

Article

Validation of the Merton Distance to the Default Model under Ambiguity

Wei-ling Chen and Leh-chyan So *

Department of Quantitative Finance, National Tsing Hua University, Hsinchu 30013, Taiwan;
E-Mail: qf@my.nthu.edu.tw

* Author to whom correspondence should be addressed; E-Mail: lcs0@mx.nthu.edu.tw;
Tel.: 886-3-5742223.

*Received: 13 December 2013; in revised form: 15 February 2014 / Accepted: 17 March 2014 /
Published: 25 March 2014*

Abstract: Bharath and Shumway (2008) provide evidence that shows that it is the functional form of Merton's (1974) distance to default (DD) model that makes it useful and important for predicting defaults. In this paper, we investigate whether the default predictability of the Merton DD model would be affected by taking investors' ambiguity aversion into consideration. The Cox proportional hazard model is used to compare the forecasting power of Bharath and Shumway's naive model, which retains the functional form of the Merton DD model and computes the default probability in a naive way, with our new model, which treats investors' ambiguity aversion as additional information. We provide evidence to show that our new model performs better than Bharath and Shumway's naive model. In addition, our empirical results show that the statistical significance of Bharath and Shumway's naive default probability is retained in the credit default swap (CDS) spread regressions, though the sign of the coefficient is changed. However, both the sign and the statistical significance of our model are retained in the CDS spread regressions.

Keywords: Merton's distance to default model; naive model; default probability; ambiguity aversion; consumer confidence index; statistical significance

JEL classification: G11, G32, G33

1. Introduction

Default can happen if the value of the issuer's total assets is less than the value of the debt obligations should the issuer be unable to make the required payments. There are many different approaches to evaluate default probabilities. The structural model and reduced-form model are the main measures of default probabilities. In the structural model, or the Merton distance to default (DD) model, which is inspired by Merton's [1] bond pricing model, a default-triggering event is explicitly defined as a firm's failure to pay debt obligations by means of modeling the equity value of the firm as a call option on the firm's value, with the firm's face value of the debt as a strike price. On the other hand, for the reduced-form model, a default-triggering event is defined as an unexpected event that is governed by an exogenous default-intensity process.

The Merton model is widely used to measure the distance to default by Moody's KMV (Kealhofer, McQuown and Vasicek). This was used previously to provide quantitative credit analysis tools to creditors and investors until the acquisition in 2002 by Moody's Analytics and is now part of Moody's Analytics Enterprise Risk Solutions. However, the original formula is complicated. Campbell, Hilscher and Szilagyi [2] concluded that the strong predictive power of the Merton DD model came from the functional form imposed by the Merton model. Bharath and Shumway [3] also provide evidence to show that it is the functional form of the Merton DD model that makes it useful and important for predicting defaults. Rather than simultaneously solving two nonlinear equations or implementing a complicated iterative procedure, as suggested by Crosbie and Bohn [4] and Vassalou and Xing [5] to obtain default probabilities, they proposed a naive default forecasting model, which retains the structure of the Merton DD model, but with a simpler and easier calculation procedure.

It is recognized that simplicity does not mean simplistic. Recently, significant works in the literature have focused on the issue of ambiguity aversion. However, the discussion of ambiguity aversion is not a novel issue, but has a long history. We use the following example, which was adapted from Ellsberg's [6] famous experiment, to illustrate the meaning of ambiguity aversion. The experiment subjects are provided with two urns, each of which contains several red and white balls. There is no difference in the total number of balls in each urn. The balls in one urn are half red and half white. However, the ratio in the other urn is unobservable to the subjects. The rule is that \$1 will be awarded when a red ball is drawn. Although the two urns promise the same expected return, an ambiguity-averse subject will be reluctant to draw a ball from the latter, because of a tendency to think pessimistically, leading to an expectation that the other urn contains lower odds.

Ambiguity aversion, or uncertainty aversion, is a wholly different concept from risk aversion. Risk aversion is an attitude that penalizes the expected return of a risky investment. Consider the following experiment. The experiment subjects are provided with two potential investments. One investment pays off either \$1 or \$0 with equal probability. Another guarantees \$0.5 for sure. Although the two investments offer the same expected return, a risk-averse investor will resist investing in the former. Ambiguity aversion, however, represents a lack of confidence in probability estimates. According to Bewley ([7], p. 1134), "... an increase in uncertainty aversion increases the size of the set of subjective distributions. An uncertainty-neutral decision maker has only one subjective distribution and so is Bayesian."

Generally speaking, the approaches of addressing ambiguity aversion may be divided roughly into two categories. One approach is called the recursive multi-priors utility approach, which addresses ambiguity aversion through the multi-prior expected utility (for example, [8–11]). Gilboa and Schmeidler [12] contribute to this field by initially establishing a well-defined max-min expected utility with multiple priors. They are famous for developing a useful axiomatic theory in this field. Following the approach of Gilboa and Schmeidler [12], Easley and O’Hara [13] discussed microstructure and ambiguity and showed that ambiguity could reduce participation by both investors and issuers. The incomplete preferences of Bewley [14], the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji [15] and the variational preferences of Maccheroni, Marinacci and Rustichini [16] are interesting extensions to this approach.

