THE GARCH STRUCTURAL CREDIT RISK MODEL: SIMULATION ANALYSIS AND APPLICATION TO THE BANK CDS MARKET DURING THE 2007-2008 CRISIS

ABSTRACT. We develop a structural credit risk model in which the asset volatility of the firm follows a GARCH process, as in Heston and Nandi (2000). We benchmark the out-of-sample model prediction accuracy against the calibrated Merton (1974) model and the Duan (1994) ML estimation of the Merton model, both using simulated data and in an empirical application to the bank CDS market in the US during the crisis period 2007-2008. The GARCH model outperforms the competitors in out-of-sample spread prediction in both cases. We document a high incidence of empirical bank CDS spread term structure inversion, and analyze its relationship with model performance.

Key words and phrases. Heston-Nandi Model; Merton Model; Risk Management; Structural Credit Risk Models; GARCH.

Structural bond pricing models are now used by institutions around the world to value the risky debt of banks and firms, as well as for applications such as the determination of capital adequacy ratios. It was Merton (1974) who first adapted the Black and Scholes (1973) and Merton (1973) option pricing framework to the valuation of corporate securities. In the "Merton model", as it is commonly known, the value of the risky debt plus the equity of a firm at any time must equal the value of the firm's assets. The risky debt can be valued as a risk-free bond minus the value of an implicit put option, whose strike price is equal to the present value of promised debt repayments discounted at the risk-free rate. The equity of the firm is valued as a European call option, with the same strike price. This strike price of the implicit options in the Merton model is commonly referred to as the "default barrier". In practice, the time series of asset values of the firm, and the volatility of asset returns, cannot be observed directly, and must be inferred from the time series of observed equity values.

There are two primary approaches to solving the problem of inferring the value and volatility of firm assets from information on firm equity. Merton (1974) solves this problem using a calibration technique. The two unknowns in the problem are the current asset value and the asset return volatility. The first of the two equations needed to solve for these unknowns is provided by the pricing equation for the implicit call option of equity. The second equation is given by the relationship between the volatility of equity returns and the volatility of asset returns, which is found by an application of Itô's lemma to the diffusion followed by the call option. Having solved for the two unknowns, one can calculate the implied value of the risky debt, as well as other risk measures, such as the spread over the risk-free rate associated with the debt of the firm, and the probability of default, over a given time horizon.

A second approach to inferring the current asset value and volatility in the Merton (1974) model is provided by Duan (1994) and Duan et al. (2004). The Duan (1994) maximum likelihood method views the observed equity time series as a transformed data set with the equity pricing formula defining the transformation. The best-known commercial implementation of the Merton (1974) structural credit risk model, which has been applied to tens of thousands of firms and banks around the world, is due to Moody's-KMV. Duan et al. (2004) show that their maximum likelihood method for estimating the parameters of structural credit risk models produces the same point estimates for the unobserved parameters as the Moody's-KMV method, in the special case of the Merton (1974) model, but that the methods will in general produce different point estimates for more general structural credit risk models. It has been shown by Ericsson and Reneby (2005), in a simulation study, that the maximum likelihood approach of Duan (1994) to estimating structural bond pricing models is superior to the calibration approach of estimating the Merton (1974) model, in the sense that it provides a less biased and more efficient estimator of asset values, asset volatilities, and spreads. In other work, Eom et al. (2004) perform an empirical analysis of the relative performance of five different structural credit risk models using actual bond data, but do not estimate any of those models using the maximum likelihood approach suggested by Duan (1994) and implemented for the Merton model by Duan et al. (2004). Eom et al. (2004), rather, use the calibration approach for estimating the asset volatility originally proposed by Merton (1974), which remains the most common approach in the academic literature and in practice, in part due to its ease of implementation.

In practice, the asset return volatility found using Merton's (1974) calibration method can exhibit significant variation over time for many firms and banks. This is also true of the asset return volatilities calculated using the maximum likelihood method of Duan (1994) and the asset return volatilities reported by the commercial software of Moody's-KMV. Significant shifts in estimated asset return volatility can induce significant shifts in risk indicators, such as spreads. There is strong reason, therefore, to believe that the volatility of firms asset returns is stochastic, rather than constant. Firms and industries alike go through periods of high levels of uncertainty regarding their future rates of asset growth, as well as periods of relative tranquility. Moreover, it has been shown that, in equity and other markets, innovations in volatility are significantly negatively correlated with spot returns: the so-called "volatility leverage effect" documented by Christie (1982) and others. It is reasonable to suspect that there might be a similar relationship between firms asset returns and asset return volatility. This has implications for bond pricing, because a fall in returns coupled with a rise in volatility can have a doubly negative impact on the valuation of firm debt.

