Journal of Banking & Finance 40 (2014) 116–129

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Ultimate recovery mixtures

Edward I. Altman [a,1], Egon A. Kalotay [b,*]

[a] Stern School of Business, New York University, Salomon Center, 44 West 4th Street, Suite 9-61, New York, NY 10012, USA

[b] Department of Applied Finance and Actuarial Studies, Macquarie University, North Ryde, NSW 2109, Australia

ARTICLE INFO

Article history:
Received 10 July 2013
Accepted 13 November 2013
Available online 27 November 2013

JEL Classification:
G17
G21
G28

Keywords:
Bankruptcy
Ultimate recovery
Loss given default
Credit risk
Mixtures of distributions
Defaulted debt

ABSTRACT

We propose a relatively simple, accurate and flexible approach to forecasting the distribution of defaulted debt recovery outcomes. Our approach is based on mixtures of Gaussian distributions, explicitly conditioned on borrower characteristics, debt instrument characteristics and credit conditions at the time of default. Using Moody's Ultimate Recovery Database, we show that our mixture specification yields more accurate forecasts of ultimate recoveries on portfolios of defaulted loans and bonds on an out-of-sample basis than popular regression-based estimates. Further, the economically interpretable outputs of our model provide a richer characterization of how conditioning variables affect recovery outcomes than competing approaches. The latter benefit is of particular importance in understanding shifts in the relative likelihood of extreme recovery outcomes that tend to be realized more frequently than observations near the distributional mean.

1. Introduction

The economic value of debt in the event of default is a key determinant of the default risk premium required by a lender and the regulatory capital charged to limit exposure to losses. The pricing of default risk insurance (CDS contracts) and the emergence of distressed debt as an investment class add further incentive to better understand the distribution of payoffs in the event of default.[2] Adding to market-driven incentives, Basel II and III provide regulatory incentives to the development of recovery models in financial institutions adopting an advanced internal ratings based (IRB) approach to computing capital requirements.

Recognizing the importance of capturing the behavior of recoveries in the event of default to quantitative models of credit risk, recent years have seen a wave of research from academics and industry professionals seeking to document the key empirical features of observed recovery outcomes. While payoffs to debt holders in the event of default depend on the interplay of many factors, often idiosyncratic, notable empirical regularities from prior research are evident.

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 (proportional) value of claims subordinate to the debt at a given seniority, known as the Debt Cushion, also seems to matter. The analysis of Keisman and Van de Castle (1999) suggests that all else equal, the larger the Debt Cushion, the higher the expected recovery outcome.
3. Recoveries tend to be lower in recessions and other periods when the rate of aggregate defaults is high. Altman et al. (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

* Corresponding author. Tel.: +61 (2) 9850 8490; fax: +61 (2) 9850 8497.
E-mail addresses: ealtman@stern.nyu.edu (E.I. Altman), egon.kalotay@mq.edu.au (E.A. Kalotay).
[1] Tel.: +1 (212) 998 0709; fax: +1 (212) 995 4220

[2] Altman and Kuehne (2013) estimate the face and market values of distressed and defaulted debt in the U.S. over time. At the end of 2012 their estimates of the size of the market exceeded $1.15 trillion face value and $678 billion market value with over 200 institutions investing in such securities.

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

0378-4266/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2013.11.021

annual default rate. Further, Frye (2000) shows that a 10% realized default rate results in a 25% reduction in recoveries relative to its normal year average.

4. Industry matters. Acharya et al. (2007) suggest that macroeconomic conditions do not appear to be significant determinants of individual bond recoveries after accounting for industry effects. More recently, Jankowitsch et al. (2012) find that the type of default, seniority of the bond and industry are as important determinants of recovery as balance sheet ratios motivated by structural credit risk models, macrovariables and transaction cost variables.
5. Variability of recoveries is high, even intra-creditor-class variability, after categorization into sub-groups. For example, Schuermann (2004) notes that senior secured bond investments have a flat distribution – indicating that recoveries are relatively evenly distributed from 30% to 80%.

Clearly, the empirical features of historical recoveries suggest the need for caution in applying popular (parametric) tools of inference – such as OLS regressions and calibrated Beta distributions. While OLS regression models provide simple, intuitive summaries of data relationships, they make strong assumptions about the conditional distribution of recovery outcomes and focus attention on variation in the mean. Alternatively, Beta distributions calibrated to historical data are used in many commercial models of portfolio risk to characterize the distribution of loss outcomes.[4] While Beta distributions offer a simple, parsimonious way of capturing a very broad range of distributional shapes over the unit interval, Servigny and Renault (2004) observe that they cannot accommodate bi-modality, or probability masses near zero and unity – important features of empirical recovery distributions.