This paper uses the other approach, called the robust control approach, to depict investors’ ambiguity aversion. This describes ambiguity through a set of priors and introduces the penalty function to a general utility function in order to capture investors’ ambiguity aversion (for example, [17–21]). Theoretically, economic agents are assumed to possess perfect knowledge of the data generating process and to know the underlying probability law exactly. However, we cannot ignore the fact that investors are not certain of the correctness of the model. Investors with higher ambiguity aversion worry about a worst-case scenario. Therefore, the investor will choose alternative models that are further from the reference model. Hence, the robust control approach assigns a smaller penalty to more distant perturbations. However, if ambiguity aversion is low, the investor will choose alternative models that are similar to the reference model. As a result, the robust model assigns a greater penalty to more distant perturbations. That is, the penalty is inversely related to the investor’s degree of ambiguity aversion.

There have been numerous studies investigating ambiguity issues in asset prices and equity premia (for example, [22–24] for a literature review on ambiguity in asset pricing and portfolio choice). By using the robust control technique, So [25] derives an adjusted Black–Scholes pricing formula, which maintains the functional form of the original Black–Scholes formula, but adds an additional parameter depicting uncertainty aversion to compute the risk-free interest rate and dividend yield. We find that managers’ uncertainty aversion about the project value raises his subjective evaluation of the real options. If we compare the firm value and the equity value to the project value and the real options in that paper, we can infer that investors’ ambiguity about the firm’s value will change the valuation of the equity. In addition, since the risk-free interest rate has been adjusted, it is possible that the default rate of the fixed income debt will also be modified. This paper could be treated as an extension of [25] from the theoretical real options analysis to the applied risk management field. The purpose of this paper is to develop a structural credit pricing model with ambiguity aversion as additional information and to derive a risk-neutral default probability under uncertainty about the firm’s value. In addition, we use the new model to examine the validation of Merton’s DD model under ambiguity. To the best of our knowledge, ambiguity aversion has not been applied to examine the default prediction issue addressed here.

As in risk aversion, investors’ ambiguity aversion is a subjective parameter, so one cannot easily observe its aggregate value or assign it a reasonable value. The Consumer Confidence Index (CCI), however, is an observed indicator, which is defined as the degree of optimism on the state of the economy that consumers are expressing through their activities surrounding saving and spending. In

the USA, the CCI is issued monthly by the Conference Board, which was started in 1967 and is benchmarked to 1985. Ait-Sahalia and Brandt [26] also mentioned that the CCI and other economic indexes are suitable to capture market uncertainty. Buraschi and Jiltsov [27] believe that the CCI has an important link with market conditions, so they used the CCI to control uncertainty and added it to construct the model and investigate option markets.

Following the literature, we use the CCI in our empirical study as an inverse proxy to the level of ambiguity aversion. We modify Bharath and Shumway’s naive model with the CCI and focus on how accurate the default predictability of our new model is compared with the original model.

The organization of the remainder of the paper is as follows. The models are presented in Section 2. The data are described in Section 3. The empirical results are analyzed in Section 4, and the conclusions are given in the final section.

2. Models

2.1. Naive Merton DD Model

In this section, we use several equations and the notation used in Bharath and Shumway’s [3] naive Merton DD model to enable comparisons to be made.

Merton’s model [1], which is regarded as a structural model, considers that the firm’s equity is treated as a call option on the underlying value of the firm’s assets, with a strike price equal to the face value of a firm’s zero-coupon bond debt maturing in T . The firm’s value, V , is assumed by following the geometric Brownian motion, with an expected return of μ and volatility of σ_V :

$$\frac{dV}{V} = \mu dt + \sigma_V dB. \tag{1}$$

Merton’s model comprises two nonlinear simultaneous equations, one being the Black–Scholes formula:

$$\begin{aligned} E &= VN(d_1) - Fe^{-rT}N(d_2) \\ d_1 &= \frac{\ln\left(\frac{V}{F}\right) + \left(r + \frac{\sigma_V^2}{2}\right)T}{\sigma_V\sqrt{T}} \\ d_2 &= d_1 - \sigma_V\sqrt{T} \end{aligned} \tag{2}$$

where V is the asset of the firm, σ_V is the volatility of the firm’s asset, E is the market value of the firm’s equity, F is the face value of the firm’s debt, r is the risk-free rate and $N(\cdot)$ is the cumulative standard normal distribution.

The other equation is from Ito’s Lemma:

$$\sigma_E E = \frac{\partial E}{\partial V} \sigma_V V = N(d_1) \sigma_V V \tag{3}$$

where σ_E is the volatility of equity.

After solving these equations, we can obtain the distance to default and the default probability as:

$$\pi_{Merton} = N(-d_2) = N(-Merton DD) \tag{4}$$

where *Merton DD* is the distance to default and equals d_2 .

