IMES DISCUSSION PAPER SERIES

A Framework for Extracting the Probability of Default from Stock Option Prices

Azusa Takeyama, Nick Constantinou, and Dmitri Vinogradov

Discussion Paper No. 2012-E-14

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

2-1-1 NIHONBASHI-HONGOKUCHO
CHUO-KU, TOKYO 103-8660
JAPAN

You can download this and other papers at the IMES Web site:
http://www.imes.boj.or.jp

IMES Discussion Paper Series 2012-E-14
October 2012

A Framework for Extracting the Probability of Default from Stock Option Prices

Azusa Takeyama*, Nick Constantinou, and Dmitri Vinogradov*****

Abstract

This paper develops a framework to estimate the probability of default (PD) implied in listed stock options. The underlying option pricing model measures PD as the intensity of a jump diffusion process, in which the underlying stock price jumps to zero at default. We adopt a two-stage calibration algorithm to obtain the precise estimator of PD. In the calibration procedure, we improve the fitness of the option pricing model via the implementation of the time inhomogeneous term structure model in the option pricing model. Since the term structure model perfectly fits the actual term structure, we resolve the estimation bias caused by the poor fitness of the time homogeneous term structure model. It is demonstrated that the PD estimator from listed stock options can provide meaningful insights on the pricing of credit derivatives like credit default swap.

Keywords: probability of default (PD); option pricing under credit risk; perturbation method

JEL classification: C12, C53, G13

***** Deputy Director and Economist, Institute for Monetary and Economic Studies, Bank of Japan (E-mail: azusa.takeyama@boj.or.jp)
****** Lectuer, Essex Business School, University of Essex (E-mail: nconst@essex.ac.uk)
******* Lectuer, Essex Business School, University of Essex (E-mail:dvinog@essex.ac.uk)

This paper is one chapter of Azusa Takeyama's PhD thesis submitted to the University of Essex. We are grateful to Emmanuel Haven, Xiaoquan Liu, Edward Tsang, John Ohara and participants on 2010 CCFEA PhD workshop and 32th JAFEE meeting. Views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan.

1 Introduction

In this paper we develop a framework to estimate the probability of default (PD) from listed stock option prices. PD plays prominent roles both in academic and practical issues of finance. Firstly, PD is one of the core elements of pricing models for defaultable financial instruments (bonds, CDS, CDS options etc.). Secondly, PD (and the co-movement of PDs) provides useful information both for investors and regulators when assessing the soundness of financial systems and individual institutions (see, e.g.; Clare, 1995 and Tabak *et al.*, 2011, for a relevant discussion).

While it is possible to calculate the arbitrage price of debt instruments such as corporate debts (Duffie and Singleton, 1999) and CDS (Duffie, 1999) from PD and loss given default (LGD), it is usually difficult to uniquely specify the parameters from the market price: observed prices are approximately given by the product of PD and LGD, and thus a continuum of parameters is consistent with a unique market price. As a practical simplification, rating agencies and financial institutions evaluated the PD of corporate debts with a 60% LGD assumption to rate and price securitized products up to 2008 (Moody's, 2001 and Krekel, 2010) although empirical research does not support the assumption (Altman *et al.*, 2005, Qi and Zhao, 2011, Das and Hanouna, 2009 and Takeyama *et al.*, 2010). Most practitioners only consider the homogeneous movement of LGD (Krekel, 2010 and Amraoui *et al.*, 2009) after the constant recovery assumption was found to be too simplistic in the turbulence of the credit market due to the failure of Bear Sterns in March 2008. It is therefore of both practical and theoretical importance to establish a reliable approach to specify the PD of individual entities.

One solution to the above problem is a joint specification of the PD with other asset prices of the same name: if the LGD is known at least for one of the assets then it is possible to specify the PD. The existing approaches therefore differ in the choice of those same-name assets. Schläfer and Uhrig-Homburg (2009) and Norden and Weber (2010) calculate PD using different seniority CDS since a positive spread on senior CDS logically implies a 100% LGD in the subordinate CDS under the absolute priority rule. Jarrow (2001) and Das and Hanouna (2009) obtain the PD from stock prices by modeling the stock as a defaultable security with full loss ($LGD = 100\%$). This idea can be extended to incorporate equity options. Linetsky (2006) and Carr and Wu (2010) suggest the joint pricing model of stock options with debt products.

We adopt the stock option approach for two important reasons. First, listed stock options provide a rich source of data. Estimating PD from the different seniority CDS is only applicable to companies that issue different seniority debts. The number of companies whose stock options are traded is larger than that of companies which issue different seniority debt.[1] On the other hand the PD estimation from the stock price requires some range of time series data of the stock price as it is univariate time series data. Therefore, short term changes in the PD are difficult to capture when the stock price data is the only available data. In contrast to the stock price data, stock options provide a large cross section of data. Therefore, even daily changes of PD can be captured for a large number of companies. The ability to monitor credit risk daily is an important aspect of risk management following the subprime crisis. The second property is that listed stock options are basically free from counterparty risk. Since CDS are over-the-counter (OTC) securities, their prices depend on counterparty risk as well as the credit risk of the reference entity. On the other hand, the daily mark to market clearing of the exposure in the clearing house minimizes the counterparty risk of the listed equity options. These two properties - richness of data and no counterparty risk - underpin our choice of financial instruments to estimate the PD.

An important ingredient in our framework is to employ an option pricing model that is sufficiently flexible (incorporating default risk, stochastic volatility and stochastic term structure) in order to model the volatility surface over a wide range of strikes. A deterioration in credit leads to a significant fall in equity prices resulting in traded options which are drastically in or out of the money. We employ the model of Bayraktar and Yang (2011), which has just these characteristics, to estimate the PD implied in listed stock options although Bayraktar and Yang (2011) calibrate the volatility surface of stock options using the PD inferred from corporate bonds. The main contribution of our paper is to construct a framework to extract the PD from listed options using the Bayraktar and Yang (2011)

[1] For example, subordinate CDS are only available in financial institutions within the members of CDX NA IG, US corporate CDS index, while equity options are available in all the listed stocks of the members.

model which we demonstrate needs to be modified with an appropriate description of the term structure that goes beyond Vasicek (1977) in order to capture the significant twist experienced during the subprime crisis in 2008-2009. The time inhomogeneous term structure model of Hull and White (1990) [henceforth HW] by construction can capture any shape of term structure and is an integral ingredient in our framework.

The framework employs a two-stage calibration procedure which is similar to the Cont and Tankov (2004) method. In the first stage of the calibration, the non-credit related parameters are estimated and these are then employed in the second stage to obtain an estimator for the PD. We consider four models within the framework, however we demonstrate that one that provides the best fit and an unbiased estimator of the PD is the case of the Bayraktar and Yang (2011) model modified by HW. Other models considered such as the Bayraktar and Yang (2011) model employing the Vasicek (1977) term structure model, may produce close agreement with the market prices under certain circumstances, but the parameters are biased in this case. An important byproduct of the framework is the ability to employ the extracted PD to derive the implied LGD. As the daily estimator of the implied LGD is positively correlated with PD and more volatile in higher rating companies, the daily estimator of the implied LGD is consistent with observation on liquidity risk of CDS (Brigo *et al.*, 2011) as well as monthly and yearly LGD (Das and Hanouna, 2009; Altman *et al.*, 2005).

For estimation we use the listed stock option data of financial institutions and non-financial companies in US and UK (Citi, Exxon, HSBC and BP) during July 2008 - July 2010. This sample period includes the financial crisis, which started from the Lehman Brother collapse in September 2008, and the oil spill in the gulf of Mexico by BP in May 2010. The data therefore contains severe re-valuation of the financial institutions' credit risk and the twisted term structure of the funding interest rate, as well as the firm specific rise on credit risk (oil spill). This period provides ideal stress-testing conditions for our approach.

The rest of the paper is organized as follows. Section 2 proposes an equity option pricing model under credit risk and derives an approximation formula of the model and related models using a perturbation method. Section 3 proposes a calibration algorithm of the option pricing models. Section 4 discusses why the fitness of our model outperforms that of the other models. Section 5 analyze some properties of the PD and LGD estimators.

2 Option pricing under credit risk

In this section we construct the option pricing model to estimate the PD implied in stock options. While we basically follow the Bayraktar and Yang (2011) model, we replace the Vasicek term structure model with HW model which is more suitable for financial derivatives pricing. Since our option pricing model is embedded with one factor stochastic volatility, default risk and stochastic interest rate, it does not have an analytical solution. Therefore, we derive an approximation of the model using the singular perturbation method.

The perturbation method based approximation allows an approximation of option pricing models without the specification of their exact functional form (Kristensen and Mele, 2011). In fact we derive an approximation formula of option pricing models just assuming that the stochastic differential equations of the option pricing model satisfy regulatory conditions for the existence of solutions. Moreover, we do not have to specify all of the parameters since the coefficients of the approximation formula are expressed as the products of the parameters. These features reduce both the computational burden and model misspecification risk.

2.1 The stochastic processes of the underlying stock price and interest rate

We describe the set up of the equity option pricing model under credit risk. First we introduce a Cox processes (Poisson processes of which intensity, λ_t is time varying)

$$\tilde{N}_t = N_t \left(\int_0^t \lambda_s ds \right) \tag{1}$$

and six correlated Brownian motions with the following correlation structure,

$$E[W_t^{(0)}, W_t^{(i)}] = \rho_i t \ , \ E[W_t^{(i)}, W_t^{(j)}] = \rho_{ij} t \ \ i, j \in \{1, 2, 3, 4, 5\}, \ \ t \geq 0. \tag{2}$$

We denote the time of the default of the reference entity in a CDS contract the stopping time τ,

$$\tilde{N}_t = \begin{cases} 0 & \tau > t \\ 1 & \tau \leq t \end{cases} \tag{3}$$

where the intensity of the Cox process, λ_t, is defined

$$\lambda_t = f(Y_t, Z_t), \tag{4}$$

$$dY_t = \frac{1}{\epsilon}(m - Y_t)dt + \frac{\upsilon\sqrt{2}}{\sqrt{\epsilon}}dW_t^{(2)}, \tag{5}$$

$$dZ_t = \delta c(Z_t)dt + g(Z_t)dW_t^{(3)}, \tag{6}$$

$f(Y_t, Z_t)$, $c(Z_t)$ and $g(Z_t)$ are respectively smooth and bounded functions of Z_t and Y_t. Y_t follows a mean reverting stochastic process characterized by positive constants m, υ and ϵ. Since we assume that $c(Z_t)$ and $g(Z_t)$ satisfy Lipschitz continuity and growth conditions, the diffusion process Z_t has a uniquely strong solution. δ and ϵ respectively control the velocity of stochastic processes (5) and (6).

Under this setting, the defaultable stock price process is defined as

$$d\bar{S}_t = \bar{S}_t \left((r_t - q_t)\, dt + \sigma(\tilde{Y}_t)dW_t^{(0)} - d\left(N_t - \int_0^{t\vee\tau} \lambda_u du \right) \right), \; \bar{S}_0 = x\,, \tag{7}$$

where q_t is the instantaneous dividend rate. The credit event indicator of (7) is the jump diffusion process, where the jump magnitude is fixed at unity.

Secondly, the stochastic volatility function is defined as a smooth and bounded function of $\tilde{Y}_t$. The stochastic process, $\tilde{Y}_t$, follows a mean reverting process,

$$d\tilde{Y}_t = \left(\frac{1}{\epsilon}\left(\tilde{m} - \tilde{Y}_t \right) + \frac{\tilde{\upsilon}\sqrt{2}}{\sqrt{\epsilon}}\Lambda(\tilde{Y}_t) \right) dt + \frac{\tilde{\upsilon}\sqrt{2}}{\sqrt{\epsilon}}dW_t^{(4)}\,. \tag{8}$$

Λ is a smooth, bounded function of one variable which represents the market price of volatility risk. The parameter $1/\epsilon$ is the rate of mean reversion of the process while ϵ also corresponds to the time scale of the process. All the other parameters in equation (5)–(8) are positive valued constants.

While the stock price drops to zero at the moment of the credit event (default) and thereafter stays at zero, the pre-default stock price dynamics has continuous paths.

$$dS_t = S_t \left((r_t + \lambda_t - q_t)\, dt + \sigma(\tilde{Y}_t)dW_t^{(0)} \right), \; S_0 = x \text{ and } \lambda_0 = \lambda\,. \tag{9}$$

Finally the risk free interest rate, r_t is assumed to follow the HW model

$$dr_t = (\alpha_t^{HW} - \beta^{HW} r_t)dt + \eta^{HW}\, dW_t^{(1)}, \quad r_0 = r\,, \tag{10}$$

where

$$\alpha_t^{HW} = \frac{\partial f^M(0,t)}{\partial t} + \beta f^M(0,t) + \frac{(\eta^{HW})^2}{2\beta^{HW}}(1 - \exp(-2\beta^{HW} t)) \tag{11}$$

and $f^M(t,T)$ is the market forward rate from time t to time T.

2.2 Option Pricing under Credit Risk

We define $\mathcal{I}_t$ as the filtration generated by $\tilde{N}_t$. In the framework, it is possible to construct the enlargement filtration $\mathcal{G}_t$ of the filtration $\mathcal{F}_t$ generated by the vector of Brownian motions and credit indicator filtration $\mathcal{I}_t$ ($\mathcal{G}_t = \mathcal{I}_t \vee \mathcal{F}_t$). We can calculate the price of the defaultable contingent claim $P(t,T)$ as the conditional expectation of the pay off, $h(\bar{S}_T)(= P(T,T))$ by Proposition 5.1.1 of Bielecki and Rutkowski (2002):

$$\begin{aligned} P(t,T) &= E\left[\exp\left(-\int_t^T r_s ds\right) h(\bar{S}_T) 1_{(\tau > T)} \middle| \mathcal{G}_t\right] \\ &= 1_{(\tau > t)} E\left[\exp\left(-\int_t^T (r_s + \lambda_s)\, ds\right) h(S_T) \middle| \mathcal{F}_t\right]. \end{aligned} \tag{12}$$

Duffie and Singleton (1999) obtain the equivalent result for defaultable bonds.

