Journal of Banking & Finance 40 (2014) 364–375

Contents lists available at ScienceDirect

Journal of Banking & Finance

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

Loss given default for leasing: Parametric and nonparametric estimations

Thomas Hartmann-Wendels *, Patrick Miller, Eugen Töws

Research Institute for Leasing, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany

ARTICLE INFO

Article history:
Received 7 January 2013
Accepted 6 December 2013
Available online 15 December 2013

JEL classification:
C14
C38
C51
G17
G28

Keywords:
Loss given default
Regression and model trees
Finite mixture models
Leasing
Forecasting

ABSTRACT

This study employs a dataset from three German leasing companies with 14,322 defaulted leasing contracts to analyze different approaches to estimating the loss given default (LGD). Using the historical average LGD and simple OLS-regression as benchmarks, we compare hybrid finite mixture models (FMMs), model trees and regression trees and we calculate the mean absolute error, root mean squared error, and the Theil inequality coefficient. The relative estimation accuracy of the methods depends, among other things, on the number of observations and whether in-sample or out-of-sample estimations are considered. The latter is decisive for proper risk management and is required for regulatory purposes. FMMs aim to reproduce the distribution of realized LGDs and, therefore, perform best with respect to in-sample estimations, but they show poor performance with respect to out-of-sample estimations. Model trees, by contrast, are more robust and outperform all other methods if the sample size is sufficiently large.

1. Introduction

The loss given default (LGD) and its counterpart, the recovery rate, which equals one minus the LGD, are key variables in determining the credit risk of a financial asset. Despite their importance, only a few studies have focused on the theoretical and empirical issues related to the estimation of recovery rates.

Accurate estimates of potential losses are essential to efficiently allocate regulatory and economic capital and to price the credit risk of financial instruments. Proper management of recovery risk is even more important for lessors than for banks because leases have a comparative advantage over bank loans with respect to the lessor's ability to benefit from higher recovery rates in the event of default. In their empirical cross-country analysis, Schmit and Stuyck noted in 2002 that the average recovery rate for defaulted automotive and real estate leasing contracts is slightly higher than the recovery rates for senior secured loans in most countries and much higher than the recovery rates for bonds. Moreover, the recovery time for defaulted lease contracts is shorter than that for bank loans. Because the lessor retains legal title to the leased asset, repossession of a leased asset is easier than foreclosure on the collateral for a secured loan. Moreover, the lessor can retain any recovered value in excess of the exposure at default. Repossessing used assets and maximizing their return through disposal in secondary markets are aspects of normal leasing business and are not restricted to defaulted contracts. Therefore, lessors have a good understanding of the secondary markets and of the assets themselves. Because the lessor's claims are effectively protected by legal ownership, the high recoverability of the leased asset may compensate for the poor creditworthiness of a lessee. Lasfer and Levis (1998) found empirical evidence for the hypothesis that lower-rated and cash-constrained firms have a greater propensity to become lessees. To leverage their potential lower credit risk, lessors must be able to accurately estimate the recovery rates of defaulted contracts.

This paper compares the in-sample and out-of-sample accuracies of parametric and nonparametric methods for estimating the LGD of defaulted leasing contracts. Employing a large dataset of 14,322 defaulted leasing contracts from three major German lessors, we find in-sample accuracy to be a poor predictor of out-of-sample accuracy. Methods such as the hybrid finite mixture models (FMMs), which attempt to reproduce the LGD distribution, perform well for in-sample estimation but yield poor results out-of-sample. Nonparametric models, by contrast, are robust in

* Corresponding author. Tel.: +49 221 470 4480; fax: +49 221 470 2305.
E-mail addresses: hartmann-wendels@wiso.uni-koeln.de (T. Hartmann-Wendels), miller@wiso.uni-koeln.de (P. Miller), toews@wiso.uni-koeln.de (E. Töws).
URL: http://www.leasing.uni-koeln.de.

0378-4266/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2013.12.006

the sense that they deliver fairly accurate estimations in-sample, and they perform best out-of-sample. This result is important because out-of-sample estimation has rarely been performed in other studies – with the notable exceptions of Han and Jang (2013) and Qi and Zhao (2011) – although out-of-sample accuracy is critical for proper risk management and is required for regulatory purposes.

Analyzing estimation accuracy separately for each lessor, our results suggest that the number of observations within a dataset has an impact on the relative performance of the estimation methods. Whereas sophisticated nonparametric estimation techniques yield, by far, the best results for large datasets, simple OLS-regression performs fairly well for smaller datasets.

Finally, we find that estimation accuracy critically depends on the available set of information. We estimate the LGD at two different points in time, at the execution of the contract and at the point of contractual default. This procedure is of particular importance for leasing contracts because the loan-to-asset value changes during the course of a leasing contract. Furthermore, the Basel II accord requires financial institutions using the advanced internal ratings-based approach (IRBA) to update their LGD estimates for defaulted exposure. To the best of our knowledge, an analysis of this type of update has been neglected in the literature thus far.

The remainder of our study is organized as follows. We review the related literature in Section 2. Section 3 provides an overview of the dataset, defines the LGD measurement, and presents some descriptive statistics. In Section 4, we introduce the methods used in this study. Section 5 reports the empirical results, and Section 6 presents the conclusions of the study.