Rather than solving these two nonlinear simultaneous equations to obtain the value and the volatility of the firm’s asset, Bharath and Shumway [3] directly assign:

$$V = E + F \tag{5}$$

and use an approximation to estimate the volatilities:

$$naive \sigma_D = 0.05 + 0.25 \times \sigma_E \tag{6}$$

where $naive \sigma_D$ $naive \sigma_D$ means the volatility of the firm’s debt in the naive model. Hence, the firm’s asset volatility in the naive model, called $naive \sigma_V$, can be represented as:

$$\begin{aligned} naive \sigma_V &= \frac{E}{E + F} \sigma_E + \frac{F}{E + F} naive \sigma_D \\ &= \frac{E}{E + F} \sigma_E + \frac{F}{E + F} (0.05 + 0.25 \times \sigma_E) \end{aligned} \tag{7}$$

In addition, they replace the risk-free rate in the original model with $naive \mu$, which is the firm’s equity return over the previous year. Finally, the naive distance to default is given as follows:

$$Naive DD = \frac{\ln \left[\frac{(E + F)}{F} \right] + (naive \mu - 0.5 \times naive \sigma_V^2) T}{(naive \sigma_V) \sqrt{T}} \tag{8}$$

Therefore, we obtain the naive default probability as:

$$\pi^{naive} = N(-Naive DD) \tag{9}$$

2.2. Merton’s DD Model under Ambiguity

So [25] derived an adjusted Black–Scholes pricing formula to filter uncertainty from risk in real options analysis by using the robust control technique. If we compare the firm value to the project value in that paper, a similar framework could be applied to solve the problem considered here. In this section, we briefly summarize the theoretical model mentioned in [25]. To save space and concentrate our focus, we omit mentioning proofs here. We recommend that interested readers could refer to [25] for the more technical parts.

In our model, it is assumed that the risk-free interest rate is r , a constant, and the firm’s value, V , follows a geometric Brownian motion with an expected return μ , dividend yield q and volatility σ_V :

$$\frac{dV}{V} = (\mu - q)dt + \sigma_V dB \tag{10}$$

When an investor is under ambiguity and is not confident about probability estimates, a misspecification problem may arise. We use the robust control technique to depict the investor’s ambiguity aversion. The investor would consider an alternative model, Q^ξ , which is not the same as the reference model, Q . Following Uppal and Wang [18], we treat the investor as an expected utility-maximizer facing a three-security portfolio (containing the firm, the equity and riskless asset) allocation problem. The Bellman equation is as follows:

$$\sup_{C, \omega} \inf_{\delta} \left\{ \begin{aligned} & u(C) - \rho J_t + W J_w \left[r + \omega(v - r) - \frac{C}{W} \right] + \frac{1}{2} W^2 J_{ww} \omega^2 \sigma_V^2 \\ & + J_w W \omega \sigma_V^2 \delta + \frac{\Theta(J)}{2} \delta^2 \phi \sigma_V^2 \end{aligned} \right\} = 0 \tag{11}$$

where J refers to the expected utility in continuous time. The investor has to choose consumption, C , and asset allocation, ω , to his wealth, W , to maximize his utility. The first term in the second line arises from the change in measure from Q to Q^ξ , and the term δ adjusts the drift term. The second term in the second line is related to the penalty function, where $\Theta(J)$ converts the penalty to units of utility. In addition, we use ϕ , the penalty parameter, to depict the investor’s confidence about the reference model.

When the investor is under ambiguity, the tendency is to worry about a worst-case scenario. The investor tends to choose alternative models that are more distant from the reference model. Therefore, the robust control approach specifies a lower penalty to more distant perturbations. On the other hand, when the level of the investor’s uncertainty is low, there is a greater likelihood of choosing alternative models that are similar to the reference model. Therefore, the robust control approach specifies a higher penalty to more distant perturbations. The higher is the level of the investor’s ambiguity aversion, the lower is the value of ϕ . Therefore, the penalty is inversely relative to the investor’s ambiguity aversion, and the reciprocal of ϕ can be used to describe the investor’s ambiguity aversion.

We assume that the investor has a power utility of $u(C_s) = \frac{C_s^{1-\gamma}}{1-\gamma}$ and that the value function takes the form:

$$J(W) = \kappa \frac{W^{1-\gamma}}{1-\gamma} \tag{12}$$

where κ is a constant depending on the parameter of the whole environment. According to Maenhout [20] and Uppal and Wang [18], it is assumed that $\Theta(J) = \frac{1-\gamma}{\gamma} J(W) = \frac{W^{1-\gamma}}{\gamma} \kappa$. After taking derivatives of Equation (11) with respect to ω and some computation, we obtain the optimal investment in the firm for the ambiguity-averse investor:

$$\omega^* = \left(\frac{\phi}{1+\phi} \right) \frac{(\mu - r)}{\gamma \sigma_V^2} \tag{13}$$

The investor’s ambiguity aversion causes a slight downgrade in the risk premium for the firm. Hence, the optimal investment in the firm would be lower than a situation where the data generating process of the firm’s value was known exactly.