In light of these issues, we propose a structural credit risk model that allows for the stochastic volatility of asset returns. The tools for the construction of such a model, in fact, exist in the option pricing literature. In the work following the seminal paper of Black and Scholes (1973) and Merton (1973), it was recognized that the assumption of constant asset return volatility in option pricing is too restrictive. For this reason, the academic literature set itself to the task of pricing options on an underlying whose volatility can be time-varying (Engle, 1982; Bollerslev, 1986; Jacquier et al., 1994; Nicolato and Venardos, 2003; Heston and Nandi, 2000; Duan, 1995). However, in the majority of time-varying volatility models, there exists no closed-form solution for the option price. As a result, one has to use Monte Carlo methods instead to calculate option prices (Christoffersen and Jacobs, 1995). In order to circumvent that problem, Heston and Nandi (2000) proposed a closed-form option pricing model in which asset returns follow a GARCH process. The resulting option pricing formula closely resembles the one derived in Black and Scholes (1973).

Due to the analytical convenience of the Heston and Nandi (2000) model, which we henceforth refer to as HN, it is that model we shall adapt for the construction of our structural credit risk model of the firm. We refer to our model as the "GARCH structural credit risk model", in light of the fact that firm asset return volatility (in fact, volatility squared) is assumed to follow a GARCH process, as in HN. The main analytical challenge of implementing our model, as usual, is the need to estimate the current value of assets and the parameters of the asset return and return volatility processes using only the observed values of firm equity. To that end, we derive an Expectations

Maximization (EM) algorithm to compute the maximum likelihood estimates for the model parameters, in a manner equivalent to that used by Duan (1994) and Duan et al. (2004) in the case of the Merton model with constant volatility. We choose this approach due to the demonstrable superiority of maximum likelihood methods for estimation, as documented by Ericsson and Reneby (2005) in the case of several previous structural credit risk models, including that of Merton (1974).

The rest of the paper proceeds as follows. In Section 2, we briefly review the HN model for pricing options on assets whose return volatility follows a GARCH process. Section 3 lays out the GARCH structural credit risk model and describes our EM algorithm for estimating the model. Section 4 benchmarks our model against the Merton model estimated using Merton's (1974) calibration technique, and against the Merton model estimated using Duan's (1994) maximum likelihood technique using simulated data. We consider both the case where the data generating process exhibits stochastic volatility, and the case where volatility is constant. Section 5 presents a practical application of our method to the estimation of fair spreads on the debt of selected investment banks during the 2007-2008 credit crunch in the United States. Section 6 concludes.

I. THE HESTON AND NANDI (2000) MODEL FOR OPTION PRICING

We now recall briefly the main assumptions and results of the HN option pricing model. If S_t is the price of the underlying at time t, define the log return at time t as $r_t \equiv \log\left(\frac{S_t}{S_{t-\Delta}}\right)$, where returns are calculated over a time interval of length Δ. The joint dynamics of the log returns and the return volatility are given as follows (Heston and Nandi, 2000; Rouah and Vainberg, 2007):

$$r_t = r + \lambda\sigma_t^2 + \sigma_t z_t \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \beta_i \sigma_{t-i\Delta}^2 + \sum_{i=1}^{q} \alpha_i (z_{t-i\Delta} - \gamma_i \sigma_{t-i\Delta})^2 \tag{2}$$

Here r is the risk-free interest rate, σ_t^2 is the conditional variance at time t, z_t is a standard normal disturbance, and $\omega, \beta_i, \alpha_i, \gamma_i,$ and λ are the HN model parameters. Note that the conditional variance $h(t)$ appears in the mean as a return premium, with coefficient λ, which allows the average spot return to depend on the level of risk. Henceforth, we will focus on the first order case of the above model, with $p = q = 1$, and drop the i subscripts on the α, β and γ parameters, as there is only a one-period lag.

