Journal of Empirical Finance 35 (2016) 43–67

Contents lists available at ScienceDirect

Journal of Empirical Finance

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

Using Merton model for default prediction: An empirical assessment of selected alternatives☆

Zvika Afik [a,*], Ohad Arad [b,1], Koresh Galil [b,2]

[a] Guilford Glazer faculty of Business and Management, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
[b] Department of Economics, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel

ARTICLE INFO

Article history:
Received 7 June 2013
Received in revised form 25 September 2015
Accepted 29 September 2015
Available online 13 October 2015

JEL classification:
G17
G33
G13

Keywords:
Credit risk
Default prediction
Merton model
Bankruptcy prediction
Default threshold
Assets volatility

ABSTRACT

It is surprising that although four decades passed since the publication of Merton (1974) model, and despite the development and publications of various extensions and alternative models, the original model is still used extensively by practitioners, and even academics, to assess credit risk. We empirically examine specification alternatives for Merton model and a selection of its variants, concluding that default prediction goodness is mainly sensitive to the choice of assets expected return and volatility. A Down-and–Out Option pricing model and a simple naïve model outperform the most common variants of the Merton model, therefore we recommend using the simple model for its easy implementation.

1. Introduction

Merton (1974) and Black and Scholes (1973) presented the basic approach for the valuation of stocks and corporate bonds as derivatives on the firm's assets. Merton (1974) is a structural model used for default prediction and corporate bond pricing, viewing the firm's equity as a call option on its assets, because equity holders are entitled to the residual value of the firm after all its obligations are paid. Many theoretical studies suggested models that relax some of the Merton model restrictive assumptions.[3] However, empirical literature on default prediction mainly focused on the application of the original model.

The application of the Merton model for default prediction requires the specification of three parameters: assets value, assets volatility and assets expected returns. The most simple and common approach in the literature is the equation-system approach in which the first two parameters are solved using an equation-system and the third parameter is set independently The equation

☆ We gratefully acknowledge comments from Kevin Aretz, Doron Kliger, Gal Zehavi, Fernando Moreira, and the anonymous reviewers of this paper. We also thank seminar participants in 2012 conference of ORSIS in Jerusalem and 2012 conference of MFS in Krakov for valuable comments and suggestions.
* Corresponding author. Tel.: +972 8 647 2205; fax: +972 53 955 6352.
E-mail addresses: afikzv@som.bgu.ac.il (Z. Afik), ohadarad@bgu.ac.il (O. Arad), galilk@bgu.ac.il (K. Galil).
[1] Tel.: +972 54 4479340; fax: +972 8 647 2 941.
[2] Tel.: +972 8 647 2310; fax: +972 8 647 2 941.
[3] See for example Black and Cox (1976), Geske (1977), Longstaff and Schwartz (1995), Collin-Dufresne and Goldstein (2001), Hsu et al. (2010), Leland (1994), Leland and Toft (1996), Acharya and Carpenter (2002).

http://dx.doi.org/10.1016/j.jempfin.2015.09.004
0927-5398/

system was initially proposed by Jones et al. (1984) and Ronn and Verma (1986) and was recently used by Campbell et al. (2008). A different approach is the maximum likelihood (ML) method developed by Duan (1994) and applied to the estimation of structural credit risk models by Ericsson and Reneby (2005) among others, allowing the joint estimation of all three parameters.[4] Another strand of the literature deals with hybrid models, combining structural models with accounting variables (e.g. Agarwal and Taffler, 2008; Bauer and Agarwal, 2014; Li and Miu, 2010; Wu et al., 2010). This paper focuses on the application of the equation-system method in default prediction and leaves further evaluation of the ML method and hybrid models for future research.

A major benchmark in these studies is the KMV model. KMV was founded in 1989 offering a commercial extension of Merton's model using market-based data. In 2002 it was acquired by Moody's and became Moody's-KMV. KMV published a number of papers which reveal some of its methods (see Crosbie and Bohn, 2003; Keenan and Sobehart, 1999; Sobehart et al., 2000). Some of the specifications made by KMV were adopted by the academic literature. Vassalou and Xing (2004), Campbell et al. (2008), Aretz and Pope (2013) are examples for such studies.