While stylized models and a growing body of empirical evidence reveal much about the important influences on debt recovery outcomes, they also serve to highlight the challenges inherent in building a quantitative model to account for: characteristics specific to the defaulted instrument, borrower characteristics, macroeconomic conditions at the time of default, and the idiosyncrasies of recovery distributions' shape. Building on insights from empirical research and the findings of recent studies documenting the relative merits of non-parametric and regression based approaches, we present in this paper a novel approach to modeling recoveries on defaulted debt using mixtures of Gaussian distributions.

More specifically, our paper makes three contributions to the literature. First, we present an approach to modeling recovery distributions that retains the flexibility of non-parametric methods while providing transparency with respect to the economic sources of variation in recovery outcomes. Second, we estimate and evaluate the out-of-sample performance of our model using Moody's Ultimate Recovery Database spanning a 25 year sample period ending in 2011. As noted by Bastos (2010) and Qi and Zhao (2011), very few studies to date have evaluated the predictive performance of alternative modeling methodologies. While they present tests of non-parametric approaches relative to regression-based alternatives, neither of the studies consider semi-parametric models. Third, our model provides further clarity on the role and importance of economic influences on recovery outcomes.

The remainder of our study proceeds as follows. We provide in Section 2 an overview of recent approaches to recovery modeling and an overview of the approach proposed in this paper. In Section 3 we describe the ultimate recoveries data used in this study and we detail the econometric approach in Section 4. We report model estimates and comparative performance metrics in Section 5 and summarize our findings in Section 6.

2. Recovery modeling approaches

Recent studies have investigated the forecasting performance of non-parametric estimation approaches relative to a variety of parametric regression specifications. Using loss data on defaulted Portuguese bank loans, Bastos (2010) finds that non-parametric regression trees tend to outperform parametric regression-based forecasts over shorter (annual) horizons. Similarly, using a larger US sample of defaulted loans and bonds, Qi and Zhao (2011) find that forecasts based on regression trees and neural networks outperform those of parametric regression models. Importantly, they attribute the success of non-parametric models to their ability to accommodate non-linear associations between debt recoveries and continuous conditioning variables. Similarly, recent work by Loterman et al. (2012) underscores the importance of models that incorporate non-linearities in predictive relations.

In demonstrating the predictive properties of non-parametric techniques relative to regression models, the studies by Bastos (2010) and Qi and Zhao (2011) also serve to highlight the potential shortcomings of the approaches. Qi and Zhao (2011) acknowledge a basic criticism of neural networks, namely, that they do not provide any insight to the economic relationships underpinning the forecasts. While regression trees are more transparent and intuitive they can become unwieldy in size and incorporate relationships that are difficult to reconcile with a priori expectations.

In modeling Moody's data on ultimate recoveries between 1985 and 2008, Qi and Zhao (2011) build a tree with 342 splits. While the regression trees built using the much smaller dataset employed by Bastos (2010) contain between 1 and 3 splits only, they suggest a primary role for loan size as a driver of expected recovery outcome – a strong finding that appears specific to the data used in the study. More recently, Bastos (2013) suggests that ensembles of regression trees, obtained through varying the estimation sample, outperform trees estimated using a single historical data set.

Given the empirical properties of recoveries and the relative merits of regression-based and non-parametric modeling techniques, we present in this paper a simple semi-parametric approach based on mixtures of distributions. Our approach is flexible enough to capture the distinctive features of recovery distributions while providing insight to the economic relationships from which predictions are derived. Instead of trying to force-fit a parametric distribution, we adopt a Bayesian perspective and model the distribution of recoveries using mixtures of Gaussian distributions.[5] By taking the appropriate probability weighted average of Gaussian components, we are able to accommodate the unusual defining features of such distributions. By explicitly modeling the assignment of recovery outcomes mixture components using an ordered probit regression we accommodate non-linearities in the relation between continuous conditioning variables and recovery outcomes suggested in earlier work.[6]

Similar to Hu and Perraudin (2002), we commence by transforming ultimate recoveries r from the unit interval to the real line such that

$$y = \Phi^{-1}(r), \tag{1}$$

[4] Portfolio Manager (Moody's KMV), Portfolio Risk Tracker (Standard and Poor's) and CreditManager (MSCI Inc.) [formerly CreditMetrics (J.P. Morgan)] are all based on the assumption that losses in the event of default are described by a Beta distribution.

[5] Recent work by Hagmann et al. (2005); Hlawatsch and Ostrowski (2011) and Zhang and Thomas (2012) present alternative semi-parametric approaches to modeling recoveries on defaulted debt. We discuss the benefits of our approach to these alternatives in Section 4.3.

[6] For example, regressions reported in Altman et al. (2005) suggest a non-linear relation between aggregate recoveries and the contemporaneous default rate.