Equation (12) provides some useful special cases. First, if $h(\bar{S}_T) \equiv 1$, then $P(t,T)$ is a (non-defaultable) risk free discount bond price, $B^0(t,T)$.

$$B^0(t,T) = E\left[\exp\left(-\int_t^T r_s ds\right) \middle| \mathcal{F}_t\right] \tag{13}$$

In the setting described by (10), it is possible to obtain the explicit solution of (13) as the forward start bond price

$$B^0(0,t,T) = \exp(a^{HW}(t,T) - b^{HW}(t,T) r_t), \tag{14}$$

where

$$\begin{aligned} b^{HW}(t,T) &= \frac{1 - \exp(\beta^{HW}(T-t))}{\beta^{HW}}, \\ a^{HW}(t,T) &= \frac{f^M(0,T)}{f^M(0,t)} \left(b^{HW}(t,T) f^M(0,t) - \frac{(\eta^{HW})^2}{4\beta^{HW}}(1 - \exp(-2\beta^{HW} t)) b^{HW}(t,T)^2 \right). \end{aligned}$$

If $h(\bar{S}_T) \equiv 1_{(\tau>T)} + 1_{(\tau\leq T)}(1-l)P(t,\tau_1-)$ and l is the rate of loss at default, then $P(t,T)$ is a defaultable discount bond price, $B^c(t,T)$.

$$\begin{aligned} B^c(t,T) &= E\left[\exp\left(-\int_t^T r_s ds\right)(1-1_{(\tau\leq T)}l)\middle|\mathcal{G}_t\right] \\ &= E\left[\exp\left(-\int_t^T (r_s + l\lambda_s)\,ds\right)\middle|\mathcal{F}_t\right]. \end{aligned} \tag{15}$$

Finally, if $h(\bar{S}_T) \equiv (S_T - K)^+$, $P(t,T)$ represents the price of a European call option with strike price K $(K > 0)$ which reduces to

$$\begin{aligned} Call(t,T) &= E\left[\exp\left(-\int_t^T (r_s+\lambda_s)\,ds\right)(X_T-K)^+\middle|\mathcal{F}_t\right] \\ &= xN(d_1) - K\,E\left[\exp\left(-\int_t^T (r_s+\lambda_s)\,ds\right)\middle|\mathcal{F}_t\right]N(d_2), \end{aligned} \tag{16}$$

where $N()$ is the standard normal distribution function and

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{x}{KB_0^c(t,T)}\right) + \frac{1}{2}\sigma(t,T)}{\sqrt{\sigma(t,T)}}, d_2 = d_1 - \sqrt{\sigma(t,T)}, \\ \sigma(t,T) &= E\left[\int_t^T \sigma\left(\tilde{Y}_t\right)^2 dt\middle|\mathcal{F}_t\right], \quad B_0^c(t,T) = E\left[\exp\left(-\int_t^T (r_s+\lambda_s)\,ds\right)\middle|\mathcal{F}_t\right]. \end{aligned}$$

The price of the put option is obtained by using put-call parity

$$\begin{aligned} Put(t,T) &= -x\,N(-d_1) + K\,E\left[\exp\left(-\int_t^T (r_s+\lambda_s)\,ds\right)\middle|\mathcal{F}_t\right]N(-d_2) \\ &+ K\,E\left[\exp\left(-\int_t^T r_s\,ds\right) - \exp\left(-\int_t^T (r_s+\lambda_s)ds\right)\middle|\mathcal{F}_t\right]. \end{aligned} \tag{17}$$

2.3 Volatility Surface Modeling Using the Singular Perturbation Method

The implementation of the perturbation method is achieved following a two-step process. First, we formulate an analytically tractable model, P_0, as the simplified modification of the original option pricing model $\tilde{P}_{\epsilon,\delta}$. By Corollary 3.2 of Bayraktar and Yang (2011), we derive the call option price P_0 assuming that the volatility of underlying stock price process and the intensity of the hazard rate are fixed at $\bar{\sigma}_1^2$ and $\bar{\lambda}$ respectively.

$$P_0(t,T) = xN(d_1) - KE\left[\exp\left(-\int_t^T \left(r_s+\bar{\lambda}\right)ds\right)\middle|\mathcal{F}_t\right]N(d_2), \tag{18}$$

where

$$
\begin{aligned}
d_1 &= \frac{\log\left(\frac{x}{K\bar{B}_0^c(t,T)}\right)+\frac{1}{2}\bar{\sigma}(t,T)}{\sqrt{\bar{\sigma}(t,T)}}\,, d_2 = d_1\sqrt{\bar{\sigma}(t,T)}\,,\\
\bar{\sigma}(t,T) &= \bar{\sigma}_1^2(T-t)+(\eta^{HW})^2\int_t^T (b^{HW}(s,T))^2\,ds + 2\eta^{HW}\bar{\rho}\bar{\sigma}_1\int_t^T b^{HW}(s,T)\,ds\,, \quad (19)\\
\bar{B}_0^c(t,t,T) &= E\left[\exp\left(-\int_t^T\left(r_s+\bar{\lambda}\right)ds\right)\middle|\mathcal{F}_t\right]\,,\\
&= \exp\left(a^{HW}(t,T)-b^{HW}(t,T)r_t-\bar{\lambda}(T-t)\right)\,, \quad (20)
\end{aligned}
$$

and $\bar{\rho}$ is the correlation between the stock price process and risk free interest rate. The put option price in the above assumption is given by put-call parity.

In the second step, we calculate the model option price $\tilde{P}_{\epsilon,\delta}$ using the first order asymptotics on the (small) positive constants, $\sqrt{\epsilon}$ and $\sqrt{\delta}$, to approximate the actual option price P,

$$\tilde{P}_{\epsilon,\delta} = P_0 + \sqrt{\epsilon}P_{1,0} + \sqrt{\delta}P_{0,1}\,, \quad (21)$$

where $P_{1,0}$ and $P_{0,1}$ are respectively the first order derivatives of $\tilde{P}_{\epsilon,\delta}$ with respect to $\sqrt{\epsilon}$ and $\sqrt{\delta}$.

Fouque *et al.* (2003) demonstrate the validity of this approximation in their Theorem 3.6. The model price of a stock option, $\tilde{P}_{\epsilon,\delta}$ is then explicitly given by

$$
\begin{aligned}
\tilde{P}_{\epsilon,\delta}(T,K,\bar{\lambda}(z),V) &= P_0(T,K,\bar{\lambda}(z)) + V_1^\epsilon g_1(T,K,\bar{\lambda}(z)) + V_2^\epsilon g_2(T,K,\bar{\lambda}(z))\\
&+ V_3^\epsilon g_3(T,K,\bar{\lambda}(z)) + V_4^\epsilon g_4(T,K,\bar{\lambda}(z)) + V_5^\epsilon g_5(T,K,\bar{\lambda}(z))\\
&+ V_6^\epsilon g_6(T,K,\bar{\lambda}(z)) + V_7^\epsilon g_7(T,K,\bar{\lambda}(z)) + V_1^\delta g_8(T,K,\bar{\lambda}(z))\\
&+ V_2^\delta g_9(T,K,\bar{\lambda}(z))\,. \quad (22)
\end{aligned}
$$

As shown in Appendix A, all coefficients V (=(V_1^ϵ, V_2^ϵ, V_3^ϵ, V_4^ϵ, V_5^ϵ, V_6^ϵ, V_7^ϵ, V_1^δ, V_2^δ)) in (22) are the polynomials of the model parameters defined in (2.1)–(2.7) of Bayraktar and Yang (2011). It follows that it is possible to estimate the option price under credit risk without specification of the individual parameters, but rather the coefficients that arise in the perturbation expansion, V (see Bayraktar and Yang, 2011 for details).

Before we calibrate the model to the actual data, we note some properties of the model, especially P_0. The default intensity $\bar{\lambda}$ plays the same role as the risk free interest rate r_t

in call options, while it partially explains higher option premium in out of the money put option (Figure 4). The following terms of (22) correct P_0 to fit with the volatility skewness with the asymptotic terms g_1 to g_9 which are also the 'Greeks' of P_0 or the partial derivatives of P_0.

2.4 Related option pricing models

The framework of the PD estimation defined in Section 2.1 is equivalent to the model defined in (2.1)–(2.7) of Bayraktar and Yang (2011) except for the term structure model of interest rates. They employ the Vasicek term structure model

$$dr_t = (\alpha^V - \beta^V r_t)dt + \eta^V dW_t^{(1)}, \quad r_0 = r\,, \tag{23}$$

where α^V, β^V and η^V are real valued constants and $W_t^{(1)}$ is a standard Brownian motion. Using (23), $T-t$ term risk free discount bond price is given by

$$B_V^0(t,T) = \exp(a^V(t,T) - b^V(t,T)r_t)\,, \tag{24}$$

where

$$\begin{aligned} b^V(t,T) &= \frac{1-\exp(\beta^V(T-t))}{\beta^V}\,, \\ a^V(t,T) &= \left(\left(\frac{\alpha^V}{\beta^V} - \frac{(\eta^V)^2}{2(\beta^V)^2}\right)(b^V(T-t) - T + t) - \frac{(\eta^V)^2}{4\beta^V} b^V(T-t)^2\right)\,. \end{aligned}$$

While stochastic interest rate models do not contribute to the goodness of fit in equity option pricing models significantly (Bakshi *et al.*, 1997), the selection of the term structure model is critical in the estimation of the PD implied in equity options for two reasons. First the HW model is an analytically tractable term structure model to reproduce the current term structure of interest rates exactly.[2] [3] Thus, the perturbation method based numerical

[2]We can find other time inhomogeneous term structure models in which discount bond prices are analytically solvable, such as CIR++ model (See Brigo and Mercurio, 2006 for details). However, the HW model is the only time inhomogeneous and mean reverting term structure model which provides analytical solutions of both discount bond prices and option prices (Gaussian model, see Musiela and Rutkowski, 2006 for details) as far as we know.

[3]As discussed in Section 5.2 of Brigo and Mercurio (2006), the HW term structure model belongs to the

approximation is applicable to our model, while our model is equivalent to the use of the actual interest rates which is adopted by recent articles on advanced option pricing in terms of the interest rate modeling. Secondly the HW model accounts for the volatility of interest rates. The primary reason of the insignificant contribution of the term structure models in stock option pricing is that the sensitivity of the interest rate volatility change on option prices is generally much smaller than that of the underlying stocks. On the other hand, the short term interbank interest rates volatility reached record high level (Figures 1 - 2). Since Moody's (2010) documents that the historical average of senior debt's LGD is about 20-50%, the average CDS spread in CDS index such as CDX.NA.IG, 1-2% implies 2-8% PD (Figure 6). Therefore, the volatility of the interest rate is too high to neglect in the estimation of the PD implied in equity options.

It is worthwhile recapping the model that has been developed here. It is a stochastic volatility model that incorporates default risk and captures the term structure via HW. This model is referred to as model I. It is an extension of other existing models, denoted II, III and IV, which are now described. The original Bayraktar and Yang (2011) model contains stochastic volatility, credit risk and time homogeneous term structure model described (the Vasicek model). This model is denoted as model II, a special case of our model with α_t^{HW} fixed at constant α^V. In this case, V_7^ϵ is zero and equation (22) becomes identical to equation (4.6) of Bayraktar and Yang (2011). Models III and IV do not contain credit risk but incorporate stochastic volatility and stochastic interest rates. They are the special cases with $\bar{\lambda}_t = 0$ in the underlying stock price process, with the HW (model III) and the Vasicek (model IV) term structure model. Models III and IV test the effect of the term structure model and both are generalizations of Fouque *et al.* (2003), who assume a constant interest rate. In model III, $\bar{\lambda}_t$, V_1^ϵ, V_3^ϵ, V_1^δ and V_2^δ of (22) become zero, and in model IV, V_7^ϵ becomes zero in addition. The four types of pricing models are listed in Table 1 and their non-zero expansion coefficients listed in Table 2.

In the following sections we provide a detailed comparison of our key model I with the

class of forward rate models although the model is originally proposed as one of the (time inhomogeneous) equilibrium model. In fact, when t equals 0, (14) is always equal to the actual discount bond price which is independent of β^{HW} and η^{HW}.

three other models. The overriding requirement is to demonstrate that a stable and consistent PD is obtained from model I. This is demonstrated by performing a fitting comparison against models II, III and IV.

3 Data and Calibration Procedure

3.1 Data

We calibrate an approximation formula of the option pricing model (22) to actual option prices and estimate the PD implied in stock options. The sample contains UK and US financial institutions and non-financial companies; HSBC Holdings (HSBC), Citi Group (Citi), BP Group (BP) and Exxon Mobil corporation (Exxon) during July 2008-July 2010. In this sample period, we can test the efficiency of the models under two different types of market conditions which impacted credit risk; the subprime crisis and the operational crisis that engulfed BP subsequent to the Gulf of Mexico oil spill. Citi faced a severe drop in market valuation during the financial crisis in 2008. BP in contrast, experienced a drastic change of market valuation due to the oil spill in the Gulf of Mexico in April 2010, a period when financial markets were relatively stable. We also investigate HSBC and Exxon which experienced, in contrast, little or no adverse market valuation during this period.

The sources of data we use are listed below:

- Daily stock price data are obtained from Datastream.
- Daily listed stock option price data of the UK companies are obtained from Datastream. The data contains options of all expiration dates and strike prices which were traded on Euroclear over the sample period. Daily listed stock option price data of the US companies traded in CBOE are obtained from OptionMetrics.
- Stock dividends of the sample companies are obtained from each web site of the sample companies. Unless the companies announce that they would cease or reduce the dividend payment[4], the sum of the last four quarterly dividends is assumed to be the dividend in a year.

[4] US Department of the Treasury has decided to restrict the dividend of Citi in the capital injection on

- The interbank interest rates, LIBOR and swap rate are obtained from DataStream .
- Daily 1 year CDS spreads are obtained from Datastream.

We estimate the PD over one year horizon using listed stock option data. However, options with a maturity of exactly one year are not always available as Euroclear and CBOE basically operate a three month expiring cycle. Hence we employ call and put options with reamining maturities from 6 months to 18 months. Therefore, we effectively estimate the average of the PD between 6 months and 18 months.

The listed options are American style and we convert the prices into European style by extracting the early exercise premium following the methodology of Bayraktar and Yang (2011) and Carr and Wu (2010). We calculate the European option prices substituting the implied volatility of American options into the Black-Scholes formula. While the implied volatility of the US companies is available in OptionMetrics dataset, we calculate the implied volatility of the UK companies' options from DataStream using the Bjerksund and Stensland (2002) model.[5] In the dataset from Datastream, we exclude the data which does not satisfy the following arbitrage condition of American type options,

$$\begin{cases} P_t > \max\left(S_t - K, 0\right) & if\, P_t \ is\ the\ price\ of\ a\ call\ option \\ P_t > \max\left(K - S_t, 0\right) & if\, P_t \ is\ the\ price\ of\ a\ put\ option \end{cases}$$

and where OptionMetrics does not provide the implied volatility of these options.

