Journal of Empirical Finance 17 (2010) 120–137

Contents lists available at ScienceDirect

Journal of Empirical Finance

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

Predicting issuer credit ratings using a semiparametric method☆

Ruey-Ching Hwang[a,*], Huimin Chung[b], C.K. Chu[c]

[a] Department of Finance, National Dong Hwa University, Hualien, Taiwan
[b] Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
[c] Department of Applied Mathematics, National Dong Hwa University, Hualien, Taiwan

ARTICLE INFO

Article history:
Received 22 October 2008
Received in revised form 29 June 2009
Accepted 27 July 2009
Available online 8 August 2009

JEL classification:
C14
G20

Keywords:
Industry effect
Issuer credit rating
Market-driven variable
Ordered linear probit model
Ordered semiparametric probit model

ABSTRACT

This paper proposes a prediction method based on an ordered semiparametric probit model for credit risk forecast. The proposed prediction model is constructed by replacing the linear regression function in the usual ordered probit model with a semiparametric function, thus it allows for more flexible choice of regression function. The unknown parameters in the proposed prediction model are estimated by maximizing a local (weighted) log-likelihood function, and the resulting estimators are analyzed through their asymptotic biases and variances. A real data example for predicting issuer credit ratings is used to illustrate the proposed prediction method. The empirical result confirms that the new model compares favorably with the usual ordered probit model.

1. Introduction

Credit ratings play an important role in capital markets. Under the New Basel Capital Accord (Basel II), credit ratings will play an even more central role than they have so far. There are two basic types of credit ratings, the bond rating and the issuer credit rating. While the former measures the likelihood of the default or delayed payment of a bond issue, the latter is an overall assessment of the creditworthiness of a company. Currently, there are many widely recognized credit rating agencies, such as Moody's Investors Service and Standard and Poor's Ratings Services (S&P's), etc. They routinely provide credit ratings for bonds and companies.

This study focuses on the S&P's long-term issuer credit rating (LTR). According to the definition given by S&P's, the LTR focuses on the obligor's capacity and willingness to meet its long-term financial commitments. Based on the Compustat North America (COMPUSTAT) database, in year 2007, there were 8010 companies listed on the New York Stock Exchange, American Stock Exchange, or NASDAQ. However, among those 8010 companies, there were only 18.96% (1519) companies having S&P's LTRs. This result indicates that most of companies listed on those stock exchanges do not have S&P's LTRs, which makes their rating predictions quite valuable to practitioners and regulators. Accordingly, the purpose of this paper is to forecast ratings for those companies "without" S&P's LTRs. For two reasons, we do not pursue the issue of rating forecast for companies "with" S&P's LTRs. First, if a company is once rated by S&P's, then it will be rated again unless a special event, for example bankruptcy, happens to the company. Second, the continuously rated companies have relatively unchanged rating categories in general (Galil, 2003; Pettit et al., 2004). Thus it seems less interesting in predicting the ratings of these companies.

☆ The authors thank the reviewer and the editor for their valuable comments and suggestions which greatly improve the presentation of this paper. The research was supported by National Science Council, Taiwan, Republic of China.

* Corresponding author. Tel.: +886 38633139; fax: +886 38633510.
E-mail address: rchwang@mail.ndhu.edu.tw (R.-C. Hwang).

0927-5398/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2009.07.007

There are several well-known statistical techniques for constructing credit rating predictions. These techniques include multiple regression analysis (Horrigan, 1966; Pogue and Soldofsky, 19692; West, 1970), multiple discriminant analysis (Pinches and Mingo, 1973, 1975; Altman and Katz, 1976), ordered linear probit model (OLPM; Kaplan and Urwitz, 1979; Ederington, 1985; Gentry et al., 1988; Hwang et al., 2008), and ordered and unordered linear logit models (Ederington, 1985), etc. Altman et al. (1981) provides a detailed introduction of statistical classification models. The common principal of these approaches is that they are developed using single-period data and parametric models. Credit rating forecasting models based on multiple-period data with independent assumption include, for example, Blume et al. (1998), Poon (2003), and Güttler and Wahrenburg (2007) employing the idea of OLPM. Other approaches based on machine learning techniques, for example, Bayesian networks (Wijayatunga et al., 2006) and support vector machines and neural networks (Huang et al., 2004) were also considered in the literature for credit rating prediction. Basically, the latter approaches based on machine learning techniques are more complicated in computation and interpretation than those based on parametric models.