Only a few studies attempted to evaluate the accuracy of Merton's model for default prediction under these specifications. Hillegeist et al. (2004) compared the predictive power of the Merton model to Altman (1968) and Ohlson (1980) models (Z-score and O-score) and concluded that the Merton model outperforms these models. Duffie et al. (2007) showed that macroeconomic variables such as interest rate, historical stock return and historical market return have default prediction ability even after controlling for Merton model's distance to default. Campbell et al. (2008), using a hazard model, combined Merton model default probability with other variables relevant to default prediction. They also found that Merton model probabilities have relatively little contribution to the predictive power. Bharath and Shumway (2008) presented a "naïve" application of Merton model that outperformed the iterative application of Merton model (based on presumably Moody's-KMV specifications).[5] Another line of literature, such as Jones et al. (1984), Ericsson and Renault (2006), Eom, Helwege and Huang (2004), Huang and Huang (2012), and Feldhütter and Schaefer (2015), examined structural models ability to explain credit spreads.

In this paper we examine the sensitivity of Merton model default prediction performance to its parameter specifications. We assess the causes for this sensitivity and evaluate the performance of a wide range of model alternatives, including those suggested by other recent studies. We conclude by providing a few prescriptions to enhance the model accuracy and suggesting a very simple model, which provides excellent discriminatory power for a low computation effort. Model wise we evaluate the textbook two-equation Merton model, its down and out (DaO) barrier alternative, the iterative model which is widely believed to be that of KMV, and single equation models and shortcuts including Bharath and Shumway (BhSh) naïve model, Charitou et al. (CDLT), and our simple naïve model (SNM).[6] In each model we focus on its three main components: the default threshold, the expected return on the firm assets and the firm assets return volatility (hereafter, asset volatility).[7] For this purpose we construct a sample with annual observations of firms rated by S&P from the merged CRSP/Compustat database during the period 1989 to 2012. We also gather information on default events during 1990 to 2013 from Standard and Poor's (S&P) and Moody's rating agencies reports. After filtering our sample includes 26,579 annual observations of 2,534 firms, of which 306 observations defaulted in the following year.

For each specification of each assessed model, we construct a Receiver Operating Characteristic (ROC) curve. This method is relatively common for the comparison of prediction models since it does not require setting a priori the desired cutoff point between cost of type I error and cost of type II error. Another advantage of using ROC curves, compared to methods used in some prior studies, is that it enables statistical inference with the non-parametric test suggested by DeLong, DeLong and Clarke-Pearson (1988), testing the statistical significance of the differences between the ROC curves (of two models). For robustness, we also include partial area under the curve (pAUC) calculations and test for pAUC differences, often at a few false positive rate levels. Prior studies, such as Bharath and Shumway (2008), focused mainly on the rate of defaulters within the first deciles of firms (highest predicted default probabilities) and did not offer a robust statistical test for differences between models.

Another approach we use to understand the adequacy of various specifications is the study of firms' characteristics changes on a path to default. For this purpose, we focus on 101 defaulting firms with data available for the five years preceding the default event and compare their level of debt, stock returns, equity volatility and assets volatility to those of a group of 101 non-defaulting firms.

We find that Merton model accuracy is only slightly sensitive to the specification of the default threshold. We explain that this is a result of the calculated assets value and volatility dependence on the default threshold. On the one hand, ceteris paribus, a low setting of default threshold for risky firms reduces their probability of default. On the other hand, such misspecification also causes overestimation of assets volatility and underestimation of assets value, thus increasing the default probability. Therefore, a deviation of the default threshold from the common practices has a relatively small effect on the model accuracy.

We also show that, not-surprisingly, using historical equity return as a proxy for expected assets return is questionable.[8] In particular, realized returns for risky firms are low and sometimes negative. While negative stock returns may be a predictive

[4] See also Duan et al. (2004).

[5] Chava and Purnanandam (2010) used the naïve model as a proxy for credit risk.

[6] Charitou et al. (2013) is a comprehensive study, similar to this work, aiming to compare various specifications of Merton model. The potential spectrum of methods and specification is too broad to be included in a single paper, hence we regard their work (denoted hereafter by CDLT) and our research as complementary with some essential overlap. This overlap is required to ensure that method comparison is based on identical database. A similar overlap exists also between CDLT, Bharath and Shumway (2008), and other preceding papers, each repeats some of the methods incorporated in its respective prior literature. In the same vein, we include CDLT proposed methodology to estimate asset volatility (σ_{CDLT}) and asset drift (μ_{CDLT}) in our study.

[7] The terms *default barrier*, *default boundary*, *default point* and *default threshold* are interchangeably used in the literature. In this paper we use *default threshold* to avoid confusion with the barrier of barrier options.

[8] This specification was used by Bharath and Shumway (2008), in their naïve model.