By using Ito’s Lemma, we can derive the dynamic process of the investor’s subjective stochastic discount factor under ambiguity for the evaluation of the equity value (that is, the call price):

$$\begin{aligned}
 dJ_w &= \frac{\partial J_w}{\partial t} dt + \frac{\partial J_w}{\partial W} dW + \frac{1}{2} \frac{\partial^2 J_w}{\partial W^2} (dW)^2 \\
 &= \kappa W^{-\gamma} \left\{ (-\gamma) \left[r + \omega^* (\mu - r) + (\omega^*)^2 \sigma_V^2 \delta^* - \frac{C^*}{W} \right] + \frac{1}{2} \gamma (1 + \gamma) (\omega^*)^2 \sigma_V^2 \right\} dt \\
 &\quad + (-\gamma) \kappa W^{-\gamma} \omega^* \sigma_V dB^\xi
 \end{aligned} \tag{14}$$

where dB^ξ is a Brownian motion under the new measure, Q^ξ . By applying $J(W) = \kappa \frac{W^{1-\gamma}}{1-\gamma}$, the dynamic process of the investor's marginal utility function can be shown to be:

$$\begin{aligned}
 \frac{dJ_w}{J_w} &= \left\{ (-\gamma) \left[r + \omega^* (\mu - r) + \omega^* \sigma_V^2 \delta^* - \frac{C^*}{W} \right] + \frac{1}{2} \gamma (1 + \gamma) (\omega^*)^2 \sigma_V^2 \right\} dt \\
 &\quad - \gamma \omega^* \sigma_V dB^\xi
 \end{aligned} \tag{15}$$

By substituting the optimal values of ω^* , δ^* , and C^* we can derive the dynamic process of the investor's stochastic discount factor under ambiguity as:

$$\frac{dJ_w}{J_w} = \left\{ -r - \frac{(\mu - r)^2 \phi (1 - \gamma)}{2(1 + \phi)^2 \gamma \sigma_V^2} \right\} dt - \frac{(\mu - r)}{(1 + \frac{1}{\phi}) \sigma_V} dB^\xi \equiv -r^* dt - \frac{(\mu - r)}{(1 + \frac{1}{\phi}) \sigma_V} dB^\xi \tag{16}$$

where $r^* \equiv r + \frac{(\mu - r)^2 \phi (1 - \gamma)}{2(1 + \phi)^2 \gamma \sigma_V^2}$.

We then derive the equity value, $E(V,t)$, under ambiguity. Applying the martingale approach, we have:

$$E^{Q^\xi} [d(J_w E(V,t))] = 0 \tag{17}$$

that is,

$$J_w E^{Q^\xi} \left[E_V dV + E_t dt + \frac{1}{2} E_{VV} (dV)^2 + E \frac{dJ_w}{J_w} + \frac{dJ_w}{J_w} E_V dV \right] = 0 \tag{18}$$

By using Equations (16) and (18), this can be rearranged as:

$$J_w \left[E_V V (v - q) + E_t + \frac{1}{2} E_{VV} V^2 \sigma_V^2 - Er^* - E_V V \frac{(\mu - r)}{1 + \frac{1}{\phi}} \right] dt = 0 \tag{19}$$

The partial differential equation is given as:

$$\frac{1}{2} E_{VV} V^2 \sigma_V^2 + E_V V (r^* - q^*) - r^* E + E_t = 0 \tag{20}$$

where $r^* = r + \frac{(\mu - r)^2 \phi(1 - \gamma)}{2(1 + \phi)^2 \gamma \sigma_V^2}$ and $q^* = q + \left(\hat{r} - \frac{r}{1 + \frac{1}{\phi}} \right) - \frac{\mu}{1 + \phi}$.

Equation (20) can be treated as a typical Black–Scholes partial differential equation with an adjusted risk-free interest rate, r^* , and an adjusted dividend yield, q^* . Therefore, the equity value under ambiguity can be shown immediately as:

$$\hat{E} = V e^{-q^* T} N(d_1^*) - F e^{-r^* T} N(d_2^*) \tag{21}$$

where F denotes the face value of the firm’s debt, and $r^* = r + \frac{(\mu - r)^2 \phi(1 - \gamma)}{2(1 + \phi)^2 \gamma \sigma_V^2}$,

$$q^* = q + \left(r^* - \frac{r}{1 + \frac{1}{\phi}} \right) - \frac{\mu}{1 + \phi}, d_1^* = \frac{\ln\left(\frac{V}{F}\right) + (r^* - q^* + \frac{\sigma_V^2}{2})T}{\sigma_V \sqrt{T}} \text{ and } d_2^* = d_1^* - \sigma_V \sqrt{T}.$$

2.3. Default Probability under Ambiguity

From Equation (21), the default probability under ambiguity is $\pi^* = N(-d_2^*)$, where

$$d_2^* = \frac{\ln\left(\frac{V}{F}\right) + m^* T}{\sqrt{T}}, \text{ and } m^* = \frac{r^* - q^* - \frac{\sigma_V^2}{2}}{\sigma_V}.$$

We can see that the default probability of the fixed income,

debt, is modified due to investors’ ambiguity aversion. Similarly, the actual probability that a firm will

default is computed as $\pi = N(-d_2)$, where $d_2 = \frac{\ln\left(\frac{V}{F}\right) + m T}{\sqrt{T}}$, and $m = \frac{\mu - q - \frac{\sigma_V^2}{2}}{\sigma_V}$.

