlements in bankruptcy strongly reflect this principle of priority, though they do not usually follow it exactly. For example, senior debt holders may agree to give a higher recovery to junior bond holders than they would otherwise receive in order to speed up the resolution process and to reduce the outflow of recovered amounts towards legal fees.

In this report we propose different methods for estimating recovery rates for a given default and we compare the relative performance of these methods. The rest of this report

is organized as follows. In section (2) we introduce our mathematical notation and develop a simple compact formula for evaluating recovery rates given a firm's capital structure. In section (3) we use maximum likelihood estimation to fit a beta distribution to our empirical data and we evaluate the validity of the fit. In section (4) we introduce three different regression techniques for estimating recovery rates from economic factors. In section (5) we combine both problems of distribution-fitting and regression into one non-convex problem and we show that how by solving this problem one is able to predict recovery rates, and finally in section (6) we conclude this report with some final remarks.

2 Mathematical Setup

In this section we introduce our mathematical notation. We denote the capital structure of a firm at the time of default by vector $\mathbf{s}$:

$$\mathbf{s} = (s_1, s_2, ..., s_n)^{\mathrm{T}} \in \mathbf{R}_+^n \tag{1}$$

We assume that elements of $\mathbf{s}$ are ordered based on their seniorities, so s_1 is the most senior claim and s_2 is the second most senior claim and so on. We define the actual tranche recovery rates after a 30-day period by vector $\tilde{\mathbf{r}}$:

$$\tilde{\mathbf{r}} = (\tilde{r}_1, \tilde{r}_2, ..., \tilde{r}_n)^{\mathrm{T}} \in [0\ \ 1]^n \tag{2}$$

Knowing $\mathbf{s}$ and $\tilde{\mathbf{r}}$ we can calculate the firm recovery (or simply recovery) as follows:

$$r = \frac{\mathbf{s}^{\mathrm{T}}\tilde{\mathbf{r}}}{\sum_{i=1}^n s_i} \tag{3}$$

Note that unlike $\tilde{\mathbf{r}}$, r is a scalar. We also define our predicted values of tranche recoveries at the time of default to be $\hat{\mathbf{r}}$:

$$\hat{\mathbf{r}} = (\hat{r}_1, \hat{r}_2, ..., \hat{r}_n)^{\mathrm{T}} \in [0\ \ 1]^n \tag{4}$$

Now assume that recoveries follow a distribution with pdf $f(r)$ and define vector $\mathbf{u}$ as follows:

$$\mathbf{u} = (u_1, u_2, ..., u_n, u_{n+1})^{\mathrm{T}} = \frac{1}{\sum_{i=1}^n s_i}(0, \sum_{i=1}^{1} s_i, \sum_{i=1}^{2} s_i, ..., \sum_{i=1}^{n-1} s_i, \sum_{i=1}^{n} s_i)^{\mathrm{T}} \tag{5}$$

Note that $\mathbf{u}$ is a $(n+1)$-vector. With this setup we can evaluate estimated recoveries for each tranche using the following equation:

$$\hat{r}_i = \int_{u_i}^{u_{i+1}} \frac{r - u_i}{u_{i+1} - u_i} f(r) dr \tag{6}$$

The mean squared error (MSE) between actual recoveries and predicted recoveries are:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\tilde{r}_i - \hat{r}_i)^2 \tag{7}$$

We also define the average absolute error (AAE) to be:

$$AAE = \frac{1}{n}\sum_{i=1}^{n}|\tilde{r}_i - \hat{r}_i| \tag{8}$$

Note that we have assumed an arbitrary pdf $f(r)$ in equation (1). In next section we deal with a suitable choice for the distribution $f(r)$.