3.2 Calibration Procedure

The objective of the calibration is to obtain an accurate and unbiased estimate of the PD implied in listed stock options. In the calibration of option pricing models, we have two choices to calibrate the equilibrium term structure models of the interest rate. First following Bayraktar and Yang (2011) means the calibration of the equilibrium term structure model to the actual term structure separately and the calibration of the option pricing model to option prices using the parameters of the term structure model. On the other hand Bakshi

14th October 2008 (Table 3). BP announced to stop the dividend payment due to the loss of the oil spill in the Gulf of Mexico on 18th June 2010 (Table 6)

[5] The method is available in the financial derivatives toolbox of Matlab.

et al. (1997) calibrate both the equilibrium term structure model (Cox *et al.*, 1985 model) and option pricing models to S&P index option prices simultaneously. As Merton (1973) shows, the volatility function of option prices (19) depends on the volatility of the risk free interest rate as well as that of the underlying stock price in option pricing with a stochastic interest rate model. It implies that the parameters of the Vasicek model in model II and IV must be calibrated to fit the term structure of interest rates and the volatility function (19). As the best parameters to describe the term structure of interest rates does not necessarily match those that best describe the term structure of interest rate volatilities, the separate calibration with the term structure of interest rates is not always optimal in the option pricing calibration. Moreover, the Vasicek model does not capture the term structure of interest rates well in our sample period (figure 7). Therefore, we follow Bakshi *et al.* (1997) and calibrate simultaneously the term structure model and the option pricing model to stock option prices. In this procedure, we can optimize the term structure model parameters both in terms of the volatility function (19) and the term structure of interest rates. We expect that the fitness of the option pricing model would be better than the result from the separate calibration, while the result may be inconsistent with the term structure of interest rates. On the other hand, as the HW fits the current term structure perfectly, the parameters of the HW term structure model are calibrated just to the volatility function (19). From an optimization point of view, the term structure of interest rate is not a part of the calibration objective but one of the constraints. This, as we demonstrate, is a key advantage of using the HW model in option pricing models.

We calibrate the option pricing model via the optimization of the model parameter vector Θ (V of (22), $\bar{\lambda}_t$ and the term structure model parameter vector Ψ^h) with an algorithm applied to the Merton (1976) model calibration by Cont and Tankov (2004). In this procedure we minimize the weighted sum of absolute error to avoid the bias caused by the difference of the number of options, $N_{i,t}$, at individual maturity i.

1. First, we optimize $V(= (V_1^\epsilon, V_2^\epsilon, V_3^\epsilon, V_4^\epsilon, V_5^\epsilon, V_6^\epsilon, V_7^\epsilon, V_1^\delta, V_2^\delta))$ and Ψ_t^h with $\bar{\lambda}_{t-1}$[6]

$$[\bar{V}\ \bar{\Psi}^h] = \arg\min_{V,\,\Psi^h} \sum_i \sqrt{\sum_j \left(\sigma_{impl}(T_i, K_{i,j}) - \sigma_{impl}^M(T_i, K_{i,j}))\right)^2 / N_{i,t}}\,, \qquad (25)$$

where j refers to the strike of the option and

$$\sigma_{impl}(T_i, K_{i,j}) - \sigma_{impl}^M(T_i, K_{i,j}) \approx \frac{P(t, T_i, K_{i,j}) - \tilde{P}_{\epsilon,\delta}(T_i, K_{i,j}, \bar{\lambda}_{t-1}, V,\ \Psi^h)}{vega(\sigma_{impl}(T_i, K_{i,j}))}\,, \qquad (26)$$

$N_{i,t}$ is the number of the maturity T_i options traded at t, the implied volatility $\sigma_{impl}(T_i, K_{i,j})$ and the option price sensitivity of the volatility $vega(\sigma_{impl}(T_i, K_{i,j}))$ are calculated using the Black-Scholes formula and

$$\Psi^h = \begin{cases} \alpha^V,\ \beta^V\, and\ \eta^V & h = II, IV \\ \beta^{HW}\, and\, \eta^{HW} & h = I, III\ . \end{cases} \qquad (27)$$

2. Second, we estimate PD with the parameters $\bar{V}$ and $\bar{\Psi}^h$ from (25),

$$\bar{\lambda}_t = \arg\min_{\lambda_t} \sum_i \sqrt{\sum_j \left(\sigma_{impl}(T_i, K_{i,j}) - \sigma_{impl}^M(T_i, K_{i,j}))\right)^2 / N_{i,t}}\,, \qquad (28)$$

where

$$\sigma_{impl}(T_i, K_{i,j}) - \sigma_{impl}^M(T_i, K_{i,j}) \approx \frac{P(t, T_i, K_{i,j}) - \tilde{P}_{\epsilon,\delta}(T_i, K_{i,j}, \lambda_t, \bar{V}, \bar{\Psi}^h)}{vega(\sigma_{impl}(T_i, K_{i,j}))}\,. \qquad (29)$$

As model III and IV do not contain credit risk, the calibration terminates at the first step (25). The calibration of models I and II consists of two steps, (25) and (28).

The estimator $\bar{\lambda}_t$ in (28) is not yet equivalent to the one implied in credit instruments for two reasons. The first is the credit risk implied in funding interest rates. The option pricing models in Table 1 assume r_t is a non-defaultable interest rate (a risk free interest rate. See Musiela and Rutkowski, 2006 for details). However, the actual funding rate, r_t obviously contains two kinds of credit risks, the credit risk of the government and that of financial institutions. Thus the estimated value $\bar{\lambda}_t$ in (28) is possibly biased.

[6]The initial value of the probability of default, $\bar{\lambda}_0$, is calculated from 1 year CDS spread under 60% LGD. To eliminate the dependency on the initial value we ignore the first 20 days results.

We adjust the estimation result by accounting for the first type of credit risk, the credit risk of the government. No government is perfectly free from credit risk. The CDS spread of the UK and US government, whose bonds were rated AAA in 2008-2010, was non-zero even before the financial crisis. This implies that $\bar{\lambda}_t$ estimated in (28) can be lower than the true value λ_t if the government bond yield is assumed to be the risk free interest rate. That is the calibration optimizes the level of the sum, $r_t + \bar{\lambda}_t$ regardless of the credit risk embedded in r_t. To modify the bias of the credit risk contained in the funding rate, we estimate the intensity of the Cox process, $\hat{\lambda}_t$ using the estimated value, $\bar{\lambda}_t$ and CDS spread of the UK and US government, $CDS(t, T)$,

$$\hat{\lambda}_t = \bar{\lambda}_t + CDS(t, T) \,. \tag{30}$$

Therefore, even if $\bar{\lambda}_t$ is zero, it implies that the credit risk does not influence the stock option prices explicitly and the implied PD of the company is identical to the CDS spread of the government.

The second adjustment is the risk premium implied in listed stock options. The adjusted estimator $\hat{\lambda}_t$ is the risk neutral intensity of the hazard process. Thus, it is not equivalent to the intensity of the hazard rate in the credit instrument yet. As Berg (2010) demonstrates, $\hat{\lambda}_t$ does not contain a risk premium. Denote the intensity obtained in (30) $\hat{\lambda}_t^{RN}$ to emphasize it under risk neutral one. We follow Berg (2010) and adjust the risk premium with Sharpe ratio (SR_t) to obtain the intensity $\hat{\lambda}_t^{act}$ under the actual probability measure

$$\hat{\lambda}_t^{act} = \Phi\left[\Phi^{-1}\left(\hat{\lambda}_t^{RN}\right) - SR_{t,i}\right] \,, \tag{31}$$

where Φ is a cumulative distribution function of the standard normal distribution.

The problem in the conversion (31) is the calculation of the Sharpe ratio since the expected excess return of risky assets is often negative in short run estimation. Since Fama and French (2002) report the long run Sharpe ratio of US stock market during 1951-2000, SR^M is 0.4, Berg (2010) calculate the Sharpe ratio of individual stocks with the long run Sharpe ratio and the correlation between the individual stock price and the stock index, $Cor_{i,t}$,[7]

$$SR_{i,t} = Cor_{i,t} SR^M \,. \tag{32}$$

[7] The correlation between the individual stock price and stock index is calculated using the 20 days historical data.

We calculate the Sharpe ratio of US and UK firms' stocks using the results of Kyriacou *et al.* (2006) who report that the Sharpe ratio of US and UK equity are 0.36 and 0.24.

4 The analysis of the model fitness

In this section we compare the fitness of the four option pricing models listed in Table 1 to determine which option pricing model is the most appropriate model to capture the volatility surface. The criterion we employ to determine the performance of the models, is the statistical properties of the time series of the weighted sum of absolute error, (28) for model I and II, and (25) for model III and IV.

Before the comparison of the performance of option pricing models, we demonstrate the comparison of the fitness between the term structure model calibration methods (Table 5). As discussed in Section 3.2, we adopt the simultaneous calibration of the term structure model with the other parameters of option pricing models since it is suitable for the option pricing model calibration. In fact, the performance of the simultaneous calibration is better than that of the separate calibration, especially in the calibration of the models with credit risk (model II). In the separate calibration results, the fitness of model IV is better than that of model II in all of the samples although model II is an extended model of model IV. In general the choice of the term structure models does not matter in equity option pricing as the volatility of interest rates is much smaller than that of the underlying stock. However, Table 5 implies that the choice of the term structure models is one of the key issues in the calibration of the option pricing model with credit risk.

However, as discussed later in this section, the better performance of the simultaneous calibration does not necessarily imply more precise estimation of the parameters. The Vasicek model is a kind of the linear projection of the instantaneous interest rate on the term structure of interest rates (equation 24). In other words, the term structure model produces a term structure of interest rates via the calibration although the actual term structure of interest rates is directly observable. As a result, the simultaneous calibration may intentionally choose a term structure which is inconsistent with the actual one when this intentional error in the fitness of the term structure model contributes to the improvement of the option pricing model fitness. This intentional selection of wrong parameters brings about several problems shown later.

The calibration results show that the fitness of the key model, model I is consistently the most appropriate model, with the lowest mean, median and standard deviation among the four models for all the sample companies. The results are tabulated in Table 6, and demonstrate that a time inhomogeneous term structure model and default risk intensity are both key to produce an improved fitness of the option pricing models to option prices using the methodology developed in this paper. On the other hand, we cannot determine the second best model since the performance of the other models depends on the particular situation of the company it is applied too.

The fitness of model I and III is generally better than that of model II and IV respectively in all the sample companies. This means that the time inhomogeneous term structure model (HW model) can improve the fitness of the option pricing models. Moreover, the difference between the median and mean in model II and IV is large while it is small in model I and III. This implies that the option pricing models with equilibrium term structure model (Vasicek model) do not perform well particularly when the term structure of interest rates is twisted as it was, at the height of the subprime crisis September 2008 - March 2009 (Figures 1 and 2). As the discount bond price (24) is a function of the instantaneous interest rate r_t (overnight (1 day) LIBOR rate), the twist of the term structure, positive spreads in 1 day - 6 month and negative spread in 6 month – 1 year, results in the significant pricing error by the equilibrium term structure model.

On the other hand the impact of credit risk on the option pricing model performance is not straightforward. In the results of BP and Exxon, the fitness of model III is better than that of model II. This implies that credit risk is not necessarily a critical risk factor in pricing equity options of these companies. In fact, Figures 5 and 6 show that BP and Exxon did not suffer severe market valuation of their credit, except for BP at the moment of the oil spill in the Gulf of Mexico in April 2010.

Moreover, it is a counter-intuitive result that the fitness of model II is worse than that of model IV in BP, Citi and Exxon. This means that the introduction of default risk intensity $\bar{\lambda}_t$ does not contribute to the improvement of the option pricing model fitness. As we demonstrate above, the fitness of model IV is generally worse than that of model III due to the insufficient fitness of the equilibrium term structure model. In the calibration of model II, we estimate the default intensity $\bar{\lambda}_t$ in (28) after the calibration in (25) which is equivalent to the calibration of model IV. As we calibrate model II with the PD of the previous day, $\bar{\lambda}_{t-1}$, in (25), the poor fitness of model II, especially the term structure model, in the previous day possibly expands the error of the PD estimation of the following day. As a result the fitness of model II is worse than that of model IV although model II is a credit risk adjusted model of model IV.

Finally we note that the combination of the poor fitness of the Vasicek model and the credit risk causes model IV to outperform model III for Citi. As we calibrate the term structure model parameter vector Ψ^V and the vector of coefficients V simultaneously, the calibration result of the term structure model is not necessarily that which best fits the term structure of interest rates, but which best fits to option pricing model. In this case, overfitness of the term structure plays a complementary role of the credit risk in option pricing, since the default intensity has a similar impact on option prices except for out of the money put options (Figure 4). In contrast the fitness of model III is worse than that of model IV as there is no room of the adjustment via the misidentification in model III. Of course model IV cannot outperform model I as the contribution of this lucky misidentification is limited.

In both cases, the goodness of fit of models II – IV is worse than that of model I since the relatively good performance of these models is to some extent the local optimum due to the part of misidentification of the option pricing models. That is, the poor fitness of the term structure model in (25) brings about the misidentification of $\hat{\lambda}_t$ in (28). Although the behavior of the risk free interest rate r_t is similar to that of the default intensity in the option pricing model (4), they are not perfectly identical to each other. Therefore, the term structure model can produce a critical misspecification which the default intensity never cancels out. This implies that the calibration result of models II and IV is the failure of the optimization, the local minimum.

These results demonstrate that the use of the time inhomogeneous term structure model is a necessary condition to estimate PD accurately from equity options. Since Bayraktar and

Yang (2011) estimate the volatility surface using the PD inferred from corporate bonds, the time homogeneous term structure model does not bring about any serious problems. That is, the term structure model does not affect the performance of the option pricing models given the default intensity. In contrast the intentional error in the time homogeneous term structure model can cause some serious misidentification in the estimation of PD from equity options. We divide the calibration procedure into the two steps, (25) and (28) to improve the efficiency of the parameter identification in our reverse modeling of PD. In this division, we have to improve the fitness of the term structure model as the error of the term structure model in (25) can bias the estimation of PD in (28). Therefore, the replacement of the term structure model is not just a minor sophistication of the option pricing model but a crucial improvement in the PD estimation.

5 The analysis of the PD estimators

The results in Section 4 confirmed that model I is the most appropriate model in comparison to the other models listed in Table 1. In this section we demonstrate that the implied PD extracted from model I is a reliable estimator and then discuss the numerical results.

To test whether the implied PD is unbiased, we investigate the property of the vector of the term structure parameters Ψ. As the implied PD is not directly observable, we have no criteria to examine the bias of the implied PD. As an alternative, we investigate the properties of the other parameters, specifically the term structure model parameters Ψ in the option pricing model. If we can demonstrate that all the parameters (except for the implied PD) in model I are unbiased, we can then deduce that the implied PD is unbiased. To do this we impose the theoretical requirement that in the calibration of the model, instruments within the same market obtain statistically equivalent term structure parameters. Table 7 shows that the goodness of fit test (Kolmogorov and Smirnov test) does not reject the null hypothesis "the difference between the parameters follow a normal distribution with zero mean, $N(0, \sigma_i^2)$" only in model I (see Lehmann and Romano, 2005 for details of the test). The implication is that models II, III and IV are possibly misspecified even when the fitness of these models is close to that of model I. We can therefore be confident that the PD obtained from model I is unbiased.

The properties of the PD extracted from model I are now discussed. For comparison we also present the PD extracted from model II, and the PD obtained directly from the CDS spreads. For the latter we use the model for the CDS spread employed by Bluhm *et al.* (2002) with the LGD fixed at 60% (we call this the constant LGD model).