The information of the risk-neutral default probability is critical for pricing newly issued debt or other credit risk derivatives. However, it is not a trivial matter to obtain the risk-neutral default probability when debt is newly issued and only limited market data are available. It is suggested in the literature that the risk-neutral default rate can be implied from the actual default probability by using the mapping: $\hat{\pi} = N(-\hat{d}_2) = N\left[N^{-1}(\pi) + \eta \sqrt{T}\right]$, where $\hat{\pi}$ is the risk-neutral default probability, and $\pi = N(-d_2)$ is the actual default probability (refer to [28–30] for more information of survival analysis). As the default probability under ambiguity is difficult to explore, we applied this approach to imply the default probability under ambiguity from the actual default probability, as follows. The mapping between the two default probabilities is:

$$\pi^* = N(-d_2^*) = N\left[N^{-1}(\pi) + \eta^* \sqrt{T}\right] \tag{22}$$

By setting $N^{-1}(\pi^*) = N^{-1}(\pi) + \eta^* \sqrt{T}$, we find that $\eta^* = \left(\frac{\mu - r}{\sigma}\right) \left(\frac{\phi}{1 + \phi}\right)$, where ϕ is the penalty parameter, used to depict the investor’s confidence about the reference model.

As mentioned before, investors’ ambiguity aversion is a subjective parameter, so one cannot easily observe its aggregate value or assign it a reasonable value. Following the literature, we use CCI in our empirical study as a proxy for the penalty parameter, that is:

$$\eta^{CCI} = \left(\frac{\mu - r}{\sigma} \right) \left(\frac{CCI}{1 + CCI} \right) \tag{23}$$

By treating the naive default probability as the actual default probability and using CCI as the value of the penalty parameter, the default probability under ambiguity can be constructed as follows:

$$\pi^{CCI} = N \left[N^{-1}(\pi^{naive}) + \eta^{CCI} \sqrt{T} \right] \tag{24}$$

2.4. Cox Proportional Hazard Model

Hazard models are widely used regression models in a range of survival analyses. Survival analysis focuses on the time period it takes for a specific event to occur. Most importantly, in the literature, such survival models attempt to examine the correlation of survival, or death, with its predictors (called covariates). By definition, the hazard rate is the instantaneous risk of demise at time t . It is also known as the conditional default probability, for example:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr \left[(t \leq T < t + \Delta t) \mid T \geq t \right]}{\Delta t} \tag{25}$$

where T is referred to as a random survival time. In order to examine the relationship between survival and its covariates, the Cox proportional hazard model assigns a linear-form for the log hazard, as follows (see [28–30] for further information about survival analysis):

$$\log \lambda_i(t) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \tag{26}$$

where i is used to identify different individuals, and x ’s are the covariates. $\alpha(t) = \log \lambda_0(t)$ is referred to as the log-baseline hazard, as it is the value of the log-hazard when all covariates have no effect and are common to all individuals. This model is also useful for an individual-to-individual comparison. The hazard ratio between individual i and j can be defined as:

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t) e^{\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}}}{\lambda_0(t) e^{\beta_1 x_{j1} + \beta_2 x_{j2} + \dots + \beta_k x_{jk}}} = \frac{e^{\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}}}{e^{\beta_1 x_{j1} + \beta_2 x_{j2} + \dots + \beta_k x_{jk}}} \tag{27}$$

It should be noted that the covariates may vary over time, and the log-hazard with time-dependent covariates is modified as:

$$\log \lambda_i(t) = \alpha(t) + \beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \dots + \beta_k x_{ik}(t) \tag{28}$$

In order to compare the predictability of π^{naive} with π^{CCI} , we use a Cox proportional hazard model in our analysis. We want to show that there may be a possible variable other than π^{naive} in a hazard model that is a statistically significant covariate, so that π^{naive} is not sufficient for forecasting default. In addition, we have to find evidence to show that π^{CCI} performs better in a hazard model than π^{naive} . Such findings would support our suggestion that π^{CCI} should be taken into account in forecasting default. As both Campbell, Hilscher and Szilagyi [2] and Bharath and Shumway [3] have emphasized

that the functional form of the Merton DD model is useful, we also attempt to show that either π^{naive} or π^{CCI} should retain their statistical significance in a hazard model together with all the variables used to calculate these quantities.

2.5. Credit Default Swap Spread Regression

We also use a market-based default probability variable as a robust test to support the results from the hazard model. A credit default swap (CDS) is a contract to protect bond holders if a default situation occurs. The higher is the probability of default, the higher are the insurance fees payable. Hence, the CDS spread can be considered as a measure of default probability. In this paper, we run regressions of log-CDS spread on log- π^{naive} or log- π^{CCI} to highlight that π^{CCI} has superior default predictive power.