3 Maximum Likelihood Estimation of Recoveries

Many rating agencies such as Moody's model recovery rates using a beta distribution. Beta distributions are a family of continuous time distributions defined on the interval [0 1]. Since values of recoveries also fall in this same interval, the domain of the beta distribution can be viewed as a recovery value. The pdf of a beta distribution is:

$$f(r) = f(r, \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta-1} \tag{9}$$

where $0 \leq r \leq 1$ represents recovery, Γ is the gamma function $\Gamma(n) = (n-1)!$, and $\alpha > 0$ and $\beta > 0$ are called shape parameters. We define the likelihood function for a set of recoveries $r_1, r_2, ..., r_k$ as:

$$L(\alpha, \beta) = \prod_{i=1}^{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r_i^{\alpha-1}(1-r_i)^{\beta-1} \tag{10}$$

The maximum likelihood estimator finds values of α and β such that the likelihood function $L(\alpha, \beta)$ is maximized. Often for computational purposes it is easier to maximize the logarithm of $L(\alpha, \beta)$. In this case we have:

$$l(\alpha, \beta) = k\log\Gamma(\alpha+\beta) - k\log\Gamma(\alpha) - k\log\Gamma(\beta) + (\alpha-1)\sum_{i=1}^{k}\log r_i + (\beta-1)\sum_{i=1}^{k}\log(1-r_i) \tag{11}$$

where $l(\alpha, \beta) = \log L(\alpha, \beta)$ is called the log-likelihood function. Using information for 1063 recoveries from our dataset, we used the maximum likelihood estimator [Wessa:2011] to find estimates of values of α and β. Our estimated values are $\hat{\alpha} = 1.11$ and $\hat{\beta} = 1.66$. Figure (1) shows the Q-Q plot for comparing this distribution with the distribution of our recovery data. Notice that the points in Q-Q plot approximately lie on the line $y = x$, which means that our estimated values $\hat{\alpha}$ and $\hat{\beta}$ are consistent with our data. Figure (2) shows the pdf of our estimated beta distribution for recoveries.

4 Recovery Estimation Using Regression

In this section we perform multiple regression techniques on recovery rates for each bond in Moody's database. Our data consist of 196 default events with more than one tranche of debt from February 28, 1975 to April 8, 2009. We sort the data by chronological order and choose the first 154 defaults as our training set and the most recent 42 defaults as our test set. We use the same training and test sets in our subsequent analysis.

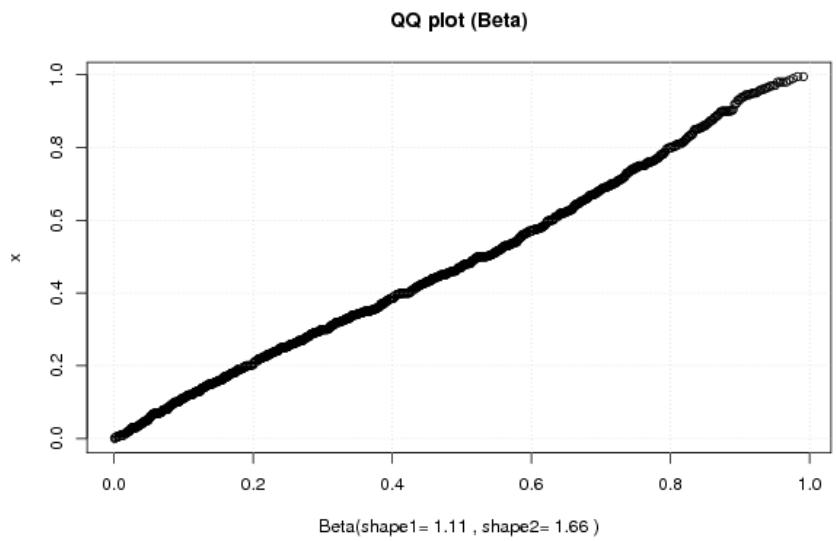

Figure 1: Q-Q plot for maximum likelihood estimator of recovery distribution.