2. Review of the literature

There are two major challenges in estimating recovery rates for leases with respect to defaulted bank loans or bonds. First, estimates of LGD on loans or bonds take for granted that the recovery rate is bounded within the interval $[0,1]$, which assumes that the bank cannot recover more than the outstanding amount (even under the most favorable circumstances) and that the lender cannot lose more than the outstanding amount (even under the least favorable circumstances). Although the assumption of an upper boundary is justified for bank loans, it does not apply to leasing contracts. As the legal owner of the leased asset, the lessor may retain any value recovered by redeploying the leased asset, even if the recoveries exceed the outstanding claim. In fact, there is some empirical evidence that recovery rates greater than 100% are by no means rare. For example, Schmit and Stuyck reported in their study from 2002 that up to 59% of all defaulted contracts in their sample had a recovery rate that exceeded 100%. Using a different dataset, Laurent and Schmit (2005) found that recovery rates were greater than 100% in 45% of all defaulted contracts. The lower boundary of the recovery rate rests on the implicit assumption of a costless workout procedure. In fact, most empirical studies have neglected workout costs (presumably) because of data limitations. Only Grippa et al. (2005) have accounted for workout costs in their study of Italian bank loans and found that workout costs average 2.3% of total operating expenses. The Basel II accord, however, requires that workout costs are included in the LGD calculation. Thus, when workout costs are incorporated, there is no reason to assume that workout recovery rates must always be non-negative. The second challenge in estimating recovery rates is the bi-modal nature of the density function, with high densities near 0 and 1. This property of workout recovery rates has been well documented in almost all empirical studies, whether of bank loans or leasing contracts (e.g., Laurent and Schmit, 2005).

Because of the specific nature of the recovery rate density function, standard econometric techniques, such as OLS-regression, do not yield unbiased estimates. Renault and Scaillet (2004) applied a beta kernel estimator technique to estimate the recovery rate density of defaulted bonds, but they found that it was difficult to model its bi-modality. Calabrese and Zenga (2010) extended this approach by considering the recovery rate as a mixed random variable obtained as a mixture of a Bernoulli random variable and a continuous random variable on the unit interval and then applied this new approach to a large dataset of defaulted Italian loans. Qi and Zhao (2011) compared fractional response regression to other parametric and nonparametric modeling methods. They concluded that nonparametric methods – such as regression trees (RTs) and neural networks – perform better than parametric methods when overfitting is properly controlled for. A similar result was obtained by Bastos (2010), who compared the estimation accuracy of fractional response to RTs and neural networks.

Despite the growing interest in the modeling of recovery rates, little empirical evidence is available on this topic. Several studies (e.g., Altman and Ramayanam, 2007; Friedman and Sandow, 2005; Frye, 2005) have relied on the concept of market recoveries, which are calculated as the ratio of the price for which a defaulted asset is traded some time after default to the price of that asset at the time of default. Market recoveries are only available for bonds and loans issued by large firms. Workout recoveries were used by Khieu et al. (2011), Dermine and Neto de Carvalho (2005), and Friedman and Sandow (2005). However, Khieu et al. (2011) found evidence that the post-default price of a loan is not a rational estimate of actual recovery realization, i.e., it is biased and/or inefficient. According to Frye (2005), many analysts prefer the discounted value of all cash flows as a more reliable measurement of defaulted assets because: (1) cash flows ultimately become known with certainty, whereas the market price is derived from an uncertain forecast of future cash flows; (2) the market for defaulted assets might be illiquid; (3) the market price might be depressed; and (4) the asset holder might not account for the asset on a market-value basis.

Schmit et al. (2003) analyzed a dataset consisting of 40,000 leasing contracts, of which 140 were defaulted. Using bootstrap techniques, they concluded that the credit risk of a leasing portfolio is rather low because of its high recovery rates. Similar studies were conducted by Laurent and Schmit (2005) and Schmit (2004). In 2002, Schmit and Stuyck found considerable variation in the recovery rates of 37,000 defaulted leasing contracts of 12 leasing companies in six countries. Average recovery rates depend on the type of the leased asset, country, and contract age. De Laurentis and Riani (2005) found empirical evidence that leasing recovery rates are inversely correlated with the level of exposure at default. However, recovery rates increase with the original asset value, contract age, and existence of additional bank guarantees. Applying OLS-regressions to forecast LGDs in that study led to rather poor results: the unit interval was divided into three equal intervals, and only 31–67% of all contracts were correctly assigned in-sample. With a finer partition of five intervals, the portion of correctly assigned contracts decreased even further. These results clearly indicate that more appropriate estimation techniques are needed to accurately estimate recovery rates.

Our study differs from the LGD literature in several crucial aspects. First, we calculate workout LGDs and consider workout costs. Second, we perform out-of-sample testing at contract execution and default, which meets the Basel II requirements for LGD validation. Third, by separately analyzing the datasets of three lessors, we gain insight into the robustness of the estimation techniques.

3. Dataset

This study uses datasets provided by three German leasing companies, which shall be referred to herein as companies A, B,