Figures 8 and 9 show the implied PD for UK and US sample companies. It is noted that the PD extracted from model I is less volatile than that obtained via model II in all the sample companies. As discussed in Section 4, the calibration of model II failed to satisfy the goodness of fit criteria due to misspecification of the Vasicek term structure model and the default intensity. Thus the PD estimator gets more biased day by day. On the other hand, the parameters except for the PD estimator are consistent between the sample companies while model I satisfies the goodness of fit criteria better than model II. This implies that the estimator obtained via model I is more reliable than the one obtained via model II.

Before we examine the LGD implied in CDS obtained via model I (Figure 10), it should be noted that there is a conceptual difference of LGD between CDS and bonds. In contrast to the bond default data, the LGD implied in CDS is the expectation of LGD at credit events. As early CDS contracts are designed to settle the payment of CDS via physical settlement,[8] the exchange of the defaulted bond or loan and notional amount of cash after a credit event, the similar movement of the LGD in CDS implies that the market efficiency of bond and CDS. However, ISDA (2009) proposes the transfer of all the existing CDS contracts into the cash settlement (the Big Bang in CDS). Although Howlege *et al.* (2009) demonstrate that the results of the cash settlement in 2007-2008 are close to the bond prices at the date of

[8] There are three types of settlement in the CDS market. In a physical settlement, CDS buyers deliver the debt of the reference entity (typically a loan or bond) and receive the same amount of money as the notional amount of the CDS. In a cash settlement, sellers pay the difference between the market value of the debt and the notional amount of the CDS to buyers. If the market value of the debt is not uniquely available, the dealers of the CDS will determine the market value in the auction. Finally, in a fixed amount settlement, sellers pay a fixed amount of money to buyers regardless of loss due to the credit event. BBA (2006) showed that the market share of CDS contracts which are designed to settle in a cash settlement has increased from 13% to 26% in 2006 with the growth of CDS although the share of the contracts with physical settlement is still 73%.

the auction, the auction results and bond prices are not necesarily close to the final recovery rate.

In fact, the implied LGD implies that CDS spreads reflect non credit risk factors such as counterparty risk and liquidity risk. The fluctuation range of the daily estimator of LGD is consisitent with that of the monthly variation obtained by Das and Hanouna (2009). One of the causes of the relatively high volatility is liquidity risk. Brigo *et al.* (2011) document that the CDS of AAA and AA rated companies reflect liquidity risk more than that of A and BBB rated companies. The implied LGD of Exxon which was rated AAA during 2008-2010 is more volatile than the other sample companies which were rated AA or A (Table 8) although Exxon did not experience systemic risk as did financial institutions, or a firm specific distressed situation as did BP (Tables 3 and 6).

While the implied LGD obtained with the PD of model I reflects non-credit risk factors as well as the expectation of the LGD at credit events, it is consistent with the observation by the previous studies. Previous studies document the positive correlation between PD and LGD (the yearly data of corporate bonds in Altman *et al.*, 2005 and monthly CDS data in Das and Hanouna, 2009). The correlation between PD and LGD is positive in all the sample companies (Table 9). This means that the implied LGD obtained via the PD of model I basically reflects the expectation of the LGD at credit events.

The implied LGD of the financial institutions indicates the impact of the regulation and government intervention. The LGD of the financial institutions decreased after the Lehman Brothers collapse on September 15, 2008 since the implied PD rose quicker than the CDS spreads. This implies that the series of government actions to stabilize the financial system contributed to the support of the financial institutions' credit (Table 3). Debt investors assessed that if banks were short of capital, they could raise capital via the mechanism provided by governments to support banks during the financial crisis. In addition it is important to consider the situation of financial institutions, especially commercial banks, in greater detail, given the greater complexity of their capital structure. Most major banks in the industrialized countries issue various kinds of hybrid securities such as subordinate debts, preferred trust security and preferred stock in addition to senior debt and common stock to meet the regulatory capital requirement. Moreover, they have to keep their risk adjusted capital adequacy ratio and Tier I capital ratio greater than 8% and 4% respectively in the operating regulatory capital requirement during the period considered in this paper. If these ratios are below the required level and the financial institution cannot raise capital immediately, a credit event (usually nationalization) is triggered. Under these specific conditions to banks, the PD of commercial banks is expected to be higher while the LGD of the senior debt of commercial banks remains low. This is indeed what is shown in Figure 10 for HSBC and Citi. Therefore, the implied PD and LGD contradict the constant recovery assumption such as Moody's (2001).

6 Concluding Remarks

We have developed a framework to estimate the PD implied in listed stock options, which consists of a flexible option pricing model and an implementation procedure. Our choice of the HW term structure model allows for the perfect fit to the actual term structure of interest rates. Thus it is possible to calibrate the term structure model only to the term structure of interest rate volatilities instead of calibrating the term structure model to that of the interest rate and interest rate volatility. It reduces the complexity of the calibration, since the term structure of interest rates is one of the calibration constraints rather than a part of the calibration objective from an optimization point of view. Moreover, this implementation succeeds in removing the spiral of increasing error due to the poor performance of the Vasicek model where the actual term structure exhibits twists. The arbitrage term structure model is therefore not just a refinement of the term structure modeling but a crucial component of the unbiased estimation of the PD. Our approach demonstrates several useful properties: (1) the fitness of the option pricing model in this class of models, (2) an unbiased estimate of the PD, (3) relatively smooth PD resulting from *daily* observations. The latter property is both attractive from the perspective of risk management (a highly volatile PD is not very informative about the actual risk of default) and indicative about the ability of markets to properly price the PD.

As a byproduct of the PD estimation framework and as a consistency check we derive the LGD embedded in CDS spreads. Our estimator confirms that the constant 60% LGD assumption is generally invalid both in the analysis of the individual entity's credit risk and the valuation of credit portfolio such as CDOs. The implied LGD of CDS obtained via the PD implied in listed stock option prices is not constant both for the financial institutions and non-financial companies. Moreover, the analysis of the bailed out financial institutions indicates that the PD inferred from the CDS spread with the constant LGD assumption does not plausibly reflect the default risk. It implies that the stochastic recovery modeling with the homogeneous mark down cannot capture the risk of the heterogeneous recovery change in responce to the event of the individual entities.

Finally we note the following two avenues for further research. The first one is the impact of a series of the government actions to stabilize the financial system on the credit risk of the financial institutions. Since the PD is common among the stock and debt of the same name and the stock return is a predictor of the CDS spread in the sense of Granger causality, we might expect that government interventions have different effects on the credit risk of the securities of different seniorities. This implies that it should be possible to test the effect of the policy actions on the PD clean of moral hazard or conflict of stakeholders' interest.

The second issue refers to the impact of heterogeneous LGD on the pricing and rating of CDO. Moody's (2001) claim the robustness of their rating methodology with the constant recovery assumption because there is no significant difference between CDO rating and pricing with the constant recovery assumption and those with heterogeneous recovery rate. Even after the constant recovery CDO pricing models failed to appropriately price CDO tranches, especially super senior tranches in March 2008, practitioners considered the implementation of the homogeneous mark-up of LGD in the stochastic LGD modeling (Krekel, 2010; Amraoui *et al.*, 2009; Amraoui and Hitier, 2008) although they already develop the heterogeneous one. However, the homogeneous mark-up of LGD is not feasible for extremely high PD like that of the financial institutions bailed out by the government. For example the credit risk of financial institutions is accounted in the counterparty risk of the underlying CDS contracts as well as in the credit risk of the underlying contract in pricing synthetic CDO. Our framework allows the implementation of the heterogeneous LGD modeling.

References

[1] Altman, E. I., Brady, B., Resti, A., Sironi,A., 2005. The link between default and recovery rates: theory, empirical evidence, and implications. Journal of Business 78, 2203-2228.

[2] Amraoui, S., Cousot., Hitier, S., Laurent, J., 2009. Pricing CDOs with state dependent stochastic recovery rate, Working Paper.

[3] Bakshi, G., Cao, C., Chen, Z., 1997. Empirical Analysis of alternative option pricing models, Journal of Finance, 52, 2003-2049.

[4] Bayraktar, E., Yang, B., 2011. A unified framework for pricing credit and equity derivatives. Mathematical Finance, 21, 493-517

[5] Berg, T., 2010. From actual to risk-neutral default probabilities: Merton and Beyond. Journal of Credit Risk 6, 55-86.

[6] Bielecki, T. R., Rutkowski, M., 2002. Credit Risk: Modeling, Valuation and Hedging. Springer Verlag: New York.

[7] Bjerksund, P., Stensland, G., 2002. Closed form valuation of American options, working paper NHH Bergen Norway.

[8] Bluhm, C., Overbeck, L., Wagner, C., 2002. An Introduction to Credit Risk Modeling. Chapman & Hall/CRC .

[9] Brigo, D., Mercurio, F., 2006. Interest Rate Models - Theory and Practice Second Edition. Springer Verlag: New York.

[10] Brigo, D., Predescu, M., Capponi, A., 2011. Liquidity modeling for credit default swaps: an overview. in *Bielecki, T., Brigo, D., Patras , F., 2011. Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings, and Liquidity (Bloomberg Financial). Bloomberg Press.*

[11] British Bankers Association., 2006. BBA Credit Derivatives Report 2006.

[12] Carr, P., Wu, L., 2010. Stock options and credit default swaps: A joint framework for valuation and estimation. Journal of Financial Econometrics 8, 409-449.

[13] Clare A., 1995. Using the arbitrage pricing theory to calculate the probability of financial institution failure: Note, Journal of Money, Credit and Banking 27 , pp. 920–926.

[14] Cont, R., Tankov, P., 2004. Non-parametric calibration of jump diffusion option pricing models. Journal of Computational Finance 7, 1-49.

[15] Cox, J., Ingersoll, J., Ross, S., 1985. A theory of the term structure of interest rates, Econometrica 53, 385-408.

[16] Das, S. R., Hanouna, P., 2009. Implied Recovery. Journal of Economic Dynamics and Control 33, 1837-1857.

[17] Duffie, D., 1999. Credit swap valuation. Financial Analyst Journal 55, 73-87.

[18] Duffie, D., Singleton, K., 1999. Modeling term structure of defaultable bonds. Review of Financial Studies 12 , 687-720.

[19] Fama, E, F., French, K, R., 2002. The Equity Risk Premium. Journal of Finance 62, 637-659.

[20] Fouque, J., Papanicolaou, G., Sircar, K., 2000. Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press: New York.

[21] Fouque, J., Papanicolaou, G., Sircar, K., Solna, K., 2003. Multiscale stochastic volatility asymptotics. SIAM Journal of Multiscale Modeling and Simulation 2, 22-42.

[22] Helwege, J., Maurer, S., Sarkar, A., Wang, Y., 2009. Credit default swap auctions and price discovery. Journal of Fixed Income 19, 34-42.

[23] Hull, J., White, A., 1990. Pricing interest rate derivative securities. Review of Financial Studies 3, 573-592.

[24] ISDA (International Swap and Derivatives Association)., 2009. ISDA Announces Successful Implementation of 'Big Bang' CDS Protocol; Determinations Committees and Auction Settlement Changes Take Effect. News Release April 8, 2009.

[25] Jarrow, R., 2001. Default parameter estimation using market prices. Financial Analysts Journal 57, 75-92.

[26] Krekel, M., 2010. Pricing distressed CDOs with Base Correlation and Stochastic Recovery. Risk (July), 90-94.

[27] Kristensen, D., Mele., A. 2011. Adding and subtracting Black-Scholes:A new approach to approximating derivative prices in continuous-time models. Journal of Financial Economics 102, 390-415.

[28] Kyriacos, K., Madsen, Jakob, B., Mase, B., 2006. Does inflation exaggerate the equity premium? Journal of Economic Studies 33, 344 - 356.

[29] Lehmann, E, L., Romano, J, P., 2005. Testing Statistical Hypotheses Third Edition, Springer.

[30] Linetsky, V., 2006. Pricing equity derivatives subject to bankruptcy. Mathematical Finance 16, 255-282.

[31] Merton, R, C., 1973. Theory of Rational Option Pricing. Bell Journal of Economics and Management Science 4, 141-183.

[32] Merton, R, C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125-144.

[33] Moody's., 2001. Responses to Frequently Asked CDO Questions (Second of Series).

[34] Moody's., 2010. Corporate Default and Recovery Rates, 1920-2009.

[35] Musiela, M and Rutkowski, M., 2006. Martingale Methods in Financial Modeling, Springer.

[36] Norden, L., Weber, M., 2010. When senior meets junior in credit default swap spreads of large banks. Working paper.

[37] Qi M., Zhao, X., 2011. Comparison of modeling methods for Loss Given Default, Journal of Banking and Finance 35, 2842-2855.

[38] Schlafer,T.S., Uhrig-Homburg, M., 2009. Estimating market-implied recovery rates from credit default swap premia. Working paper.

[39] Tabak, B.M., Luduvice A.V.D., Cajueiro D.O. , 2011. Modeling default probabilities: The case of Brazil, Journal of International Financial Markets, Institutions and Money 21, 513-534.

[40] Takeyama, A., Constantinou, N., Vinogradov., 2010. Do CDS spreads reflect default risks? Evidence from UK bank bailouts.http://ssrn.com/abstract=1700167

[41] Vasicek, O., 1977. Equilibrium Characterization of the Term Structure. Journal of Financial Economics 5, 177-188.