4.1 Ordinary Least Squares (OLS)

Suppose that our training set consists of m different tranche recoveries, and suppose that each tranche recovery is a linear function of a set of regressors, so we have:

$$\tilde{\mathbf{r}}_i = \mathbf{x}_i^{\mathrm{T}}\mathbf{C} \qquad 1 \leq i \leq m \tag{12}$$

where $\tilde{\mathbf{r}}_i$ is the i-th tranche recovery, $\mathbf{x}_i$ is a vector of econometric factors associated with $\tilde{\mathbf{r}}_i$, and $\mathbf{C}$ is a vector of constant coefficients we are interested to determine. For m different values of tranche recoveries, we can express equation (12) in matrix notation as:

$$\tilde{\mathbf{R}} = \mathbf{XC} \tag{13}$$

Here $\tilde{\mathbf{R}}$ is an $m \times 1$ vector where each element corresponds to one of m tranche recoveries. Suppose that we designate q econometric factors to each $\tilde{\mathbf{r}}_i$. In this case $\mathbf{X}$ would be an $m \times (q+1)$ matrix and $\mathbf{C}$ would be an $(q+1) \times 1$ vector. Note that because of the existence of an intercept term, the first column of matrix $\mathbf{X}$ would be all 1's. In this case the first element of vector $\mathbf{C}$ is called the intercept term. Ordinary least squares (OLS) finds the vector $\mathbf{C}$ which minimizes $||\mathbf{XC} - \tilde{\mathbf{R}}||$ over our training set. After finding this value of $\mathbf{C}$ we can plug it into the following equation to obtain our estimated values of tranche recoveries for the test set:

$$\hat{\mathbf{R}} = \mathbf{XC} \tag{14}$$

We used OLS on our dataset to obtain a linear fit for our recovery data. The final linear model uses the following econometric factors: the S&P 500 trailing 6 month log return, seniority, rating, default type, industry, inflation adjusted M2 money supply, number of

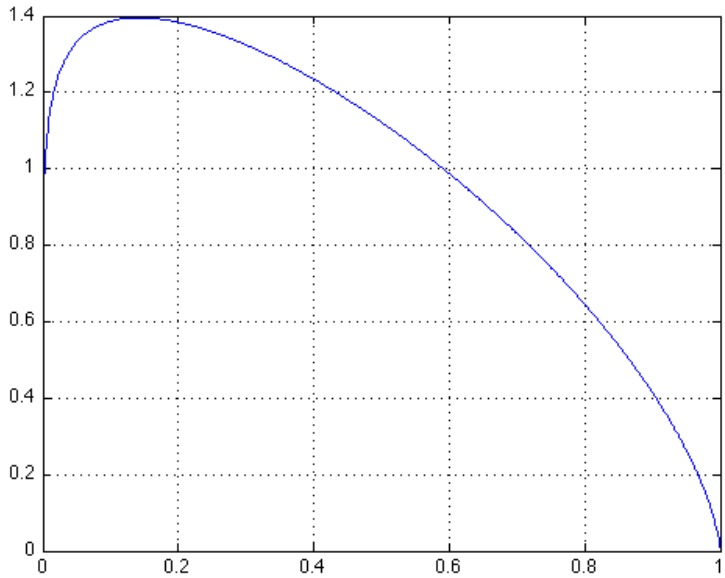

Figure 2: The pdf of estimated beta distribution for recoveries.

defaults in the past 6 months, and weekly unemployment claims. The model has an R-squared of 0.60 and an adjusted R-squared of 0.55. The averaged AAE on the test set is 0.3502 which is relatively high. Thus, even with a reasonable value for R-squared, this linear model has very limited predictive capability and we need to refine our regression model by exploring other techniques with greater predictive power.

4.2 Classification and Regression Trees (CART)

In this section we use a classification and regression tree (CART) to predict tranche recoveries. CART is a decision tree learning algorithm which is commonly used in data mining. Its goal is to create a regression model that estimates a target value based on different input variables. CART has several advantages over OLS, for example, it is resistant to irrelevant variables, i.e., we don't have to do variable selection and we don't need to worry about problems with overfitting. It handles mixed variable types and it is easily interpreted using a 2D tree. A simple explanation of such a tree is shown in figure (3).