Heston and Nandi (2000) note the following facts about the first order version of the model. First, the process is stationary with finite mean and variance when $\beta + \alpha\gamma^2 < 1$. Second, the one period ahead variance of the process, $\sigma_{t+\Delta}^2$, can be directly computed at time t as a function of the current period log return r_t, as follows:

$$\sigma_{t+\Delta}^2 = \omega + \beta\sigma_t^2 + \alpha\frac{(r_t - r - (\lambda+\gamma)\sigma_t^2)^2}{\sigma_t^2} \tag{3}$$

Third, the parameter α determines the kurtosis of the distribution and $\alpha = 0$ implies a deterministic time varying variance. Fourth, the γ parameter allows shocks to the return process to have an asymmetric influence on the variance process, in the sense that a large negative shock z_t raises the variance more than a large positive z_t. That is, the parameter γ controls the skewness of the distribution of log returns, and the distribution of log returns is symmetric when γ and λ are both equal to zero. Finally, the correlation between volatility and realized log returns is given by

$$Cov_{t-\Delta}[\sigma_{t+\Delta}^2, r_t] = -2\alpha\gamma\sigma_t^2 \tag{4}$$

Positive values for α and γ imply a negative correlation between volatility and spot returns, which is consistent with the leverage effect documented by e.g. Christie (1982).

We now turn to the issue of pricing contingent claims on the underlying S_t. In order to value such claims, we need to derive the risk-neutral distribution of the spot price. As HN show, this is accomplished by transforming model equations (1) and (2) so that the log return corresponding to the expected spot price is the risk-free rate, and invoking the assumption that the value of a call option one period to expiration is given by the Black-Scholes-Rubenstein formula. This second assumption ensures that the distribution of the transformed shocks, $z_t^\star$, is a standard normal under the risk-neutral probabilities. The risk-neutral version of the model is given by

$$r_t = r + \lambda^\star \sigma_t^2 + \sigma_t z_t^\star \tag{5}$$

$$\sigma_t^2 = \omega + \beta \sigma_{t-\Delta}^2 + \alpha (z_{t-\Delta}^\star - \gamma^\star \sigma_{t-\Delta})^2 \tag{6}$$

where the transformed parameters are

$$\begin{aligned} \lambda^\star &= -\frac{1}{2} \\ \gamma^\star &= \gamma + \lambda + \frac{1}{2} \\ z_t^\star &= z_t + \left(\lambda + \frac{1}{2}\right) \sigma_t \end{aligned}$$

Heston and Nandi (2000) show that the price of an European call option, with maturity T and strike price K, can be computed as the discounted expected value of the call option payoff function $\max[S_T - K, 0]$ under the risk neutral measure, which is computed using the formula they derive for the characteristic function of the log spot price under the risk neutral measure. The formula for the HN call option price is given by:

$$C_t = S_t \mathbb{P}_1 - K e^{-r(T-t)} \mathbb{P}_2 \tag{7}$$

where

$$\mathbb{P}_1 = \frac{1}{2} + \frac{e^{-r(T-t)}}{\pi S_t} \int_0^\infty \mathcal{R} \left[\frac{K^{-i\phi} f^\star(i\phi + 1)}{i\phi} \right] d\phi$$

and

$$\mathbb{P}_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathcal{R}\left[\frac{K^{-i\phi} f^\star(i\phi)}{i\phi}\right] d\phi$$

are the delta of the call value, and the risk neutral probability of the asset price S_T being greater than the strike price K at maturity, respectively. Here $\mathcal{R}(.)$ denotes the real part of the complex number that is its argument. The generating function of the return process in the model defined by (1) and (2) is given by

$$f(\phi) = S_t^\phi \exp\left(A_t + B_t \sigma^2_{t+\Delta}\right), \tag{8}$$

with coefficient functions defined by

$$A_t = A_{t+\Delta} + \phi r + B_{t+\Delta}\omega - \frac{1}{2}\ln\left(1 - 2\alpha B_{t+\Delta}\right) \tag{9}$$

$$B_t = \phi(\lambda + \gamma) - \frac{1}{2}\gamma^2 + \beta B_{t+\Delta} + \frac{0.5(\phi - \gamma)^2}{1 - 2\alpha B_{t+\Delta}} \tag{10}$$