Table 1: *A summary of the four option pricing models*

	Volatility	**Interest rate**	**Default Risk**
Model I	Stochastic volatility	HW	Default risk
Model II	Stochastic volatility	Vasicek	Default risk
Model III	Stochastic volatility	HW	No default risk
Model IV	Stochastic volatility	Vasicek	No default risk

Table 2: *A summary of the non-zero expansion coefficient in equation (22)*

	V_1^ϵ	V_2^ϵ	V_3^ϵ	V_4^ϵ	V_5^ϵ	V_6^ϵ	V_7^ϵ	V_1^δ	V_2^δ	$\bar{\lambda}(z)$
Model I	•	•	•	•	•	•	•	•	•	•
Model II	•	•	•	•	•	•	-	•	•	•
Model III	-	•	-	•	•	•	•	-	-	-
Model IV	-	•	-	•	•	•	-	-	-	-

Note: • *non-zero and - zero coefficient*

Table 3: *The Major events of the subprime crisis*

Date	Event	Country
09/15/2008	Lehman Brothers Bankruptcy	United States
09/15/2008	Bank of America announces the takeover of Meril Lynch	United States
09/16/2008	FRB set the lending facility for AIG.	United States
09/18/2008	Lloyds announces the takeover of HBOS.	United Kingdom
09/21/2008	Goldman Sachs & Morgan Stanley transfer the status to Bank Holding Company (BHC)	United States
09/29/2008	Bradford&Bingley (B&B) is nationalized.	United Kingdom
09/30/2008	US congress rejected TARP (Troubled Asset Relief Program)	United States
10/03/2008	UK raises the limit of the bank deposit guarantee.	United Kingdom
10/12/2008	Capital injection for RBS, Lloyds and HBOS.	United Kingdom
10/14/2008	Capital injection for the US major banks.	United States
11/23/2008	Capital Injection (2nd) and Asset Guarantee for Citi Group	United States
01/16/2009	Capital Injection (2nd) and Asset Guarantee for Bank of America	United States
01/17/2009	UK announces Asset Protection Scheme (APS).	United Kingdom
02/26/2009	2nd Capital Injection and APS for RBS	United Kingdom
02/27/2009	The Treasury Department announces the conversion of Citi's preferred stock to common stock	United States United States
03/02/2009	HSBC issued common stocks to shareholders	United Kingdom
03/07/2009	UK announces conversion of preferred share to common share and the ASP (Lloyds)	United Kingdom

Table 4: *The major events of the BP oil leak*

Date	Event
04/20/2010	Explosion and fire on the BP-licensed Transocean drilling rig
	the Deepwater Horizon in the Gulf of Mexico
04/22/2010	Deepwater Horizon rig sinks in 5,000ft of water.
	The US National Response team begins search-and-rescue operation.
05/09/2010	BP started to address the undersea leak with a method called
	"junk shot"
05/10/2010	BP announces plans to place a small containment dome, known as
	a "top hat", over the blown-out well to funnel oil to the surface
05/2/6/2010	BP pumps thousands of barrels of mud into the well in an attempt
	to plug the leak. The process, known as top kill, fails to overcome
	the flow of oil.
06/04/2010	S&P cuts BP's credit rating (AA→ AA-)
06/15/2010	Fitch cuts BP's credit rating (AA→ BBB)
	BP agrees to a $20bn downpayment towards compensation
06/17/2010	S&P cut BP's credit rating (AA-→ A)
06/18/2010	Moody's cut BP's credit rating (Aa3→ Baa1)
07/05/2010	BP announced the cost of the oil spill has now risen to over $3bn.
07/13/2010	BP successfully installed a new containment cap on the wellhead.
07/15/2010	BP stopped the flow of oil for the first time in 87 days.
07/26/2010	BP chief executive, Tony Hayward, announces to leave BP.

Table 5: *Descriptive statistics of the pricing errors in model II and IV (%)*

	Model II (sim)			Model II (sep)			Model IV (sim)			Model IV (sep)		
	Mean	Median	S.D	Mean	Median	S.D	Mean	Median	S.D	Mean	Median	S.D
BP	4.708	1.712	6.818	7.281	5.432	8.550	3.940	1.695	6.332	3.094	1.501	5.593
HSBC	1.717	1.084	2.656	7.078	4.827	7.493	2.632	1.381	4.168	2.740	2.197	3.505
Exxon	8.639	5.032	9.075	38.024	18.885	126.981	3.500	2.090	4.928	21.806	3.509	127.095
Citi	5.528	3.409	8.016	20.232	7.698	56.365	4.466	2.103	5.375	17.776	5.388	53.884

Notes: In this table, "sim" and "sep" indicate respectively the simultaneous and separate calibration of the term structure model interest rates. In the separate calibration, the parameters of the term structure model, Ψ *are separately calibrated to the term structure of interest rates in advance of the calibration in (25) while all of the parameters are calibrated simultaneously in the simultaneous calibration.*

.

Table 6: *Descriptive statistics of weighted mean absolute error(%)*

	Model I			Model II			Model III			Model IV		
	Mean	Median	S.D	Mean	Median	S.D	Mean	Median	S.D	Mean	Median	S.D
BP	1.032	0.828	0.818	4.708	1.712	6.818	1.590	1.472	0.758	3.940	1.695	6.332
HSBC	1.146	0.955	0.764	1.717	1.084	2.656	2.243	2.177	1.328	2.632	1.381	4.168
Exxon	1.873	1.465	1.271	8.639	5.032	9.075	2.791	2.378	1.626	3.500	2.090	4.928
Citi	2.931	1.535	4.037	5.528	3.409	8.016	5.834	3.477	6.328	4.466	2.103	5.375

Note: These are determined from equation (28) for model I and II and equation (25) for model III and IV.

Table 7: *Term structure model parameter test*

	Model I		Model II			Model III		Model IV		
	$\Delta\beta$	$\Delta\eta$	$\Delta\alpha$	$\Delta\beta$	$\Delta\eta$	$\Delta\beta$	$\Delta\eta$	$\Delta\alpha$	$\Delta\beta$	$\Delta\eta$
(BP-HSBC)										
Mean	0.060	0.137	-0.048	-0.483	-0.035	0.051	-0.056	-0.023	-0.795	-0.118
p-value	81.69%	13.43%	0.05%	0.00%	39.65%	21.51%	1.50%	6.17%	0.00%	0.45%
KS statistics	0.047	0.054	0.083	0.062	0.063	0.060	0.094	0.071	0.070	0.078
p-value	29.13%	13.01%	0.15%	5.13%	3.20%	3.99%	0.02%	0.98%	1.46%	0.34%
(Exxon-Citi)										
Mean	-0.286	0.379	-0.005	-0.063	-0.059	-0.266	-0.006	-0.140	0.344	-0.150
p-value	47.00%	16.99%	42.60%	13.81%	0.10%	1.37%	42.23%	0.00%	0.07%	0.00%
KS statistics	0.046	0.048	0.090	0.072	0.043	0.389	0.102	0.195	0.367	0.173
p-value	44.53%	35.03%	0.08%	1.36%	33.96%	0.00%	0.00%	0.00%	0.00%	0.00%

Notes: The p-values under the mean value are the superior values to reject the null hypothesis that the sample mean is zero. The Kolmogorov-Smirnov test examines whether both the location parameter (sample mean) shape of the empirical distributions fit normal distribution $N(0, \sigma_i^2)$. *The p-values under the KS statistics are the superior values to reject the null hypothesis that the parameter difference follows a normal distribution. To exclude outliers, we cut off both 0.5% maximum and minimum data.*

Table 8: *Standard deviation of LGD obtained via model I*

	HSBC	BP	Citi	Exxon
Daily	4.84%	2.79%	9.37%	3.74%
Weekly	8.06%	3.01%	13.36%	9.58%

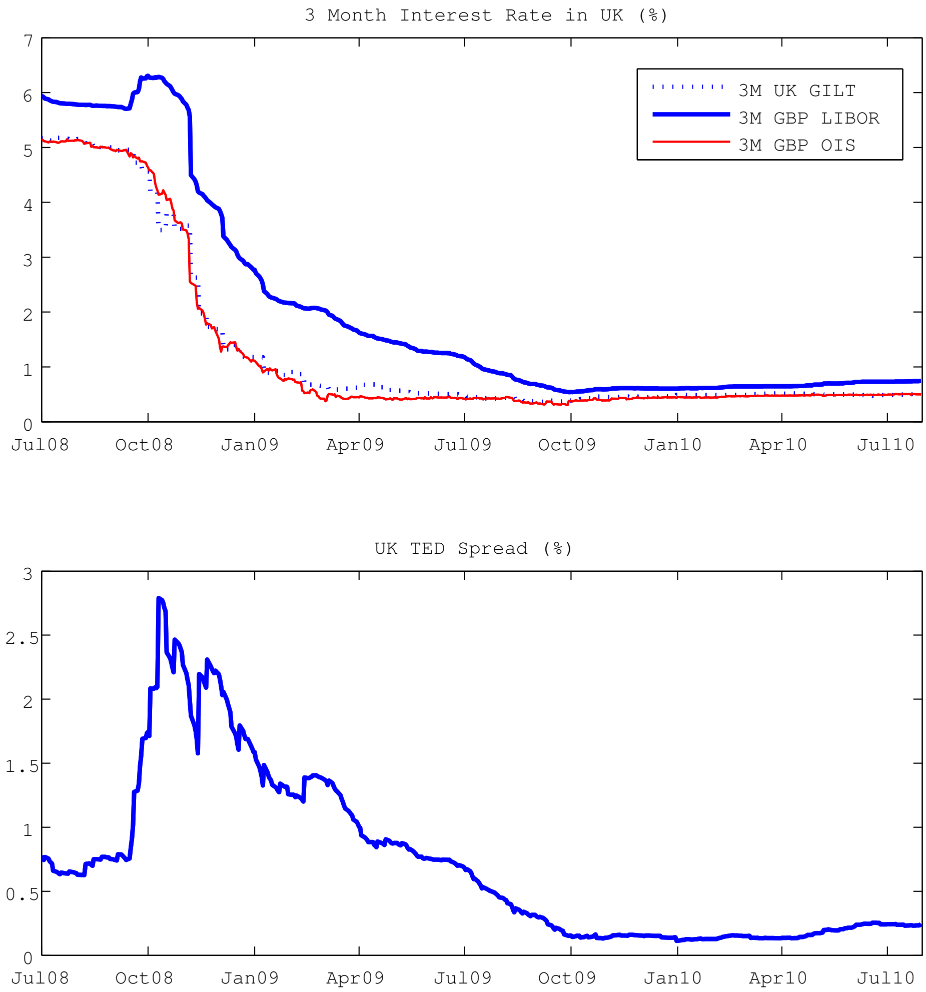
Note: The table shows the standard deviation of the daily and weekly change of the implied LGD in the sample period, July 2008 - July 2010.

Table 9: *Correlation between the implied PD and LGD*

	HSBC	BP	Citi	Exxon
Daily	0.136	0.567	0.190	0.507
Weekly	0.370	0.123	0.224	0.487

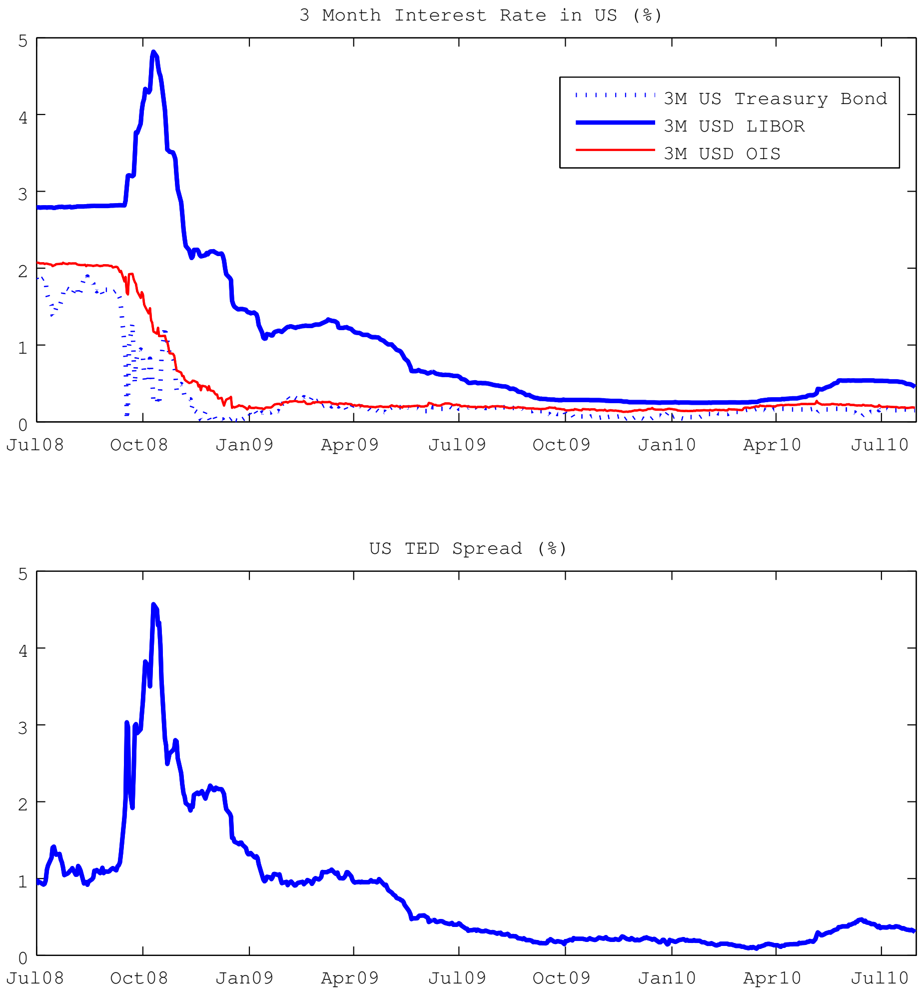
Note: The table shows the correlation coefficient between the daily and weekly change of the implied PD and LGD in the sample period, July 2008 - July 2010.

Figure 1: *Short Term Interest Rates in UK*



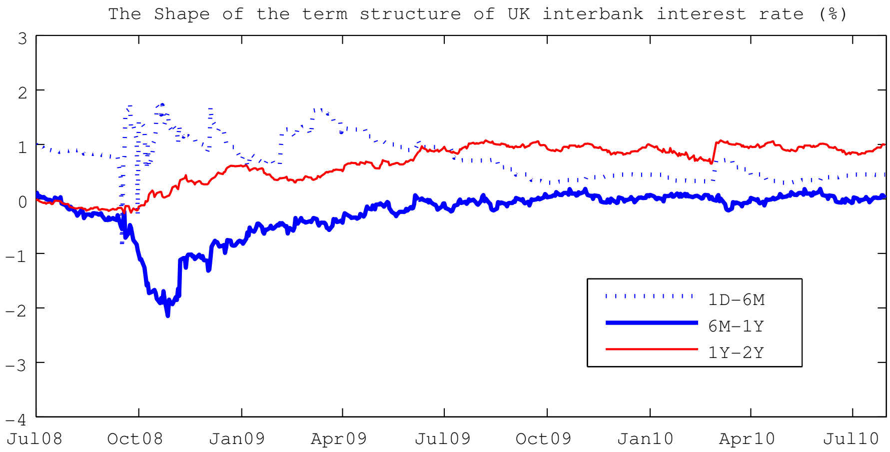
Note: We define the UK TED spread as the spread between 3 month LIBOR rate and 3 month gilt yield.

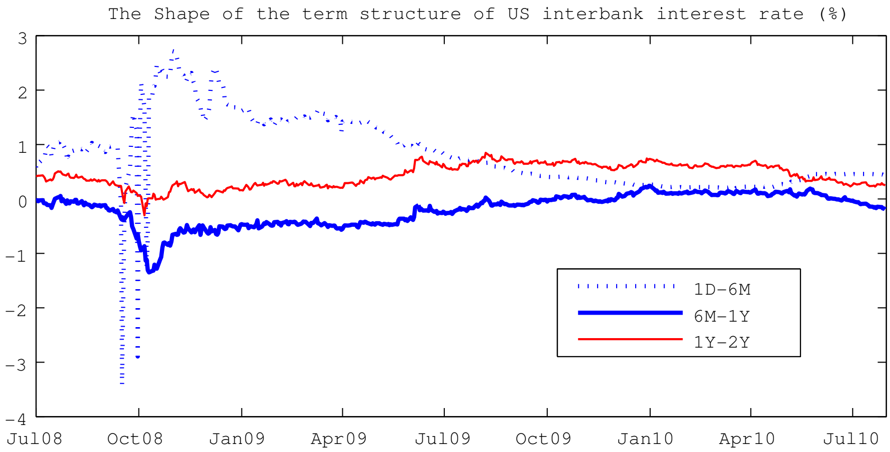
Figure 2: *Short Term Interest Rates in US*



Note: We define the US TED spread as the spread between 3 month LIBOR rate and 3 month treasury yield.