We applied the CART algorithm to our dataset to obtain a fit for our recovery data. The final model uses the following econometric factors: the S&P 500 trailing 3, 6, and 12 month log returns, seniority, rating, default type, industry, inflation adjusted M2 money supply, percent of debt above in capital structure, and yield spread of 10 year T-Bill over the federal funds rate (Compare these with the indicators we used in the previously using LOS). The averaged AAE on the test set is 0.2195 which is much lower than what we got using LOS. However, one major disadvantage of CART is the instability of the learning tree and its

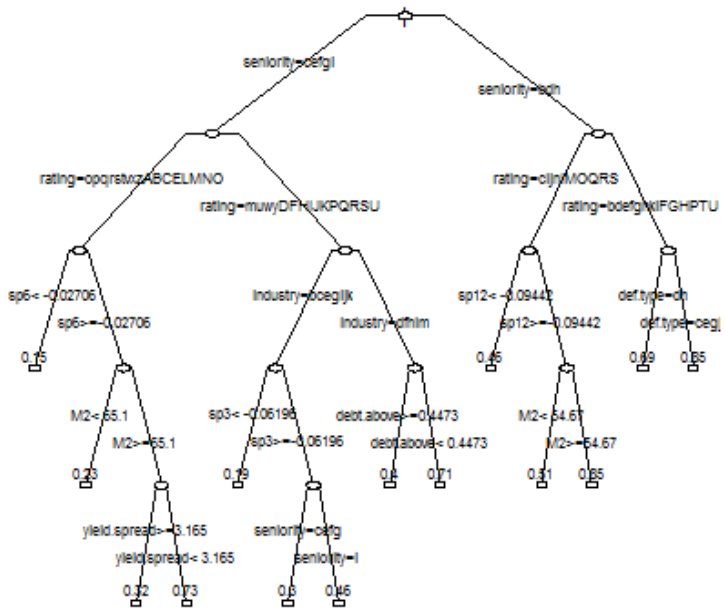

Figure 3: A CART model can be visualized as a simple tree. Each split in the tree corresponds to a yes/no question about the observation in question. For example, here the first split corresponds to whether the bond is a senior secure bond, revenue bond, or has multiple seniority.

ability to only measure higher order interactions in the data. A small change in the data will probably change the tree a lot. Thus, we pursue one further refinement of our regression.

4.3 Multiple Additive Regression Trees (MART)

In this section we use a multiple additive regression tree (MART) to predict tranche recoveries. MART is an iterative implementation of the gradient tree boosting algorithm for predictive data mining and is widely used in practice. It has several advantages over other comparable algorithms, for example, it is capable of handling mixed variable types. Many data mining and machine learning algorithms do not have a way of dealing with categorical variables such as industry type, default type, or ratings. MART is invariant to monotone transformations of input variables. In OLS, the model is sometimes extremely dependent on a clever transformation of input variables. In fact, Moody's LossCalc notes that most

economic factors are transformed. When using MART, we do not have to search for these transformations. Another advantage of MART is that it is immune to outliers, so one faulty observation will not skew the predictions. Finally, MART is robust to irrelevant regression variables. This is extremely convenient when the universe of possible regressors is large, which may be the case in a recovery estimation problem. There are many more economic and firm-specific factors that we can add into our model without corrupting our predictive power.

We applied the MART algorithm to our dataset to obtain a fit for our recovery data. We used the following econometric factors: the S&P 500 trailing 3, 6, and 12 month log returns, seniority, rating, default type, industry, inflation adjusted M2 money supply, percent of debt above in capital structure, weekly unemployment claims, monthly unemployment rate, number of defaults in the past six months, and yield spread of 10 year T-Bill over the federal funds rate. The averaged AAE on the test set is 0.1739 which is our best result so far. Figure (4) shows a plot of the averaged AAE for our training and test sets as a function of the number of iteration in MART algorithm.