The generating function under the risk neutral measure, $f^\star(\phi)$, is obtained by substituting $\lambda^\star$ and $\gamma^\star$ for λ and γ in the function $f(\phi)$ above. Finally, the coefficient functions A_t and B_t are solved recursively, given the terminal conditions $A_T = 0$ and $B_T = 0$. The function $f^\star(i\phi)$ is the characteristic function of the logarithm of the stock price under the risk neutral measure, and is calculated by replacing ϕ with $i\phi$ everywhere in the generating function $f^\star(\phi)$. Feller (1971) shows how to calculate probabilities and risk-neutral probabilities by inverting the characteristic function. In practice, once we estimate the parameters of the GARCH model given by equations (1) and (2) given data on the underlying asset returns, we simply have to compute the values of $\lambda^\star$ and $\gamma^\star$ in order to calculate option prices under the risk-neutral measure. For more details on the HN model with $p > 1$ and $q > 1$, see Heston and Nandi (2000) and Rouah and Vainberg (2007).

Using the HN model described above, one can price a series of European call options $\{C_t\}_{t=1}^n$, given the time series for the underlying asset $\{S_t\}_{t=1}^n$, as well as the time series of strike prices $\{K_t\}_{t=1}^n$ and the time series of risk-free rates $\{r_t^f\}_{t=1}^n$ over the sample. The relevant procedure for doing this is, first, to use the time series of asset prices to estimate the parameters of the valuation formula using the Maximum Likelihood Estimation (MLE) procedure employed by Bollerslev (1986), Heston and Nandi (2000), and others, and second, to apply the pricing formulas just stated, inserting the values for the estimated parameters and the relevant values for the strike prices and risk free rates at each point in time.

We will now describe the GARCH structural credit risk model and develop an estimation procedure that allows us to retrieve the time series of the asset levels and return volatility, given that we are only able to observe the time series for the equity of the firm. From now on, unless otherwise stated, we will assume that the frequency of the data, $\Delta = 1$, corresponds to one week of data.

II. THE GARCH STRUCTURAL CREDIT RISK MODEL AND ITS ESTIMATION

Consider the following re-interpretation of the HN model, in the spirit of Merton (1974). In the new setup, suppose that the underlying asset, relabeled V_t, represents the asset value of a firm at time t, which cannot be observed directly. The firm has traded debt, with market value D_t, and traded equity, with market value equal to E_t. The value of both debt and equity is observable in the market, and at all times we must have that $V_t = D_t + E_t$. The present value of the promised payments on the debt of the firm is equal to $e^{-r(T-t)}K$, where K is the dollar amount due at time T, when the model is presumed to end. When we reach the terminal date T, the firm will find itself in one of two situations. If the value of its assets $V_T > K$, it will pay off the promised value of its debt to debtholders and will pay the residual, $V_T - K$, to equityholders. Conversely, if $V_T < K$, the firm will default, leaving equityholders with nothing, and will turn over its terminal asset value V_T to debtholders, who will experience a partial default on what they are owed.

Merton (1974) showed that in this setup, the value of equity, E_t, is priced as a European call option with strike price K on the assets of the firm. The risky debt, D_t, is priced as the sum of a risk-free bond, worth the present value of promised debt payments K, minus the value of a European put option with strike price K, which represents the present value of the expected loss on the debt. The only substantial difference between the GARCH structural credit risk model and Merton's (1974) structural credit risk model, is that in our model, the unobserved asset returns of the firm exhibit stochastic volatility, and their dynamics is described by the HN model equations (1) and (2), with S_t replaced by V_t.

