Expert Systems With Applications 58 (2016) 210–220

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Nonparametric machine learning models for predicting the credit default swaps: An empirical study

Youngdoo Son [a], Hyeongmin Byun [b], Jaewook Lee [c,*]

[a] Research Institute of Engineering and Entrepreneurship, Seoul National University, Republic of Korea
[b] Department of Industrial Engineering, Seoul National University, Republic of Korea
[c] Department of Industrial Engineering, Seoul National University, Republic of Korea

ARTICLE INFO

Article history:
Received 18 December 2014
Revised 28 March 2016
Accepted 29 March 2016
Available online 4 April 2016

Keywords:
Financial forecasting
Nonparametric models
Credit default swap
Empirical analysis

ABSTRACT

Credit default swap which reflects the credit risk of a firm is one of the most frequently traded credit derivatives. In this paper, we conduct a comprehensive study to verify the predictive performance of nonparametric machine learning models and two conventional parametric models on the daily credit default swap spreads of different maturities and different rating groups, from AA to C. The whole period of data set used in this study runs from January 2001 to February 2014, which includes the global financial crisis period when the credit risk of firms were very high. Through experiments, it is shown that most nonparametric models used in this study outperformed the parametric benchmark models in terms of prediction accuracy as well as the practical hedging measures irrespective of the different credit ratings of the firms and the different maturities of their spreads. Especially, artificial neural networks showed better performance than the other parametric and nonparametric models.

1. Introduction

Credit market is one of the most important financial markets and has received wide attention especially from the credit crisis in 2008. A default probability is one of the typical measures to represent the credit risk of a firm or nation but it is difficult to determine since many firms and nations, or obligors, are linked by various contracts and obligations and thus a credit-related event, including default, of one obligor may affects many other obligors. The default probability is usually measured by the credit derivatives traded in the credit market because their prices are not highly affected by the other factors than credit risk unlike defaultable bond prices. For example, Bühler and Trapp (2009) showed that the 95% of the credit default swap (CDS) spread stems from the credit risk while only 4% from the liquidity. The mispricing of these derivatives can lead to misunderstanding of default probability; thus, accurate pricing for credit derivatives from the credit crisis period has become an important consideration.

During the last two decades, many researches have been made to price credit derivatives and their models can be categorized into two classes of models. One class of models, called *structural models*, assume that a certain stochastic process for the fundamental value of the firm and defines an event of default as the fundamental value hits a predetermined barrier (Black & Cox, 1976; Finger, Finkelstein, Lardy & Pan, 2002; Merton, 1974). the other class of models, called *reduced-form models* or *intensity-based models*, assume that the default is driven by an exogenous factors and an event of default follows a Poisson process with a stochastic intensity (Cox, Ingersoll, & Ross, 1985; Jarrow & Turnbull, 1995; Vasicek, 1977). There have also existed a large number of studies that compared those models by the predicted credit derivative prices (Bakshi, Madan, & Zhang, 2006; Duffee, 1999; Eom, Helwege, & Huang, 2004; Gündüz & Uhrig-Homburg, 2011; Jones, Mason, & Rosenfeld, 1984; Lyden & Saraniti, 2001; Ogden, 1987) but there has not been a robust conclusion that a certain model overwhelms the others for pricing and predicting credit derivatives traded in the real market.

On the other hand, nonparametric learning models have extensively been used to predict financial time series in recent years due to their flexibility which fits the models to the data well. Most of those results have been focused on the stock (Chen, Shih, & Wu, 2006; Liao & Chou, 2013; Son, Noh, & Lee, 2012; Ticknor, 2013) and its derivative markets (Han & Lee, 2008; Hutchinson, Lo, & Poggio, 1994; Park, Kim, & Lee, 2014; Park & Lee, 2012; Yang & Lee, 2011) and achieved accurate prediction results. However, relatively a few studies have been conducted for the other markets including the fixed-income market (Cao & Tay, 2003; Kim & Noh, 1997) and

* Corresponding author. Tel.: +8228807176.
E-mail addresses: hand02@snu.ac.kr (Y. Son), ragnarok@snu.ac.kr (H. Byun), jaewook@snu.ac.kr (J. Lee).

http://dx.doi.org/10.1016/j.eswa.2016.03.049
0957-4174/

the foreign exchange market (Bhattacharyya, Pictet, & Zumbach, 2002; Osuna, Freund, & Girosi, 1997). For the credit market, most of studies using learning methods have concentrated on credit rating analysis. Lee (2007) and Kim and Ahn (2012) used support vector machines to classify the rating of firm and Huang, Chen, Hsu, Chen, and Wu (2004) classified the rating of corporate bonds using both support vector machines and artificial neural networks. For credit derivatives pricing, Gündüz and Uhrig-Homburg (2011) applied the support vector regression (Drucker, Burges, Kaufman, Smola, & Vapnik, 1997) to predicting one-dimensional output corresponding five-year maturity CDS spread of one firm using the spreads of other firms at the same moment called a cross-sectional design or using the past value of spreads of the same firm called a time series design and compared it with those of the Merton model (Merton, 1974) and the constant intensity model (Jarrow & Turnbull, 1995). However, only one specific spread was used in prediction although considering and predicting the spreads of other liquid maturities at the same time are practically important and no advanced state-of-the-art machine learning models other than support vector regression were used for comparison.