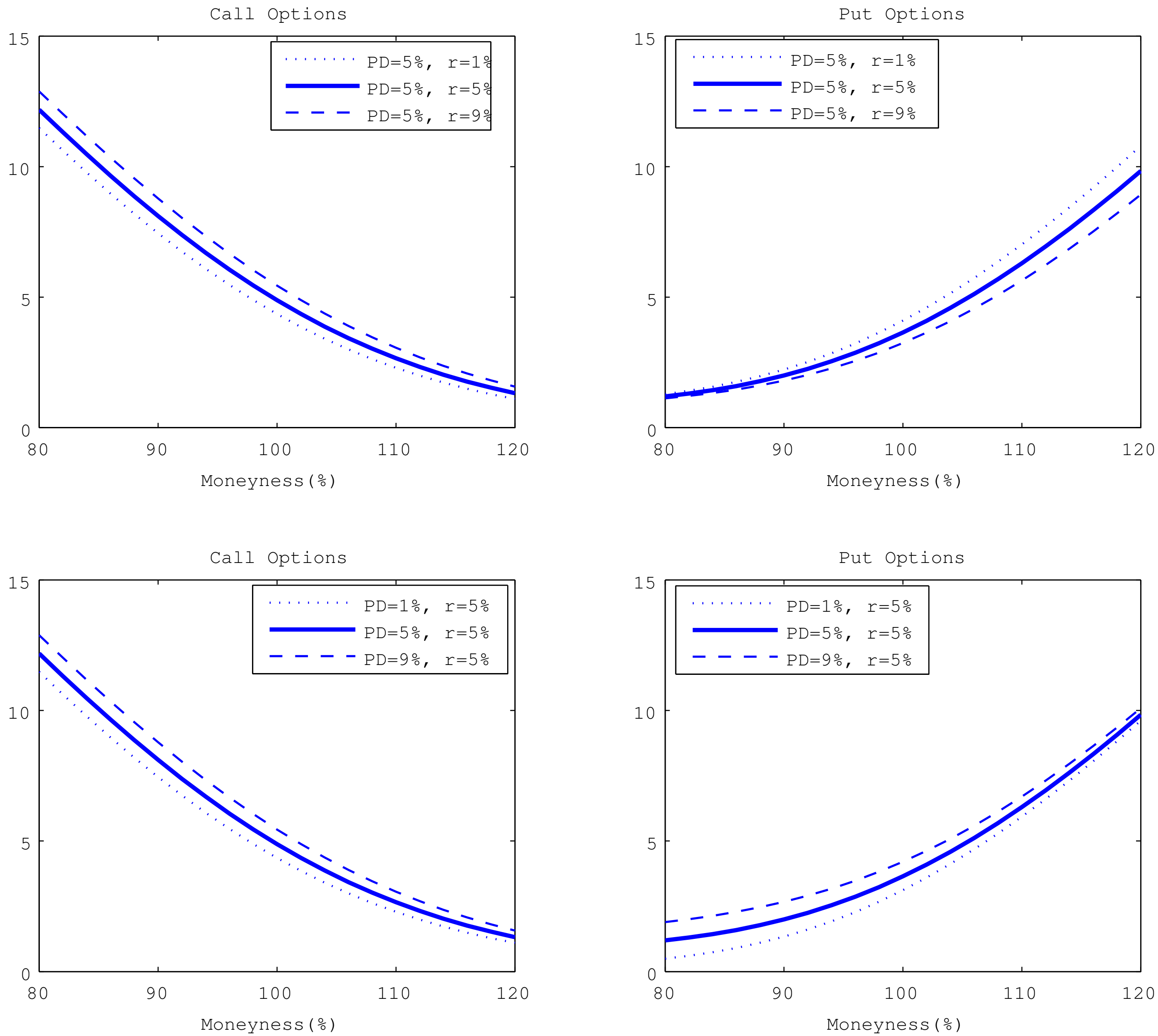
Figure 3: *The shape of term structure*





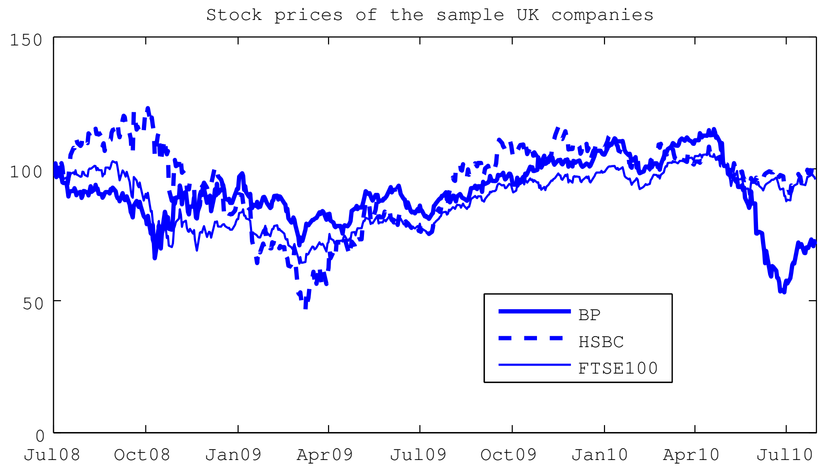
Note: Both UK and US term spreads of the interbank interest rates indicate the inversion in the term structure of interest rates around 6 months. Time homogeneous term structure models such as Vasicek (1977) often fail to reproduce the term structure due to the humps.

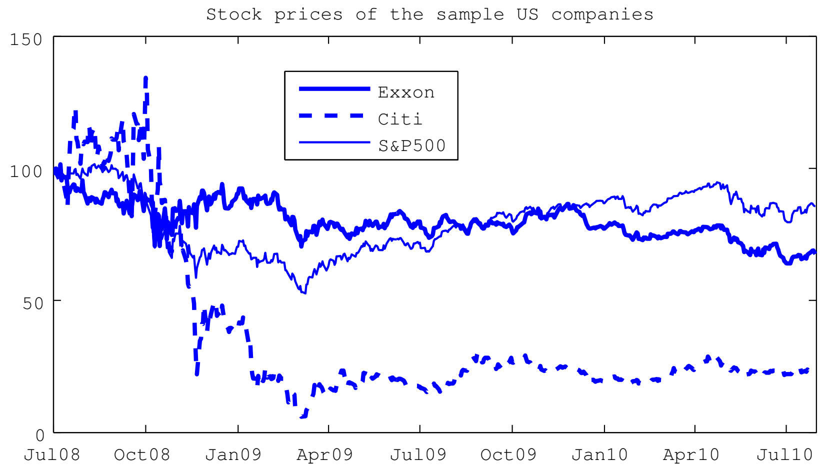
Figure 4: *The impact of credit risk on European option prices*



Note: The option prices are calculated using the common parameters (the underlying asset price: 50, maturity: 6 months and annualized volatility: 30%).

Figure 5: *The stock price of the sample companies*





Notes: All of the stock prices and the stock index are indexed at 100 on July 1, 2008. While Citi's stock price declined sharply in 2008-2009, all the other stock return of the sample companies are roughly equal to the market index except for BP in May 2010 during the BP oil spill.

Figure 6: *The 1 year CDS spread of the sample companies' senior debt*

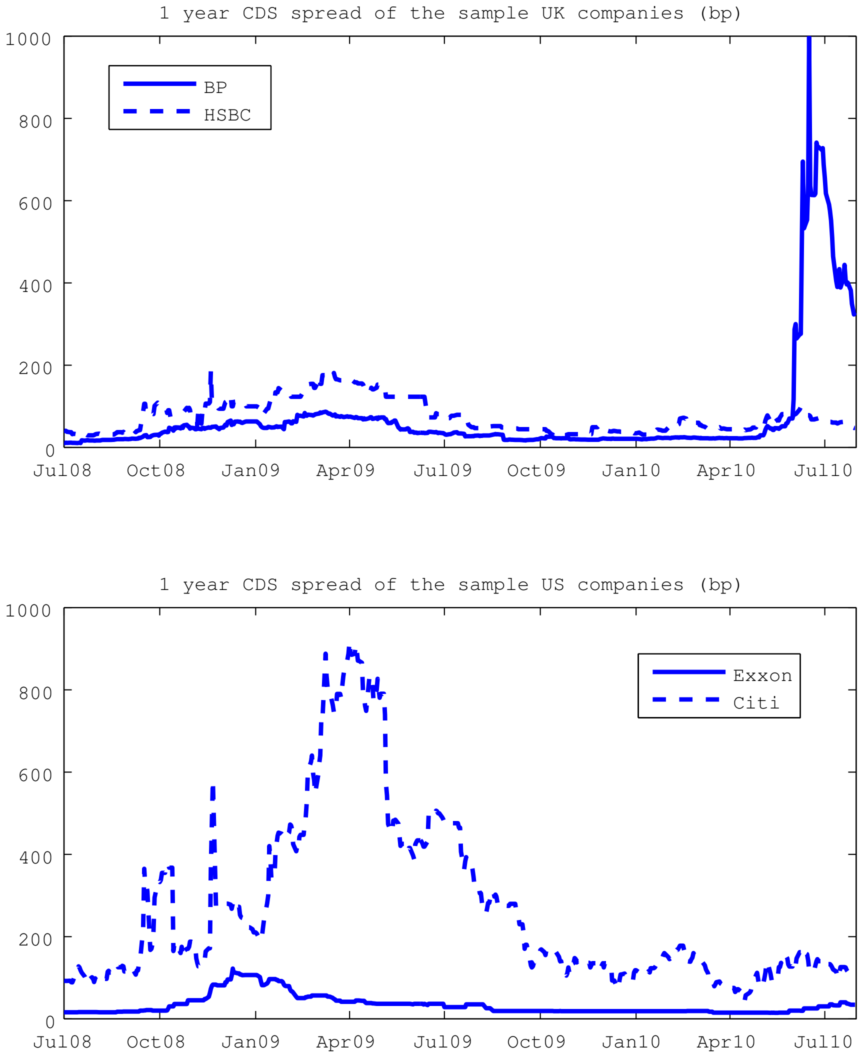
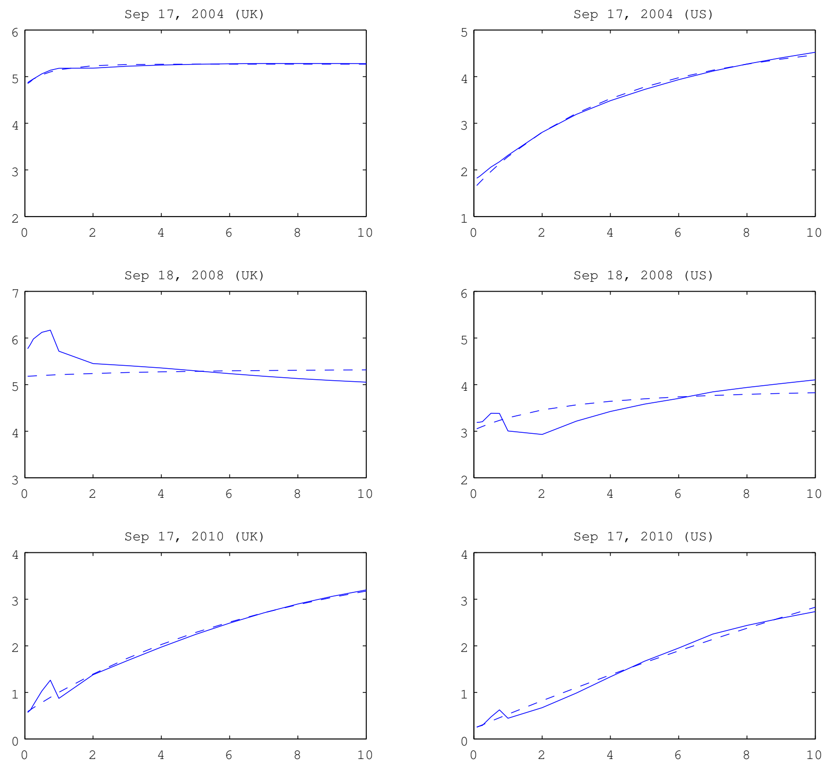
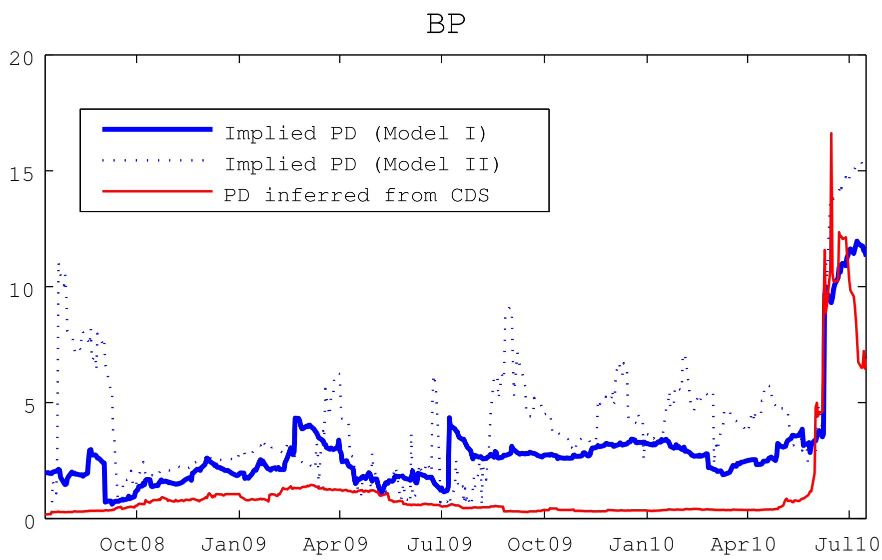


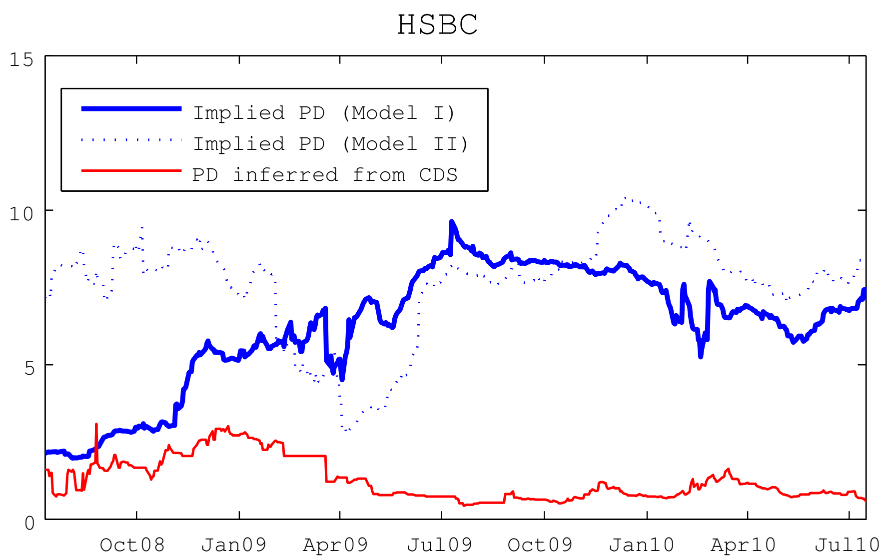
Figure 7: *The fitness of the term structure model*



Notes: The solid line is the actual term structure of LIBOR (less than 1 year) and Swap rate (longer than 1year). The dash line is the term structure of the Vasicek (1977) model. The parameters are calibrated with 3 month - 10 year data.