The problem that concerns us is that of using the time series of equity market values, $\{E_t\}_{t=1}^{n}$, to infer values for the parameters of the underlying HN model for the firm value V_t, as well as for the levels of the firm value at each point in time. Once we have estimated the aforementioned values and parameters, we may then price the risky debt D_t, using the formula

$$D_t \;=\; e^{-r(T-t)}K - P_t^{HN} \tag{11}$$

where $P_t^{HN}(V_t, K, r, T, \theta) = e^{-r(T-t)}E_t^{\star}\max[K - V_T, 0]$ is the price of a European put option, computed under the HN model, given the parameter set $\theta \equiv \{\alpha, \beta, \gamma, \lambda, \omega\}$. In particular, we will be concerned with computing the credit risk premiums, or spreads, on the risky debt of the firm at each point in time, using our estimates for the asset value and the HN model parameters derived using the time series of equity. The credit spread for a given terminal date T is defined by the relationship $s_t = y_t - r$, where y_t is the yield derived from the equation $e^{-y_t(T-t)}K = D_t$. This yields the formula

$$s_t = -\frac{1}{T-t}\ln\left(1 - \frac{P_t^{HN}}{Ke^{-r(T-t)}}\right) \tag{12}$$

Since the ability to compute fair values for spreads is so important in practice, our benchmarking of our model in the next section will include an analysis of its relative performance in the task of matching the true spreads computed under the data generating process against the methods of Merton (1974) and Duan et al. (2004). In our application section, we will compare the spreads computed using our GARCH structural credit risk model against the spreads from the Merton (1974) and Duan et al. (2004) methods on actual CDS spreads for investment banks during the 2007-2008 credit crunch.

Having framed the problem at hand, we now present a method for the estimating the firm asset levels and HN parameters from observed equity data. The estimation approach is based on an Expectation Maximization (EM) algorithm. The algorithm cycles through various steps, which are presented below; we relegate the proof of convergence to the Appendix.

EM Algorithm for the Estimation of the GARCH Structural Credit Risk Model:

(1) Set the elements of the HN parameter vector θ^0 equal to 0.001, where $\theta^0 \equiv \{\alpha^0, \beta^0, \gamma^0, \lambda^0, \omega^0\}$. Initialize the elements of the vector for the times series of asset values $\{V_t^0\}_{t=1}^n$ to any value k.

(2) Given the parameter vector θ^i computed in iteration i, compute the vector of asset values $\{V_t^{i+1}\}_{t=1}^n$ for the iteration $i+1$ by inverting the call option equation (7), given the time series of equity values $\{E_t\}_{t=1}^n$.

(3) Compute the series of log returns using the formula $r_t^{(i+1)} = \log\left(V_t^{(i+1)}/V_{t-1}^{(i+1)}\right)$ from the extracted series of asset values $\{V_t^{(i+1)}\}_{t=1}^n$. Using this log return series, express the conditional variances $\{\sigma_{t+j}^{2,(i+1)}\}_{j=1}^{n-1}$ using equation (3), as a function of the model parameters, and apply the MLE method of Bollerslev (1986) to estimate the parameter vector $\theta^{(i+1)} \equiv \{\alpha^{(i+1)}, \beta^{(i+1)}, \gamma^{(i+1)}, \lambda^{(i+1)}, \omega^{(i+1)}\}$ of the HN model.

(4) Repeat the last two steps until a tolerance level for the convergence criterion is reached.

The convergence of the above algorithm is guaranteed, but as usual in nonlinear problems with potentially multi-modal likelihood functions, convergence can be slow, and multiple solutions are possible. Thus, it is useful to experiment with multiple initial seeds of the algorithm in order to verify that the estimates of the parameters are robust to choice of initial values. We made an effort to do this in the current study, in particular in our empirical application to the banks that follows our simulation study below.

III. BENCHMARKING THE MODEL USING SIMULATED DATA

We now benchmark the GARCH structural credit risk model against the calibrated Merton (1974) model, and the Duan et al. (2004) maximum likelihood estimation of the Merton model, using simulated data. Our experimental design is as follows. We consider four scenarios, each of which represents a different, $2X2$ combination of asset volatility (business risk) and leverage ratio (financial risk) for the firm. The two values considered for the annualized asset volatility are 20% and 40%, and the initial ratios of the default barrier to assets (K/V) considered are 0.5 and 1.0. These ratios are generated by setting the initial asset value equal to 100 and setting the value of K equal to 50 or 100, respectively. Our decision to consider a scenario in which the leverage ratio is equal to 1.0, which is quite high, is motivated by our desire to generate scenarios that might give some insight into the empirical application in the following section to banks and financial companies in the US during the 2007-2008 period. In fact, the empirical evidence indicates that such a leverage ratio may even be an underestimate of the leverage ratios of some major investment banks during that period.