3. Data

The sample is from 2000 to 2011, giving 144 monthly observations, and December, 2007, to 2011 in terms of the CDS spread regression to match the CDS contract. The sample of observed firms was based on an article in *CFO Magazine* that was published in November, 2003, “Ranking America’s top debt issuers by Moody’s KMV Expected Default Frequency.” We choose 52 firms with complete data as the non-default sample and added 5 firms entering bankruptcy between 2000 and 2011.

Table 1. Summary statistics.

Panel A. Variables in the Merton distance to default (DD) model and hazard models.

Variable	Quantiles						
	Mean	SD	Min	0.25	Median	0.75	Max
E	59,213.18	74,361.40	0.45	13,575.92	28,167.40	74,009.42	519,044.42
F	41,562.61	57,808.95	684.97	17,509.28	23,839.00	38,574.00	761,203.00
r (%)	2.29	2.00	-0.02	0.21	1.72	4.04	6.39
CCI	86.94	28.22	25.30	61.40	92.95	106.13	144.70
r spread (%)	0.77	10.45	-69.01	-4.58	0.80	6.48	89.40
1/σ _E	1.91	0.44	0.52	1.58	1.92	2.24	3.13
naive σ _v (%)	37.00	10.03	17.92	29.84	35.38	42.43	76.44
π ^{naive} (%)	6.43	9.53	0.00	0.23	2.58	8.20	54.21
π ^{CCI} (%)	6.59	9.36	0.00	0.33	2.81	8.59	53.56

Panel B. Correlation.

$$\text{Corr}(\pi^{naive}, \pi^{CCI}) = 0.994$$

Note: The observation period is from 2000 to 2011. There are 52 non-default firms with 7,488 firm-month observations and 5 default firms with 418 firm-month observations. E and F are the market equity and book debt in millions of dollars, respectively. r is the 3-month risk free rate, and CCI is the Consumer Confidence Index report from the U.S. Conference Board. In addition, 1/σ_E is the inverse of equity volatility, and r spread is the difference between stock return over the last day of the previous month and the risk-free rate over the last day of the current month. Naive σ_v is the volatility of firms’ assets calculated by the naive model approach; π^{naive} is calculated via the naive model, and π^{CCI} is the default probability with CCI.

We obtained much of the data from the Bloomberg Database, including the 3-month risk-free rate, daily stock prices and shares outstanding. We then calculated the value of firms, the volatility of equity, and stock returns of firms in the USA. Following Bharath and Shumway [3], we assume that the debts in firms have a one-year horizon. The value of the firms' debts is obtained from the data of the book value of firms' total liabilities in Bloomberg. The information of CCI is also available in Bloomberg, which is monthly data, for 7,906 firm-months' of data, and the data of the firm's CDS spread were derived from the Datastream database. We adopt a 5-year CDS contract with a base date starting from December, 2007, containing 2,107 firm-months' of data.

Table 1 shows the descriptive statistics of the variables used in the paper. The information about the variables used in the Merton DD, naive DD and hazard models can be found in Panel A. A simple comparison of π^{naive} and π^{CCI} is also given in Panel A. Broadly speaking, π^{CCI} has a higher mean and lower standard deviation than does π^{naive} . From Panel B, we can see that the correlation of π^{naive} and π^{CCI} is extremely high. Therefore, we can infer that π^{naive} and π^{CCI} have a high gearing effect.

4. Empirical Results

4.1. Hazard Model Results

Table 2 shows the results of the Cox proportional hazard models. Two univariate models and four multivariate models are used. Models 1 and 2 are simple univariate models with dependent variables π^{naive} and π^{CCI} individually, and two variables are included in Model 3. It is worth noting that π^{naive} and π^{CCI} are useful for measuring default probabilities, as both are significant in Models 1 and 2, and the coefficients of the two models are negative. In Model 3, although two variables are significant, the sign of π^{naive} is changed due to the influence of π^{CCI} . Therefore, we can infer that π^{CCI} is a better explainer than π^{naive} .

Model 4 is a simple reduced form, which contains the same inputs as the naive model and the naive model with uncertainty aversion. Variable CCI is significantly positive, while r spread is not. For this reason, CCI is useful to explain default probability.

Model 5 adds π^{CCI} to Model 4. As π^{CCI} is significant, we infer that the default probability under ambiguity is a useful estimator. In addition, even though all the inputs used to calculate π^{CCI} are contained in the regression, π^{CCI} is still significant. This shows that the functional form of the Merton DD model is quite important. In Model 6, which contains all the variables, both π^{naive} and π^{CCI} are significant. However, the sign of π^{naive} is changed owing to the effect of π^{CCI} , under the same conditions as Model 3, suggesting that π^{CCI} is superior to π^{naive} . As all the variables used to calculate π^{naive} or π^{CCI} are significant and of the correct sign in Model 6, we confirm robustly that the functional form of the Merton DD model is useful when predicting defaults.

In summary, our findings correspond to those of Campbell, Hilscher and Szilagyi [2] and Bharath and Shumway [3], in that the functional form of the Merton DD model is important. Our evidence shows that either π^{naive} or π^{CCI} is useful in forecasting default. However, π^{CCI} performs slightly better than does π^{naive} , and this suggests that CCI provides crucial data in default forecasting. Hence, our new default prediction model, which treats investors' ambiguity aversion as additional information, should be a useful guide for examining default probability.