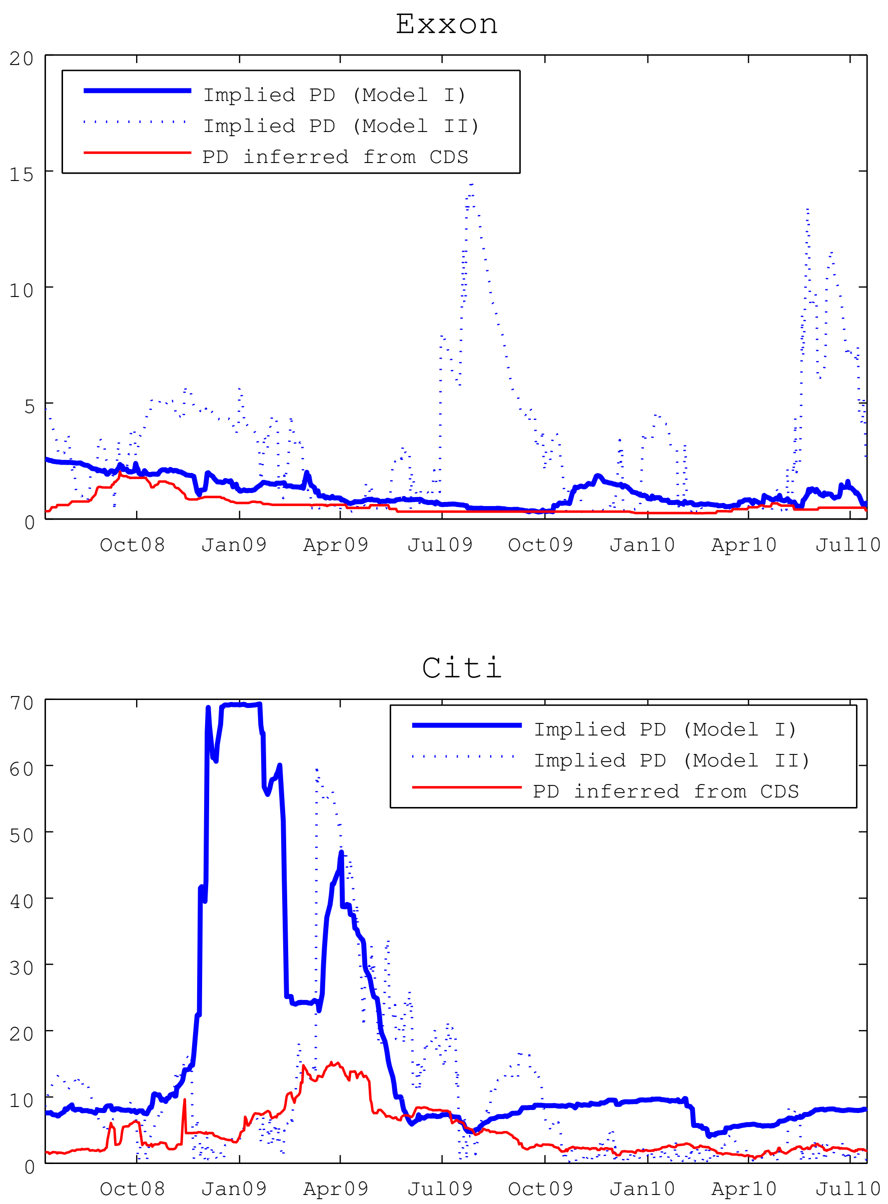
Figure 8: *The implied PD of the UK sample companies*





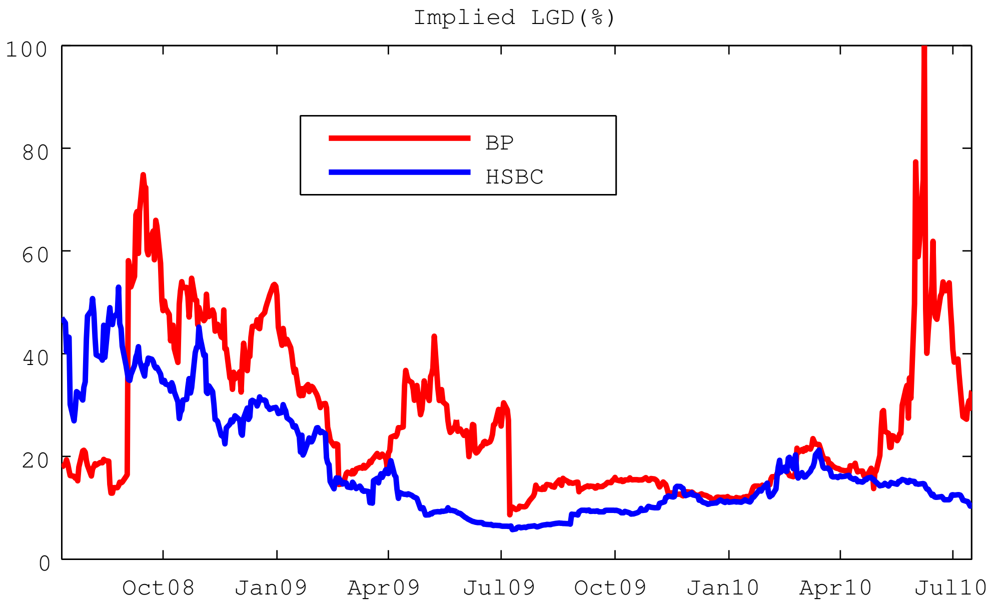
Notes: The thin curve is the PD inferred from 1year CDS spread with 60% LGD. The implied PD from model I (bold curve) and model II (dash curve), $\hat{\lambda}_t$, are adjusted using the 1year CDS spread of the United Kingdom.

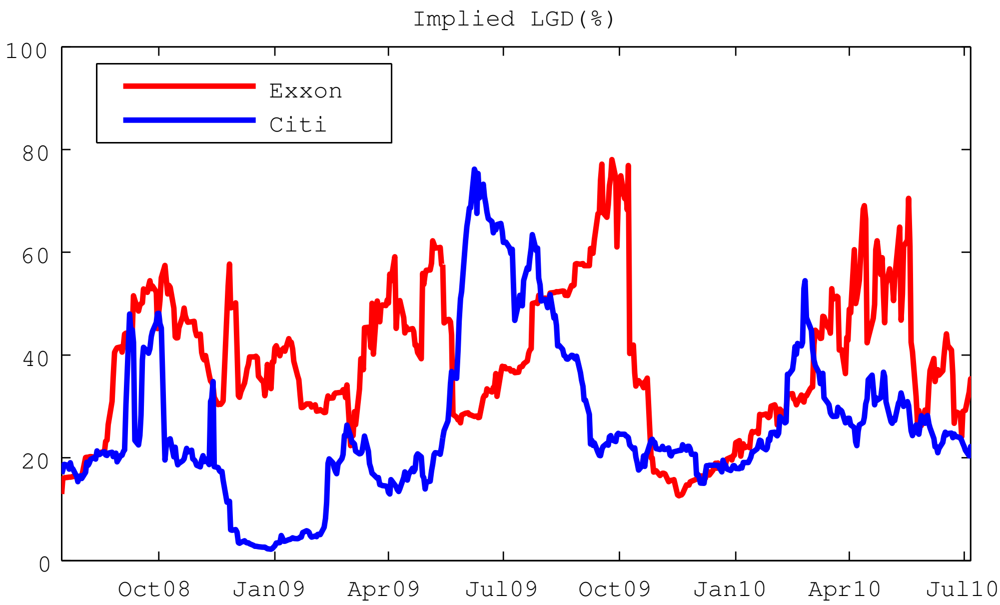
Figure 9: *The implied PD of the US sample companies*



Notes: The thin curve is the PD inferred from 1year CDS spread with 60% LGD. The implied PD from model I (bold curve) and model II (dash curve), $\hat{\lambda}_t$, are adjusted using the 1year CDS spread of the United States.

Figure 10: *The implied LGD of the senior CDS spread in the sample companies*





A The proof of equation (22)

A.1 Feynman - Kac formula for option pricing model

Here we denote the price of contingent claim of defaultable asset $h(x)$ as $P_{\epsilon,\delta}$ to emphasize the parameters, ϵ and δ. From the Feynman-Kac formula (See Karatzas and Shreve (1988) for details), $P_{\epsilon,\delta}$ is a solution of

$$\begin{aligned} \mathcal{L}^{\epsilon,\delta} P_{\epsilon,\delta}(t,x,r,y,\tilde{y},z) &= 0\,, && (33) \\ P_{\epsilon,\delta}(t,x,r,y,\tilde{y},z) &= h(x)\,, && (34) \end{aligned}$$

where the operator $\mathcal{L}^{\epsilon,\delta}$ is defined as

$$\mathcal{L}^{\epsilon,\delta} = \frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2 + \sqrt{\delta}\mathcal{M}_1 + \delta\mathcal{M}_2 + \frac{\sqrt{\delta}}{\sqrt{\epsilon}}\mathcal{M}_3\,, \tag{35}$$

in which respective operators are listed below,

$$\begin{aligned} \mathcal{L}_0 &= \nu^2\frac{\partial^2}{\partial y^2} + (m-y)\frac{\partial}{\partial y} + \tilde{\nu}^2\frac{\partial^2}{\partial \tilde{y}^2} + (\tilde{m}-\tilde{y})\frac{\partial}{\partial \tilde{y}} + 2\rho_{24}\nu\tilde{\nu}\frac{\partial^2}{\partial y\partial\tilde{y}}, \\ \mathcal{L}_1 &= \rho_2\sigma(\tilde{y})\nu\sqrt{2}x\frac{\partial^2}{\partial x\partial y} + \rho_{12}\eta\nu\sqrt{2}x\frac{\partial^2}{\partial r\partial y} + \rho_4\sigma(\tilde{y})\tilde{\nu}\sqrt{2}x\frac{\partial^2}{\partial x\partial\tilde{y}} \\ &\quad +\rho_{14}\eta\tilde{\nu}\sqrt{2}\frac{\partial^2}{\partial\tilde{y}\partial r} - \Lambda(\tilde{y})\tilde{\nu}\sqrt{2}\frac{\partial}{\partial\tilde{y}}, \\ \mathcal{L}_2 &= \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2(\tilde{y})x^2\frac{\partial^2}{\partial x^2} + (r+f(y,z))x\frac{\partial}{\partial x} + (\alpha_t-\beta r)\frac{\partial}{\partial r} + \sigma(\tilde{y})\eta\rho_1 x\frac{\partial^2}{\partial x\partial r} \\ &\quad +\frac{1}{2}\eta^2\frac{\partial^2}{\partial r^2} - (r+f(y,z))\,, \end{aligned}$$

$$\begin{aligned} \mathcal{M}_1 &= \sigma(\tilde{y})\rho_3 g(z)x\frac{\partial^2}{\partial x\partial z} + \eta\rho_{13}g(z)\frac{\partial^2}{\partial r\partial z}, \\ \mathcal{M}_2 &= c(z)\frac{\partial}{\partial z} + \frac{1}{2}g^2(z)\frac{\partial^2}{\partial z^2}, \\ \mathcal{M}_3 &= \rho_{23}\nu\sqrt{2}g(z)\frac{\partial^2}{\partial y\partial z} + \rho_{34}\tilde{\nu}\sqrt{2}g(z)\frac{\partial^2}{\partial\tilde{y}\partial z}. \end{aligned}$$

A.2 Asymptotic Expansions

We use an asymptotic expansion for the approximation of $P_{\epsilon,\delta}$ (see Bayraktar and Yang, 2011), as both $\sqrt{\epsilon}$ and $\sqrt{\delta} \to 0$, to derive an expansion of $P_{\epsilon,\delta}$ in powers of $\sqrt{\delta}$ and $\sqrt{\epsilon}$

respectively. First we expand $P_{\epsilon,\delta}$ in powers of $\sqrt{\delta}$

$$P_{\epsilon,\delta} = P_0 + \sqrt{\delta} P_{0,1} + \delta P_{0,2} + \cdots . \tag{36}$$

By substituting (36) into (33),

$$\begin{aligned}
\mathcal{L}^{\epsilon,\delta} P_{\epsilon,\delta} &= \left(\frac{1}{\epsilon}\mathcal{L}_0 P_{\epsilon,0} + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 P_{\epsilon,0} + \mathcal{L}_2 P_{\epsilon,0} \right) \\
&\quad + \sqrt{\delta} \left(\frac{1}{\sqrt{\epsilon}}\mathcal{L}_0 P_{\epsilon,1} + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 P_{\epsilon,1} + \mathcal{L}_2 P_{\epsilon,1} + \mathcal{M}_1 P_{\epsilon,0} + \frac{1}{\sqrt{\epsilon}}\mathcal{M}_3 P_{\epsilon,0} \right) \\
&\quad + \left(\sqrt{\delta}\right)^2 \left(\mathcal{L}_2 P_{\epsilon,2} + \mathcal{M}_2 P_{\epsilon,0} + \mathcal{M}_2 P_{\epsilon,0} + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 P_{\epsilon,2} + \frac{1}{\epsilon}\mathcal{L}_0 P_{\epsilon,2} \right) + \cdots \\
&= 0 .
\end{aligned} \tag{37}$$

Since $\mathcal{L}^{\epsilon,\delta} P_{\epsilon,\delta} = 0$, the first term of (37) is

$$\left(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2 \right) P_{\epsilon,0} = 0 \tag{38}$$

and given $\sqrt{\delta} \neq 0$, the second term is also zero,

$$\left(\frac{1}{\sqrt{\epsilon}}\mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2 \right) P_{\epsilon,1} = -\left(\mathcal{M}_1 + \frac{1}{\sqrt{\epsilon}}\mathcal{M}_3 \right) P_{\epsilon,0} . \tag{39}$$

Secondly, we expand the solution to (38) and (39) in powers of $\sqrt{\epsilon}$,

$$\begin{aligned}
P_{\epsilon,0} &= P_0 + \sqrt{\epsilon} P_{1,0} + \left(\sqrt{\epsilon}\right)^2 P_{2,0} + \left(\sqrt{\epsilon}\right)^3 P_{3,0} + \cdots , &(40)\\
P_{\epsilon,1} &= P_{0,1} + \sqrt{\epsilon} P_{1,1} + \left(\sqrt{\epsilon}\right)^2 P_{2,1} + \left(\sqrt{\epsilon}\right)^3 P_{3,1} + \cdots . &(41)
\end{aligned}$$

Inserting (40) into (38), we obtain

$$\begin{aligned}
&\left(\sqrt{\epsilon}\right)^{-2} \mathcal{L}_0 P_0 + \left(\sqrt{\epsilon}\right)^{-1} \left(\mathcal{L}_0 P_{1,0} + \mathcal{L}_1 P_0 \right) + \left(\mathcal{L}_0 P_{2,0} + \mathcal{L}_1 P_{1,0} + \mathcal{L}_2 P_0 \right) \\
&+ \sqrt{\epsilon} \left(\mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} \right) + \left(\sqrt{\epsilon}\right)^2 \left(\mathcal{L}_1 P_{3,0} + \mathcal{L}_2 P_{2,0} \right) + \cdots \\
&= 0 .
\end{aligned} \tag{42}$$

Since $\sqrt{\epsilon} \neq 0$, the first term of the equation (42) implies that P_0 is independent of y and $\tilde{y}$. As a result $P_{1,0}$ is also independent of y and $\tilde{y}$ ($\mathcal{L}_0 P_{1,0} = 0$) from the second term and the independence of P_0 on y and $\tilde{y}$ ($\mathcal{L}_1 P_0 = 0$).