Second, within the aforementioned setup, we evaluate the models' performances under two different data generated processes (DGPs) for the underlying asset returns and return volatility. The first DGP is the discrete time version of the Black and Scholes (1973) model, with constant volatility, found by setting $\alpha = \beta = 0$ in the HN model with one lag. This case is of interest because it allows us to measure the performance of the GARCH structural credit risk model estimation algorithm against the Merton and Duan et al. methods when the true model is actually that assumed by the latter models. The second DGP is the HN model with $\alpha > 0$ and $\beta > 0$. This second DGP is of interest because it allows us to measure the difference in performance between the Merton and Duan et al. methods and our method in a setting of stochastic volatility.

The consideration of two DGPs, together with the four combinations of business risk and financial risk mentioned above, generates eight distinct simulation experiments. In each experiment, we proceed as follows. The time horizon is assumed to be one year throughout, and the risk-free rate is constant and set equal to 5%. In the case that the DGP of the firm's assets is a geometric Brownian motion, we set the asset value, the default barrier, and the asset volatility according to the desired combination of business and financial risk, simulate 52 weeks of asset values, and then simulate 52 weeks of equity values by pricing equity as a call option on the firm's assets using the Black-Scholes formula. We then take that 52 week time series of equity values as an input to our three models under consideration, the calibrated Merton model, the Merton model estimated using the maximum likelihood method of Duan et al., and the GARCH structural credit risk model proposed in the current paper. We produce estimates of the asset value, the asset volatility, and the fair spread on the firm's debt, priced according to each of our three models, for the last week in the 52 week sample. These estimates can be compared directly against the true asset value, asset volatility, and spread calculated under the Merton model for the last day in the sample, and such comparisons serve as a measure of the out-of-sample accuracy of each model under study.

In the case where the DGP of the asset time series is a GARCH process, the above procedure is repeated, except that it is necessary to parameterize the GARCH process to produce the asset volatility required by the experiment. There are multiple parameter combinations that accomplish this, and we report the combinations used in our experiments as a footnote to the results. For each unique experiment, we ran 100 simulations of the 52 week period necessary to log one out-of-sample vector of assets, asset volatility, and spreads per model. Using the histogram of these 100 out-of-sample vectors of results, along with the true asset values, volatilities, and spreads, we calculate the sample mean and standard deviation of the difference between the predicted and true asset value, asset volatility, and spread for each model, in each of the eight scenarios described above. These results are summarized in Table III.

[INSERT TABLE III HERE]

Our simulation study reveals several noteworthy results regarding the relative performance of the three models. Let us focus first on the case where the DGP for the firm asset is geometric Brownian motion. The first noticeable trend is that increasing business risk and increasing financial risk are associated with larger average out-of-sample estimation errors for the asset, the volatility, and the spread for each of the three models, as well as larger error standard deviations. Second, the Duan and GARCH models outperform the calibrated Merton model in all cases, in terms of achieving a lower average error in estimation of assets, asset volatility, and the spread, except for the low business risk, low financial risk case, in which the performance of the three models is similar in their estimation of the asset level and volatility. This result is in general consistent with the comparison of the Merton and Duan methods in Ericsson and Reneby (2005). The GARCH and Duan models still outperform the calibrated Merton model in their estimation of spreads in that case, however. Third, the Duan model has a lower average prediction error for spreads than the GARCH model in three of the four scenarios, although in all cases, and for all three of the variables studied, the difference in the average prediction error between the Duan and the GARCH model is small. Overall, we surmise that the slight under-performance of the GARCH model versus the Duan model in the case where the asset follows a geometric Brownian motion is that the extra complexity of the GARCH model produces a slight cost in terms of estimation precision when asset volatility is constant.