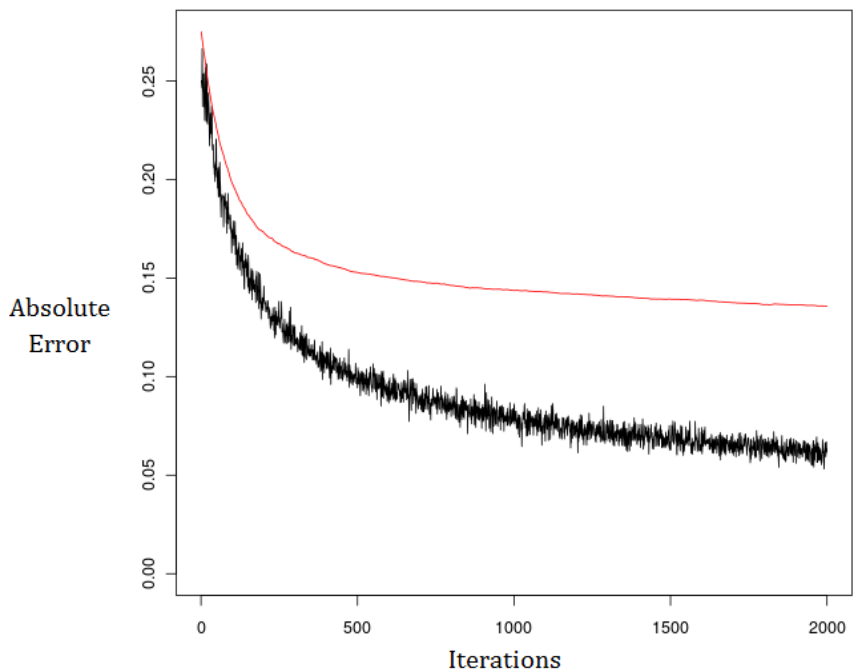

Figure 4: Plot of the averaged AAE for our training data (black) and test data (red) as a function of the number of iteration in MART algorithm.

One of the nice features of the MART algorithm is that it generates an importance measure for each input variable, based on the reduction in error when adding the variable into the model. This feature of MART allows for some interesting and insightful interpretations of our data. We provide three figures to emphasize this point further. Figure (5) shows

the four most important regressors in recovery prediction using MART model. According to this figure, the four most important regressors are seniority, rating, industry, and percent of debt above in capital structure. Figure (6) shows a plot of average partial dependence

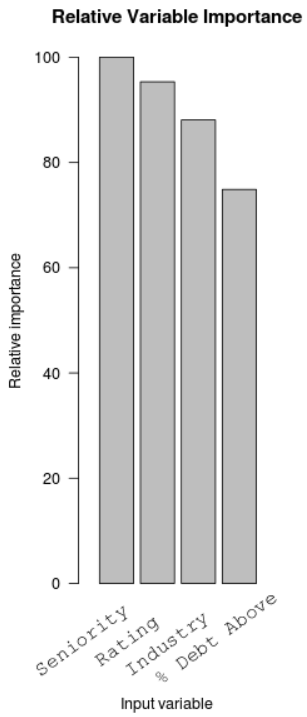

Figure 5: The four most important regressors in the MART model are seniority, rating, industry, and percent of debt above in capital structure.

of recovery rates on percent debt above in capital structure. We observe a strong negative trend here, meaning that recovery rates are lower for relatively junior tranches that have lots of relatively senior debt above them. And finally, figure (7) shows a plot of average partial dependence of recovery rates on industry. It shows that technology sector has a much lower recovery rate relative to other industries. We also observe that financial (non-bank) and utilities sectors have lower than average recovery rates while banking and transportation sectors have relatively higher recovery rates.