Table 2. Hazard model estimates.

Dependent Variable: Time to default						
Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
π^{naive}	-0.183 * (0.018)		0.733 * (0.062)			0.504 * (0.077)
π^{CCI}		-21.065 * (1.981)	-97.848 * (7.726)		-22.196 * (1.879)	-76.093 * (9.071)
CCI				0.021 * (0.002)	0.015 * (0.002)	0.013 * (0.002)
ln(E)				-0.758 * (0.035)	-0.721 * (0.035)	-0.698 * (0.035)
ln(F)				0.746 * (0.048)	0.938 * (0.052)	0.920 * (0.051)
$1/\sigma_E$				-2.208 * (0.168)	-2.170 * (0.167)	-2.177 * (0.164)
r spread				0.814 (0.342)	-3.800 * (0.636)	-3.783 * (0.638)

Note: There are 7,906 firm-months in the sample from 2000 to 2011; π^{naive} is calculated via the naive model; π^{CCI} is the default probability with CCI, and ln(E) and ln(F) are the natural logarithms of market equity and book debt, respectively; $1/\sigma_E$ is the inverse of equity volatility; r spread is the difference between stock return over the last day of the previous month and the 3-month risk free rate over the last day of the current month. Standard errors are shown in parentheses. * Denotes significance at the 1% level.

Table 3. CDS spread regressions.

Panel A. Summary statistics.

Variable	Quantiles						
	Mean	SD	Min	0.25	Median	0.75	Max
CDS spread	134.01	405.39	10.00	43.28	61.00	102.50	10,255.00
π^{naive}	5.07	11.86	0.00	0.00	0.03	2.77	68.63
π^{CCI}	5.02	11.79	0.00	0.00	0.02	2.59	68.59

Panel B. Regressions.

Variable	Dependent variable: log(CDS spread)		
	Model 1	Model 2	Model 3
Constant	-1.8494 * (0.0011)	-1.6551 * (0.0169)	-1.3999 * (0.0776)
log(π^{naive})	0.1478 * (0.0050)		-0.2025 * (0.0599)
log(π^{CCI})		0.1735 * (0.0058)	0.4075 * (0.0695)
Obs.	2107	2107	2107
R^2	0.2919	0.2995	0.3033

Note: The observation period is from December, 2007, to December, 2011. There are 43 firms in the sample for a total of 2,107 firm-month observations. CDS spread is the credit default swap spread in basis points; π^{naive} is calculated via the naive model; π^{CCI} is the default probability with CCI. Standard errors are shown in parentheses. * Denotes significance at the 1% level.

4.2. CDS Spread Regressions

We also provide evidence from the credit derivatives market to support our empirical findings. The observation period is from December, 2007, to December, 2011, in order to match the five-year CDS contract, which has a base date starting from December, 2007. We have data on 43 firms.

Panel A in Table 3 shows the descriptive statistics of the variables. The CDS spread is in basic points, while π^{naive} and π^{CCI} are in percentages. We construct three models in this paper, and the details are given in Panel B. In Models 1 and 2, which are simple univariate models, all of the variables are significant, and R^2 for in Model 2 is slightly higher than in Model 1. In Model 3, the statistical significance of π^{naive} is retained, though the sign of the coefficient is changed. However, both the sign and the statistical significance of π^{CCI} are retained. Therefore, π^{CCI} seems to be more suitable for measuring default probability.

5. Conclusions

Campbell, Hilscher and Szilagyi [2] and Bharath and Shumway [3] provided evidence that the default predictability of the Merton DD model can be attributed to its specific functional form. Bharath and Shumway [3] constructed a naive default forecasting model. The feature of their naive model is that it retains the functional form of the Merton DD model and computes default probability in a naive way, rather than simultaneously solving two nonlinear equations or implementing a complicated iterative procedure.

Recently, the issue of ambiguity aversion has drawn a lot of attention in the academic field. Although there have been many studies dealing with ambiguity issues in asset prices and equity premia, little is available regarding the influence of ambiguity aversion on forecasting default probability. The purpose of this paper was to construct a default prediction model under ambiguity and to examine the validation of the Merton DD model under ambiguity. To the best of our knowledge, ambiguity aversion has not previously been analyzed in terms of the default prediction issue considered here.

By using CCI in our empirical analysis as an inverse proxy for the level of ambiguity aversion, we constructed a new default forecasting model. Applying the Cox proportional hazard model, this paper showed that the new model was superior to Bharath and Shumway's naive model. In addition, although the statistical significance of Bharath and Shumway's naive default probability is retained, its sign is changed due to the effect of our model in the CDS spread regressions. Therefore, our model could serve as a useful reference for modifying the Merton DD model when ambiguity aversion is accommodated.

Acknowledgments

The authors are very grateful to the Editor-in-Chief, Michael McAleer, and two referees for their very helpful comments and suggestions that improved the quality of our paper.