Now let us turn to the case where the true DGP for the asset is a GARCH process with stochastic volatility. The first noteworthy generalization is that, in all of the four cases studied, the GARCH model outperforms both the Duan and the calibrated Merton model in the out-of-sample estimation of asset values, asset volatility, and spreads, in terms of having lower average sample errors for all three variables of interest. Second, in all except the low business risk, high financial risk case, the calibrated Merton model actually achieves a lower average spread estimation error than the Duan model. However, the preceding observation must be qualified by noting that the spread error standard deviations are very high compared to the absolute value of the errors in all cases, and that the average spread estimation errors for the Merton and Duan models are generally closer to each other than they are to the average estimation error for the GARCH model. Finally, the standard deviation of the estimation error for each variable of interest, in the case of the GARCH model increases with higher business risk, and with higher financial risk, with one exception, which is a slight drop in the estimation error of the asset value in the GARCH model between the high financial risk, low business risk case and the high financial risk, high business risk case. The sample standard deviations of the estimation errors for all three variables are lower in the high financial risk, high business risk case for the Merton and Duan models than for the low business risk, high financial risk case, but otherwise are increasing in business and financial risk.

We now turn to an empirical application of the GARCH model to US banks and financial companies during the 2007-2008 credit crunch. Several of the results derived from our simulation study will be useful in understanding the output of the three models on the bank data. In particular, our simulations show that when the true asset process exhibits constant volatility, and that volatility and/or financial leverage is high, the results of the GARCH and Duan models will be similar on average, and both will tend to outperform the Merton model. This generalization seems to fit rather well the pattern we see in our on spread estimates from the three models for banks and financial companies with moderate or high observed CDS spread levels. Also, the variation in the model estimated spreads tends to be much higher, especially for Duan and GARCH, for the banks and financial companies with high CDS spreads, which are more likely to have high financial and/or business risk.

IV. EMPIRICAL APPLICATION: THE CDS MARKET FOR US BANKS AND FINANCIAL COMPANIES DURING THE 2007-08 CREDIT CRISIS

In this section, we turn our attention to the application of our GARCH structural credit risk model to the out-of-sample prediction of CDS spreads for US banks and financial companies during the 2007-2008 period. In particular, we will compare the success of the calibrated Merton model, the Duan model, and the GARCH model in terms of their average prediction errors of CDS spreads at the 1 year, 3 year, and 5 year time horizons for the US banks and financial companies for which suitable balance sheet and spread data is available. Our data provider, unless otherwise stated, is Bloomberg. Consistent with our simulation study, we will focus on out-of-sample prediction of the models, using a 52 week window of data for the estimation, for the cross section of banks on two dates of particular interest: December 14th, 2007, and August 29th, 2008. The former date, which occurred approximately five months after the revelation of some of the first widely recognized signs of the crisis in July of 2007, corresponds roughly to the high point of the US equity markets before their sustained drift downward that has continued up until the time of writing. The latter date corresponds to two weeks before the investment bank Lehman Brothers declared bankruptcy, after it became clear that a rescue package by the US government was not forthcoming.

The CDS spreads we use in our analysis are those linked to the senior debt of the banks and financial companies in question, as we only assume a two-layered liability structure in applying our models, as is conventional in the literature. To build our sample of firms, we identified all banks and financial companies domiciled in the United States at the time of writing. We then selected those firms with senior CDS spreads available in at least one of the 1 year, 3 year, and 5 year categories. Of these firms, we kept those with at least one year (52 weeks) of uninterrupted balance sheet data (short and long term debt, and market equity) necessary for running the three models at the two dates of interest. Due to the fact that not all firms had even coverage in the three different CDS maturities chosen, we obtain different final numbers of firms available for running the models for each maturity on our two dates of interest. Additionally, post-estimation, we made a distinction between firms for which the largest of the three estimated model spreads was greater than 0.1 basis points, and those for which this was not the case. The total sample size pre-estimation, and the sample size of each of the aforementioned groups post-estimation, is listed in Table II for the six cases corresponding to the two dates and three CDS maturities we consider.

[INSERT TABLE II HERE]

As is apparent, there are more firms with available CDS data for all three maturities in August of 2008 than in December of 2007, although not all firms have spread data for all three maturities on either date. It appears that the market began offering CDS quotes for shorter maturities for several banks and financial companies by late 2008 due to the increases in the likelihood of default for several institutions previously considered to be remote default risks. The average CDS spread in our sample, before the exclusion of banks with near-zero model spreads, on December 14, 2007 was 166.5 basis points for the 1 year CDS market, with a sample standard deviation of 184.3 basis points. The average CDS spread in the sample before exclusion of banks with near-zero model spreads on August 29, 2008 was 297.3 basis points, with a sample standard deviation of 423.8 basis points. On December 14, 2007, the 1 year CDS spreads in our sample ranged from a low of 16.079 basis points, for Bank of America, to a high of 665.872 basis points, for MBIA Inc.. On August 29, 2008, the 1 year CDS spreads in the sample ranged from a low of 5.69 basis points for LOEWS