5 Combined Distribution Fit and Regression (Waterfall Model)

In this section we introduce a method which combines the problems of distribution fitting and regression into one optimization problem and we use the solution of this optimization problem for recovery estimation. We call this model a waterfall model. Suppose that our training

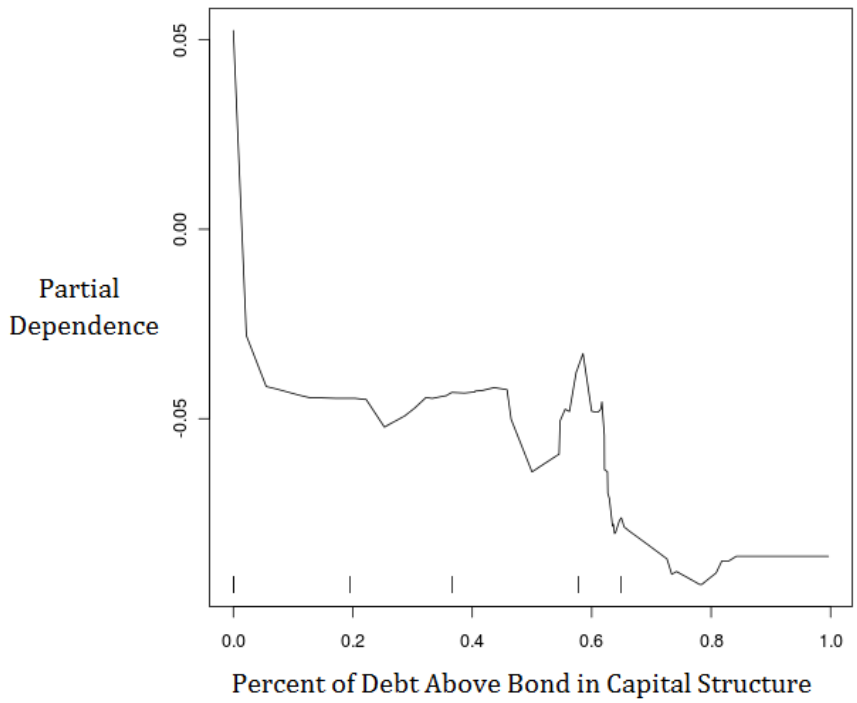

Figure 6: Partial dependence of recovery rates on the percent debt above in capital structure. There is a strong negative trend, recovery rates are lower for relatively junior tranches that have several relatively senior debt above them.

set consists of k defaults, and moreover, the recoveries of each default follow a different beta distribution. Suppose that we can write the parameters of these beta distributions as linear functions of corresponding econometric factors, so we have:

$$\alpha_j = \mathbf{x}_j^{\mathrm{T}} \mathbf{C}_1 \qquad 1 \leq j \leq k \tag{15}$$

$$\beta_j = \mathbf{x}_j^{\mathrm{T}} \mathbf{C}_2 \qquad 1 \leq j \leq k \tag{16}$$

Using equation (6), we can write the tranche recoveries for j-th default as:

$$(\hat{r}_i)_j = \int_{u_i}^{u_{i+1}} \frac{r - u_i}{u_{i+1} - u_i} f(r) dr \tag{17}$$

Here j is the default number ($1 \leq j \leq k$) and i is the tranche number within default j, and $f(r)$ is the following beta distribution:

$$f(r) = f(r, \alpha_j, \beta_j) = \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)} r^{\alpha_j - 1}(1 - r)^{\beta_j - 1} \tag{18}$$

We define our objective function as follows:

$$\sum_{j=1}^{k} \sum_{i} ((\hat{r}_i)_j - (\tilde{r}_i)_j)^2 \tag{19}$$

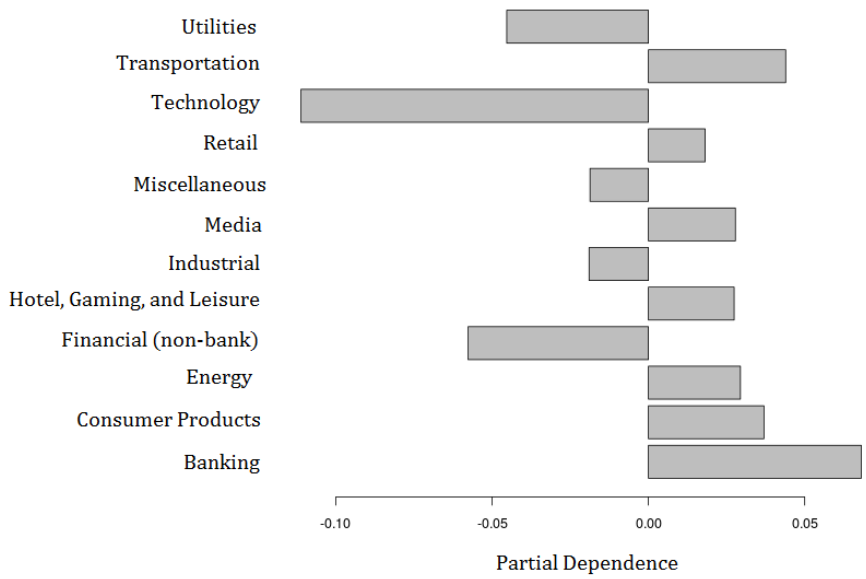

Figure 7: Partial dependence of recovery rates on industry. Note that technology sector has a much lower recovery rate relative to other industries while financial (non-bank) and utilities sectors have lower than average recovery rates and banking and transportation sectors have relatively higher recovery rates.

Where $\tilde{r}_i$'s denote the actual values of tranche recoveries in our data. The waterfall model minimizes objective function (19) subject to linear constraints (15) and (16), over all tranches and all defaults in our training data. This gives us the vectors $\mathbf{C}_1$ and $\mathbf{C}_2$. We can use these values to predict recoveries for our test data. To do so, first we estimate the parameters of beta distribution using equations (15) and (16) and then we use equation (17) to obtain our estimates for each particular tranche recovery.

Note that the objective (19) is not a convex function. To solve this optimization problem we should use several different and possibly randomized initial guesses for $\mathbf{C}_1$ and $\mathbf{C}_2$ and improve our guesses with each iteration. A major problem here is that with randomized $\mathbf{C}_1$ and $\mathbf{C}_2$, equations (15) and (16) may evaluate to negative values for either α or β or both. This will in turn halt our optimization algorithm, since beta distributions are only defined for positive values of α and β. To avoid this problem, we use the following transformations:

$$\alpha_j = 5(\frac{1}{\pi}\tan^{-1}(\mathbf{x}_j^{\mathrm{T}}\mathbf{C}_1) + \frac{1}{2}) \tag{20}$$

$$\beta_j = 5(\frac{1}{\pi}\tan^{-1}(\mathbf{x}_j^{\mathrm{T}}\mathbf{C}_2) + \frac{1}{2}) \tag{21}$$

And then minimize our objective (19) subject to constraints (20) and (21). Note that equations (20) and (21) indicate that the values of α and β are between zero and 5. This may seem contradictory at first, since in theory these values could be anything between zero

and infinity. We should note that increasing the value of scalar 5 in equations (20) and (21) does not affect our results at all. In fact, the optimal values of α and β for recovery estimations tend to be small (see, for example, the optimal values of α and β for maximum likelihood estimation in section 3). A plot of the transformation $S(z) = 5(\frac{1}{\pi}\tan^{-1}(z) + \frac{1}{2})$ is shown in figure (8).