To our knowledge, no empirical studies have been made that prices and predict the CDS spreads with diverse maturities simultaneously using several nonparametric learning models and even the earlier studies on other financial markets such as stocks or options were made to determine and predict one-dimensional outputs for its price values. In this paper we aim to conduct a comprehensive study that compares the predictive power of several nonparametric models using the real market CDS spread data of diverse maturities from January 2001 to February 2014 as well as those of different credit ratings. Since the spreads with several different maturities are predicted simultaneously, the employed models can capture the effect of the past spreads with different maturities on the current spreads. For our experiment, we applied four well-known state-of arts nonparametric learning regression models (*support vector regression* (Drucker et al., 1997), *artificial neural networks, Bayesian neural networks*, and *Gaussian processes* (Cressie, 1993; Rasmussen, 1996) to prediction of six dimensional outputs consisting of CDS spreads with six different maturities, 1, 2, 3, 5, 7, and 10 years. Also to verify the relative predictive performance of nonparametric learning models, we have applied two benchmark parametric models, called constant intensity model (Jarrow & Turnbull, 1995) and Cox-Ingersoll-Ross (CIR) intensity model (Cox et al., 1985).

The organization of this paper is as follows. In the next section, we review some literatures related to this study. Then, we give a brief description of nonparametric machine learning models and two benchmark parametric models, constant intensity model and CIR model, used to price the credit derivatives in Section 3. We describe the data used for this study and explain the design of the experiment and the performance measure and present the experimental results with some discussions on the results in Section 4 and 5. Finally we conclude this paper and provide directions for future work in Section 6.

2. Related work

Since the credit market is one of the largest markets among several financial markets, there have been a large number of literatures that related to the analysis of the CDS spreads. In this section, we reviewed some milestone literatures related to the parametric models of the CDS spreads and some recent studies related to the analysis of the CDS spreads from the different point of view.

There have been two main streams of the parametric models about credit derivatives, the *structural models* and the *reduced-form models*. In structural models, the asset value of the firm is assumed to follow a certain stochastic process, usually a geometric Brownian motion, and the even of default of the firm is defined as when the firm's asset value hits a predetermined barrier or becomes below it. Merton (1974) assumed that the asset value of the firm follows the geometric Brownian motion and the default occurs if the value of the firm is below the liability at the maturity. Therefore the default can only occur at the maturity and the formula for European option in Black and Scholes (1973) can be applied to find the theoretical value of the CDS contract. In contrast with Merton (1974), the first passage model (Black & Cox, 1976) defined the event of default of a firm as when the asset value of the firm hits the predetermined barrier. Thus the default can occur any time before maturity in this model. CreditGrades model (Finger et al., 2002) is one of the recent and widely used structural models, which imposed the randomness on the default barrier. The main framework of CreditGrades model is similar to the first passage model, but the barrier follows a lognormal distribution to reflect the uncertainty in the recovery. In reduced-form models, another main stream of credit derivative models, assumes that the default is caused by the exogenous factors which are independent of market information and there is no information about the arrival time of default. Thus, the occurence of default is usually assumed to follow the poisson process in the reduced-form models. Jarrow and Turnbull (1995) firstly proposed the reduced-form model with the constant intensity of the poisson process but the inhomogeneous poisson process like Vasicek model (Vasicek, 1977), CIR model (Cox et al., 1985), and even the intensities including jumps (Schoutens & Cariboni, 2009) have also been employed to analyze the credit derivatives. However, both of these structural and reduced-form models explained above are parametric models that assume a certain form of the solution. Although, these parametric models have an advantage in finding the financial implication like the explanation of stylized facts, they do not concentrate on fitting the data itself compared to the machine learning models. The more detailed description of these parametric models for credit derivatives is well summarized in Lando (2004) and Schoutens and Cariboni (2009).

CDS has also been extensively studied and used in recent literatures. Brigo and El-Bachir (2010) proposed the exact pricing model of defaultable swaptions under the assumption of the jump diffusion process and used the CDS spreads for calibration. Cont and Kan (2011) proposed the multivariate time series model for the CDS spreads of three different firms and found that the proposed model showed better performance than the intensity models with jumps and random walk models for predicting loss quantiles of CDS portfolios. Jarrow, Li, and Ye (2011) found the statistical arbitrage opportunities based on reduced-form models. Bianchi and Fabozzi (2015) compared several parametric models, Brownian motion, three Lèvy processes, and a Sato process, for the prediction of CDS prices. Gündüz and Uhrig-Homburg (2011) employed the support vector regression, Merton model, and the constant intensity model, which are a nonparametric machine learning model, a structural model, and a reduced-form model, respectively, to predict the CDS spreads and compared the results. However, they only used the CDS spreads with one specific maturity and no state-of-the-art machine learning models other than the support vector regression were not employed. Most of the literatures above, except Gündüz and Uhrig-Homburg (2011), also focused on or used the parametric models to find the CDS spreads. However, since these parametric models assume a special form of the model with predefined parameters, they cannot use the information which is not predefined in the model compared to the machine learning models which are versatile in the kinds and number of input variables. For example, Galil, Shapir, Amiram, and Ben-Zion (2014) found that the three variables, stock return, the change in the stock return volatility, and the change in the median CDS spreads in the rating class, were effective to the change of the CDS spreads and