Author Contributions

Both authors contributed to all aspects of this work.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Merton, R.C. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *J. Financ.* **1974**, *29*, 449–470.
2. Campbell, J.Y.; Hilscher, J.; Szilagyi, J. In Search of Distress Risk. *J. Financ.* **2008**, *63*, 2899–2939.
3. Bharath, S.; Shumay, T. Forecasting Default with the Merton Distance to Default Model. *Rev. Financ. Stud.* **2008**, *21*, 1339–1369.
4. Crosbie, P.; Bohn, J. Modeling Default Risk. 2003. Available online: <http://business.illinois.edu/gpennacc/MoodysKMV.pdf> (accessed 20 March 2014).
5. Vassalou, M.; Xing, Y. Default Risk in Equity Returns. *J. Financ.* **2004**, *59*, 831–868.
6. Ellsberg, D. Risk, Ambiguity, and Savage Axioms. *J. Econ.* **1961**, *75*, 643–669.
7. Bewley, T.F. Knightian Decision Theory and Econometric Inferences. *J. Econ. Theor.* **2011**, *146*, 1134–1147.
8. Dow, J.; Werlang, S. R. Uncertainty Aversion, Risk Aversion and the Optimal Choice of Portfolio. *Econometrica*, **1992**, *60*, 197–204.
9. Epstein, L.; Wang, T. Intertemporal Asset Pricing under Knightian Uncertainty. *Econometrica*, **1994**, *62*, 283–322.
10. Chen, Z.; Epstein, L. Ambiguity, Risk and Asset Returns in Continuous Time. *Econometrica*, **2002**, *70*, 1403–1443.
11. Epstein, L.; Miao, J. A Two-Person Dynamic Equilibrium under Ambiguity. *J. Econ. Dyn. Control*, **2003**, *27*, 1253–1288.
12. Gilboa, I.; Schmeidler, D. Maxmin Expected Utility with Non-Unique Prior *J. Math. Econ.*, **1989**, *18*, 141–153.
13. Easley, D.; O’Hara, M. Microstructure and Ambiguity. *J. Financ.* **2010**, *65*, 1817–1846.
14. Bewley, T.F. Knightian Decision Theory: Part I. *Decis. Econ. Financ.* **2002**, *25*, 79–110.
15. Klibanoff, P.; Marinacci, M.; Mukerji, S. A Smooth Model of Decision Making under Uncertainty. *Econometrica* **2005**, *73*, 1849–1892.
16. Maccheroni, F.; Marinacci, M.; Rustichini, A. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica* **2006**, *74*, 1447–1498.
17. Boyle, P.; Uppal, R.; Wang, T. Ambiguity Aversion and the Puzzle of Own-Company Stock in Pension Plans. Working Paper. Institute of Finance and Accounting London: London, UK, 2003.
18. Uppal, R.; Wang, T. Model Misspecification and Underdiversification. *J. Financ.* **2003**, *58*, 2465–2486.
19. Anderson, E.; Hansen, L.; Sargent, T. Robustness, Detection and the Price of Risk. Unpublished paper, Stanford University: Stanford, CA, USA, 2000.
20. Maenhout, P. Robust Portfolio Rules and Asset Pricing. *Rev. Financ. Stud.*, **2004**, *17*, 951–983.

21. Kogan, L.; Wang T. A Simple Theory of Asset Pricing under Model Uncertainty. Unpublished Paper, MIT Sloan School of Management: Cambridge, MA, USA, 2002.
22. Bossaerts, P.; Ghirardato, P.; Guarneschelli, S.; Zame, W. Ambiguity in Asset Markets: Theory and Experiment. *Rev. Financ. Stud.* **2010**, *23*, 1325–1359.
23. Epstein, L.; Schneider, M. Ambiguity and Asset Markets. *Annu. Rev. Financ. Econ. Annu. Rev.* **2010**, *2*, 315–346.
24. Guidolin, M.; Rinaldi, F. Ambiguity in Asset Pricing and Portfolio Choice: A Review of the Literature. *Theor. Decis.* **2013**, *74*, 183–217.
25. So, L. Are Real Options “Real”? Isolating Uncertainty from Risk in Real Options Analysis. *Ann. Financ. Econ.* **2014**, forthcoming.
26. Ait-Sahalia, Y.; Brandt, M.W. Variable Selection for Portfolio Choice. *J. Financ.* **2001**, *56*, 1297–1351.
27. Buraschi, A.; Jiltsov, A. Model Uncertainty and Option Markets with Heterogeneous Beliefs. *J. Financ.* **2007**, *61*, 2841–2897.
28. Cox, D.R.; Oakes, D. *Analysis of Survival Data*; Chapman and Hall: New York, NY, USA, 1984.
29. Allison, P.D. *Survival Analysis Using the SAS System: A Practical Guide*; SAS Publishing: Cary, NC, USA, 1995.
30. Fox, J. Cox Proportional-hazard Regression for Survival Data. Appendix to An R and S-PLUS Companion to Applied Regression. 2002. Available online: <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-cox-regression.pdf> (accessed on 1 February 2014).

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).