Next we can reduce the third term using the independence of $P_{1,0}$ on y and $\tilde{y}$ ($\mathcal{L}_1 P_{1,0} = 0$),

$$\mathcal{L}_0 P_{2,0} + \mathcal{L}_2 P_0 = 0\,. \tag{43}$$

This is a Poisson equation for $P_{2,0}$ (see Fouque et al., 2000). The solvability of the equation requires the centering condition

$$\langle \mathcal{L}_2 \rangle P_0 = 0\,, \tag{44}$$

where $\langle \cdot \rangle$ denotes the average with respect to the invariant distribution of $(Y_t, \tilde{Y}_t)$, which have joint density

$$\begin{aligned} \Psi(y, \tilde{y}) \;&=\; \frac{1}{2\pi\nu\tilde{\nu}} \exp\left(-\frac{1}{2\left(1-\rho_{24}^2\right)} \left[\left(\frac{y-m}{\nu}\right)^2 + \left(\frac{\tilde{y}-\tilde{m}}{\tilde{\nu}}\right)^2 \right.\right. \\ &\left.\left. \quad -2\rho_{24} \frac{(y-m)(\tilde{y}-\tilde{m})}{\nu\tilde{\nu}} \right] \right)\,. \end{aligned} \tag{45}$$

Let us denote the averaging of $\mathcal{L}_2$ as

$$\begin{aligned} \langle \mathcal{L}_2 \rangle \;&=\; \frac{\partial}{\partial t} + \frac{1}{2}\bar{\sigma}_2(\tilde{y}) x^2 \frac{\partial^2}{\partial x^2} + (r + \bar{\lambda}(z)) x \frac{\partial}{\partial x} + (\bar{\alpha} - \beta r)\frac{\partial}{\partial r} \\ &\quad + \bar{\sigma}_1(\tilde{y})\eta\rho_1 x \frac{\partial^2}{\partial x \partial r} - (r + \bar{\lambda}(z))\,, \end{aligned} \tag{46}$$

where $\bar{\sigma}_1 = \langle \sigma(\tilde{y}) \rangle$, $\bar{\sigma}_2 = \langle \sigma^2(\tilde{y}) \rangle$, $\bar{\lambda}(z) = \langle f(y, z) \rangle$ and $\bar{\alpha} = \langle \alpha_t \rangle$. Under the terminal condition

$$P_0(T, x, r, z) = h(x)\,, \tag{47}$$

equation (44) defines the leading order term P_0. Using (43) we can also define the solution to the Poisson equation,

$$P_{2,0} = -\mathcal{L}_0^{-1} \left(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle\right) P_0\,. \tag{48}$$

Finally the fourth term of (42), term of $\sqrt{\epsilon}$ yields a Poisson equation

$$\mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} = 0\,. \tag{49}$$

For the complete identification of $P_{1,0}$, the solvability of this equation requires

$$\langle \mathcal{L}_2 P_{1,0} \rangle = -\langle \mathcal{L}_1 P_{2,0} \rangle = \left\langle \mathcal{L}_1 \mathcal{L}_0^{-1} \left(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle\right) \right\rangle P_0 \tag{50}$$

under the terminal condition

$$P_{1,0}(T,x,r,z)=0\,. \tag{51}$$

Next, we will express the centoring condition (50) more explicitly. Using (46), we can rewrite $\mathcal{L}_0 P_{2,0}$ as

$$\begin{aligned}(\mathcal{L}_2-\langle\mathcal{L}_2\rangle)\,P_0 &= \frac{1}{2}\left(\sigma^2(\tilde{y})-\bar{\sigma}_2\right)x^2\frac{\partial^2 P_0}{\partial x^2}+\left(\sigma(\tilde{y})-\bar{\sigma}_1\right)\eta\rho_1 x\frac{\partial^2 P_0}{\partial x\partial r}\\ &\quad+\left(f(y,z)-\bar{\lambda}(z)\right)\left(x\frac{\partial P_0}{\partial x}-P_0\right)+\left(\alpha_t-\bar{\alpha}\right)\frac{\partial P_0}{\partial r}\,. \end{aligned}\tag{52}$$

This is identical to equation (3.20) of Bayraktar and Yang (2010) except for the last term. Therefore,

$$\begin{aligned}\mathcal{L}_0^{-1}\left(\mathcal{L}_2-\langle\mathcal{L}_2\rangle\right)P_0 &= \frac{1}{2}\kappa(y,\tilde{y})x^2\frac{\partial^2 P_0}{\partial x^2}+\psi(y,\tilde{y})\eta\rho_1 x\frac{\partial^2 P_0}{\partial x\partial r}\\ &\quad+\phi(y,\tilde{y},z)\left(x\frac{\partial P_0}{\partial x}-P_0\right)+\varsigma\left(y,\tilde{y},r\right)\frac{\partial P_0}{\partial r}, \end{aligned}\tag{53}$$

where ψ, κ, ϕ and ς are the solutions to the Poisson equations

$$\mathcal{L}_0\psi(\tilde{y}) = \sigma(\tilde{y})-\bar{\sigma}_1, \tag{54}$$

$$\mathcal{L}_0\kappa(\tilde{y}) = \sigma^2(\tilde{y})-\bar{\sigma}_2, \tag{55}$$

$$\mathcal{L}_0\phi(y,z) = \left(f(y,z)-\bar{\lambda}(z)\right), \tag{56}$$

$$\mathcal{L}_0\varsigma\left(y,\tilde{y},r\right) = \alpha_t-\bar{\alpha}. \tag{57}$$

Applying the differential operator $\mathcal{L}_1$ to the last expression, we can calculate $P_{1,0}$ explicitly,

$$
\begin{aligned}
\left\langle \mathcal{L}_1 \mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) \right\rangle P_0 &= \sqrt{2}\left(\rho_2 \nu \langle \sigma \phi_y \rangle (z) - \frac{1}{2}\tilde{\nu} \langle \Lambda \kappa_{\tilde{y}} \rangle \right) x^2 \frac{\partial^2 P_0}{\partial x^2} \\
&+\sqrt{2}\left(\rho_{12}\eta\nu \langle \phi_y \rangle (z) \tilde{\nu} \langle \Lambda \varsigma_{\tilde{y}} \rangle \right) \frac{\partial}{\partial r}\left(x\frac{\partial P_0}{\partial x} - P_0 \right) \\
&+\frac{\sqrt{2}}{2}\rho_4 \tilde{\nu} \langle \sigma \kappa_{\tilde{y}} \rangle x \frac{\partial}{\partial x}\left(x^2 \frac{\partial^2 P_0}{\partial x^2} \right) \\
&+\sqrt{2}\rho_1 \rho_4 \tilde{\nu}\eta \langle \sigma \psi_{\tilde{y}} \rangle x \frac{\partial}{\partial x}\left(x\frac{\partial^2 P_0}{\partial x \partial r} \right) \\
&+\frac{\sqrt{2}}{2}\rho_{14}\eta\tilde{\nu} \langle \kappa_{\tilde{y}} \rangle \frac{\partial}{\partial r}\left(x^2 \frac{\partial^2 P_0}{\partial x^2} \right) \\
&+\sqrt{2}\rho_1 \rho_{14}\eta^2 \tilde{\nu} \langle \psi_{\tilde{y}} \rangle \frac{\partial}{\partial r}\left(x\frac{\partial^2 P_0}{\partial x \partial r} \right) \\
&+\sqrt{2}\left(\rho_2 \nu \langle \sigma \varsigma_y \rangle + \rho_4 \tilde{\nu} \langle \sigma \varsigma_{\tilde{y}} \rangle - \tilde{\nu} \langle \Lambda \psi_{\tilde{y}} \rangle - \tilde{\nu} \langle \Lambda \varsigma_{\tilde{y}} \rangle \right) x \frac{\partial^2 P_0}{\partial x \partial r} \\
&+\sqrt{2}\left(\rho_{12}\eta\nu \langle \varsigma_y \rangle + \rho_{14}\eta\tilde{\nu} \langle \varsigma_{\tilde{y}} \rangle \right) \frac{\partial^2 P_0}{\partial r^2}. \qquad (58)
\end{aligned}
$$

It is possible to get the explicit expression of P_1^ϵ by inserting (41) into (39). Discussed above, $P_{0,1}$ is independent of y and $\tilde{y}$ and satisfies

$$
\langle \mathcal{L}_2 \rangle = -\langle \mathcal{M}_1 \rangle P_0, \;\; P_{0,1}(T, x, r; z) = 0 \qquad (59)
$$

Using Propositions 3.2 and 3.3 and Remark 3.1 in Bayraktar and Yang (2011), the first order expansions on $\sqrt{\epsilon}$ and $\sqrt{\delta}$ are derived respectively

$$
\begin{aligned}
\sqrt{\epsilon}P_{1,0} &= -(T-t)\left(V_1^\epsilon(z) x^2 \frac{\partial^2 P_0}{\partial x^2} + V_2^\epsilon x \frac{\partial}{\partial x}\left(\frac{\partial^2 P_0}{\partial x^2} \right) \right) \\
&+V_3^\epsilon(z)\left(-x\frac{\partial^2 P_0}{\partial x \partial \alpha} - \frac{\partial P_0}{\partial \alpha} \right) + V_4^\epsilon x^2 \frac{\partial^3 P_0}{\partial x^2 \partial \alpha} + V_5^\epsilon x \frac{\partial^2 P_0}{\partial \eta \partial x} + V_6^\epsilon x \frac{\partial^2 P_0}{\partial x \partial \alpha} \\
&+V_7^\epsilon \frac{\partial^2 P_0}{\partial r^2} \qquad (60) \\
\sqrt{\delta}P_{0,1} &= V_1^\delta(z)\frac{(T-t)^2}{2}\left(x^2 \frac{\partial^2 P_0}{\partial x^2} \right) + V_2^\delta \left[\left(x\frac{\partial^2 P_0}{\partial \alpha \partial x} - \frac{\partial P_0}{\partial \alpha} \right) \right. \\
&\left. -(T-t)\left(x\frac{\partial^2 P_0}{\partial r \partial x} - \frac{\partial P_0}{\partial r} \right) + \frac{(T-t)^2}{2}\left(x^2 \frac{\partial^2 P_0}{\partial x^2} - x\frac{\partial P_0}{\partial x} + P_0 \right) \right] \qquad (61)
\end{aligned}
$$

in which

$$\begin{aligned}
V_1^\epsilon(z) &= \sqrt{\epsilon}\left(\rho_2\nu\sqrt{2}\left\langle\sigma\phi_y\right\rangle(z)-\frac{1}{2}\tilde{\nu}\sqrt{2}\left\langle\Lambda\kappa_{\tilde{y}}\right\rangle\right)\,,\; V_2^\epsilon=\frac{\sqrt{2}}{2}\sqrt{\epsilon}\rho_4\tilde{\nu}\left\langle\sigma\kappa_{\tilde{y}}\right\rangle\,,\\
V_3^\epsilon(z) &= \sqrt{\epsilon}\left(\rho_{12}\eta\nu\sqrt{2}\left\langle\phi_y\right\rangle(z)+\sqrt{2}\tilde{\nu}\left\langle\Lambda\varsigma_{\tilde{y}}\right\rangle\right)\\
V_4^\epsilon &= -\sqrt{\epsilon}\left(\frac{1}{2}\rho_{14}\eta\tilde{\nu}\sqrt{2}\left\langle\kappa_{\tilde{y}}\right\rangle-\rho_4\tilde{\nu}\sqrt{2}\left\langle\sigma\psi_{\tilde{y}}\right\rangle\eta\rho_1+\rho_{14}\eta\tilde{\nu}\sqrt{2}\left\langle\psi_{\tilde{y}}\right\rangle\bar{\sigma}_1\rho_1^2\right)\,,\\
V_5^\epsilon &= -\sqrt{\epsilon}\left(\rho_{14}\eta\tilde{\nu}\sqrt{2}\left\langle\psi_{\tilde{y}}\right\rangle\rho_1\right)\\
V_6^\epsilon &= \sqrt{\epsilon}(-\sqrt{2}\rho_4\tilde{\nu}\left\langle\sigma\psi_{\tilde{y}}\right\rangle\eta\rho_1+\sqrt{2}\rho_{14}\eta\tilde{\nu}\left\langle\psi_{\tilde{y}}\right\rangle\bar{\sigma}_1\rho_1^2\\
&\quad -\sqrt{2}\tilde{\nu}\eta\left\langle\Lambda\psi_{\tilde{y}}\right\rangle\rho_1\rho_2\nu\left\langle\sigma\varsigma_y\right\rangle+\sqrt{2}\rho_4\tilde{\nu}\left\langle\sigma\varsigma_{\tilde{y}}\right\rangle-\sqrt{2}\tilde{\nu}\left\langle\Lambda\varsigma_{\tilde{y}}\right\rangle)\\
V_7^\epsilon &= \sqrt{\epsilon}\left(\sqrt{2}\rho_{12}\eta\nu\left\langle\varsigma_y\right\rangle+\sqrt{2}\rho_{14}\eta\tilde{\nu}\left\langle\varsigma_{\tilde{y}}\right\rangle\right)\\
V_1^\delta(z) &= \sqrt{\delta}\bar{\lambda}(z)\bar{\sigma}_1\rho_3 g(z)\,,\; V_2^\delta(z)=\sqrt{\delta}\bar{\lambda}(z)\eta\rho_{13}g(z)\,.
\end{aligned}$$