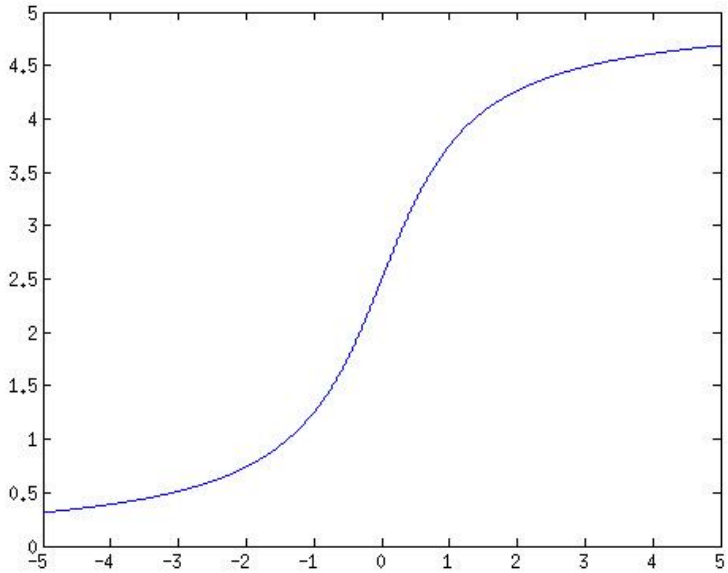

Figure 8: Plot of the transformation used in waterfall model.

We used waterfall model on our dataset to obtain a fit for our recovery data. We used several different initial values for $\mathbf{C}_1$ and $\mathbf{C}_2$ and our final results were consistent. We used the following econometric factors in our model: yield spread of 10 year T-Bill over the federal funds rate, weekly unemployment claims, monthly unemployment rate, inflation adjusted M2 money supply, and the S&p 500 trailing 6 month log returns. The averaged AAE on the training set is 0.1979 and the averaged AAE on the test set is 0.1826. For the test set, the averaged AAE for senior and junior tranches are 0.1842 and 0.1810 respectively, and the averaged AAE of the recovery spread on the test set is 0.1985.

6 Conclusion and Final Remarks

In this report we presented four different methods (OLS, CART, MART, and waterfall model) for estimating recovery rates in defaults using their capital structure and some econometric factors at the time of the default. Table (1) shows a summary of the performance of these methods. From this table we see that MART and waterfall model produce the smallest prediction errors on our dataset.

Estimation Method	Averaged AAE
OLS	0.3502
CART	0.2195
MART	0.1739
Waterfall Model	0.1826

Table 1: The averaged AAE of test data for different recovery estimation methods.

One particular difficulty in predicting recovery rates is that in practice the absolute priority rule of recoveries is often disobeyed. In many cases the bonds with same seniority have different recoveries, or even worse, in many cases junior tranches recover more than senior tranches. Unfortunately, the beta distribution is not able to model this idiosyncrasies and it is not exactly clear to us at this moment how one should proceed with modeling this anomalies in recovery rates mathematically.

Acknowledgements

We would like to thank Professor Kay Giesecke and our friendly TA Gerald Teng for their guidance on different aspects of this project during our numerous meetings and conversations.

References

1. Schläfer, Timo S. and Marliese Uhrig-Homburg, 2009, Estimating Market-implied Recovery Rates from Credit Default Swap Premia, *13th Conference of the Swiss Society for Financial Market Research*

2. Jacobs, Michael Jr. and Ahmet K. Karagozoglu, 2010, Modeling Ultimate Loss-Given-Default on Bonds and Loans

3. Gupton, Greg M. and Roger M. Stein, 2002, LOSSCALC™: Model For Predicting Loss Given Default (LGD)

4. Recovering Your Money: Insights Into Losses From Defaults. 1999. *Standard & Poor's Creditweek June 16:* 29-34

5. Altman Edward I., Andrea Resti, Andrea Sironi and Brooks Brady, 2005, The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications, *Journal of Business*, vol.78,no.6

6. Til Schuermann, 2004, What Do We Know About Loss Given Default? *Credit Risk Models and Management 2nd Edition, London, UK: Risk Books*

7. Levent Guntay, Dilip Madan and Haluk Unal, 2001, Pricing the risk of recovery in default with absolute priority rule violation, *Journal of Banking & Finance 27 (2003)*: 1001–1025

8. Wessa, P. (2011), Free Statistics Software, Office for Research Development and Education, version 1.1.23-r6, URL http://www.wessa.net/