To forecast S&P's LTRs, the prediction methods based on OLPM and an extension of OLPM will be used in this paper.[1] The OLPM is simply constructed by imposing a linear regression relationship between S&P's LTRs and predictor variables. Its important parameters are determined by maximizing a log-likelihood function. However, if the underlying regression function is not linear, then the advantages of OLPM in explaining and predicting will not be realized. To avoid this potential pitfall, we show in this paper that the idea of semiparametric logit model in Hwang et al. (2007) can be directly extended to OLPM. Specifically, we shall propose an ordered semiparametric probit model (OSPM) for credit rating prediction by replacing the linear regression function in OLPM with a semiparametric function. The proposed OSPM is built on the works of OLPM but needs not assuming any parametric form for the regression function. Thus it is much more flexible in modeling the regression function. Furthermore, the proposed method is developed under the concept of local likelihood, and it turns out that the important parameters in OSPM can be estimated by maximizing a local (weighted) log-likelihood function. Thus the required computation for OSPM is as simple as that for OLPM.

To apply OLPM and OSPM to predict S&P's LTRs, the twenty-four potential predictors in Hwang et al. (2008) for studying important predictors of S&P's LTRs in year 2005 were considered in our data analysis section. These variables include four market-driven variables (Shumway, 2001; Bharath and Shumway, 2008), nineteen accounting variables (Altman, 1968; Poon, 2003; Pettit et al., 2004), and industry effects (Chava and Jarrow, 2004; Pettit et al., 2004). The studied data were collected from COMPUSTAT and Center for Research in Security Prices (CRSP) databases. Our sample consisted of 779 companies receiving S&P's LTRs in April 2007 and having complete values of the twenty-four potential predictors. The sample was further divided into the estimation sample and holdout sample based on the longevity of S&P's LTR (Hwang et al., 2008).[2] According to S&P's Research Insight North America Data Guide (2004, p. 54), S&P's began to use the term LTR on September 1, 1998. Companies receiving S&P's LTRs in consecutive nine years (April 1999–April 2007) were classified into the estimation sample. The rest of the sampled companies were classified into the holdout sample. Based on the division principle, 413 companies were divided into the estimation sample and 366 companies into the holdout sample.

To examine whether our estimation and holdout samples induced selection bias, a procedure based on OLPM with sample selection was performed using LIMDEP 8.0 to test the null hypothesis of no selection bias caused by the above sample division principle. The result of the test shows no rejection of the null hypothesis of interest at 5% level of significance. Before performing the selection bias test, a forward selection procedure based on minimizing classification error rate on the estimation sample (Härdle et al., 2008) was used to objectively determine effective predictors for OLPM. The final list of the selected predictors includes industry effects, two market-driven variables assessing a firm's market capitalization and risk, and two accounting variables measuring a firm's financial leverage and profitability. The values of estimated coefficients of the selected market-driven and accounting variables all agree with their expected signs. This indicates that market-driven variables and industry effects are also important to determine S&P's LTRs. Our variable selection result coincides with that obtained by Hwang (2008) for predicting ratings in year 2005. On the other hand, to study the difference between unsolicited and solicited ratings, Poon (2003) suggested profitability and sovereign credit risk as two major factors in determining S&P's LTRs. Furthermore, to assess biases in credit ratings assigned by Moody's and S&P's for near-to-default issuers, Güttler and Wahrenburg (2007) used accounting and macroeconomic variables as major determinants of issuer credit ratings.

In Section 3, we describe one real data set and provide some summary statistics. The summary statistics show that the predictors under consideration have reasonable power in discriminating the creditworthiness of companies. The real data set was analyzed using methods based on OLPM and OSPM. The prediction performance of each method was measured by the total error rate obtained from the holdout sample. By the error rates summarized in Section 3, we conclude that the prediction method based on OSPM has better performance. The empirical results in Section 3 also demonstrate that the functional form between the S&P's LTR assessment and the selected continuous predictors is nonlinear. This indicates that OLPM may not be adequate in explaining and predicting S&P's LTRs. Thus, by the empirical results, OSPM has potential to be a powerful credit rating forecasting model.

[1] Due to the superiority in explaining and predicting, OLPM has been adopted for multiple-class prediction by a number of studies such as Kaplan and Urwitz (1979) and Gentry et al. (1988), etc. Also, the test procedure of sample selection bias is only available for OLPM (Greene, 2002). On the other hand, it is not suggested using discrete explanatory variables in multiple discriminant analysis (Johnson and Wichern, 2002, p. 641). In this paper, industry effects on S&P's LTRs were estimated through coefficients of six industry indicator variables. Given these industry indicator variables, it is not adequate to use multiple discriminant analysis to predict S&P's LTRs.

[2] Given the pool of companies with S&P's LTRs in April 2007, by comparing the longevity of S&P's LTR, our estimation companies solely correspond to the rated ones, and our holdout companies the newcomers. Their purified composition agrees with our purpose to forecast ratings for companies without S&P's LTRs. On the other hand, one may separate the sampled companies by random allocation. Random allocation has the advantage of eliminating the need to test for selection bias since the resulting estimation and holdout samples have the same composition structure. However, each of the latter samples contains both rated companies and newcomers. Such mixed composition does not agree with our prediction purpose.