Corporation, to a high of 2335.905 basis points for Washington Mutual. The pattern of average spreads and sample standard deviations for the 3 year and 5 year CDS maturities between these two dates is similar to the pattern for the one year maturities: both average spreads and the sample standard deviation of spreads, as well as the range of spread value, increases from December 2007 to August 2008. The full summary statistics, besides those reported in this paper, are available upon request. One other empirical feature of the CDS market that stands out, and is in contrast to the pattern of the market in the five years leading up to 2007 and 2008, is that the term structure of average spreads is U-shaped in December of 2007, and fully inverted in August of 2008. This is the result of the fact that the individual spread term structures for many banks are humped (U-shaped upward or downward) or fully inverted on the two dates we study. Besides the fact that inversion of the term structure of spreads appears to be a leading indicator of the deterioration of credit quality of banks, as evidenced by an increase in spreads across all maturities, this issue also has important implications for the structural credit risk models we test in this paper. This issue deserves special attention, and we will treat it in detail in subsection A that follows.

The results of running the three models on December 14, 2007 and August 29, 2008 using the relevant 52 week data windows in each case are shown in Tables III and IV, respectively. Each table reports the root mean squared error (RMSE), the mean absolute deviation (MAD), and the constant (α) and coefficient (β) terms of a standard OLS regression of actual CDS spreads on model spreads, along with the R^2 of the regression, with the standard errors displayed in parenthesis below the point estimates of α and β. Statistical significance of a coefficient in one of the linear regressions is denoted using three stars for significance at the 1% level, two stars for significance at the 5% level, and one star for significance at the 10% level, with the stars placed alongside the standard error of the relevant coefficient. In the upper panel of each table, we present the statistics obtained using the sample of banks whose largest post-estimation model spread was greater than 0.1 basis points. The rationale for this post-estimation selection is that it is valid to exclude banks whose model results uniformly indicate some sort of serious model mis-specification, and the inability of any of our three models to generate (effectively) nonzero spreads provides a good reason to exclude these banks from the sample. To check that the inability of the models we tested to generate non-zero spreads was indeed due to differences in the characteristics of the equity time series for the group of excluded banks vs. the group of banks included in the summary statistics of Tables III and IV, we computed the annualized equity volatility for the banks included in the August 29, 2008 sample and the banks with available data on that date that were excluded from our summary statistics due to the imposition of the 0.1 minimum spread criterion. The included banks had an average annualized equity volatility of 45.3%, versus an average annual equity volatility of only 6.5% for the excluded banks, and a t-test of difference in means rejects the null hypothesis of equality of equity volatility between the two groups at the 1% level. Thus, it is safe to conclude that the equity time series of the banks excluded from our summary statistics simply did not display enough volatility to generate nonzero spreads in our models, and tax considerations or other sorts of effects, such as jump-risk in the asset value, which are outside of our models sustain the positive and generally low spreads we observe for those banks in the CDS market.

In the sample of banks that we display summary statistics for, there remains a legitimate concern of model misspecification in the case of the Federal National Mortgage Association, commonly known as Fannie Mae, and the Federal Home Loan Mortgage Association, commonly known as Freddie Mac, because both organizations have an implicit and very public guarantee from the government to cover a large portion of their expected losses in situations that would ordinarily provoke default in a private corporation. For this reason, Tables III and IV display summary statistics for the sample discussed above, first including, and then excluding, Fannie Mae and Freddie Mac from the sample.

[INSERT TABLE III HERE]

[INSERT TABLE IV HERE]

Several patterns emerge from the results. First, the lowest out-of-sample RMSE for the models on December 14, 2007 is achieved by the GARCH model for the 5 year CDS maturity, when Fannie Mae and Freddie Mac are included in the sample. There is no data recorded for Fannie and Freddie at the 1 year and 3 year maturities in December 2007, so there are no statistics to report for those cases. When Fannie Mae and Freddie Mac are excluded from the sample on December 14, 2007, the lowest RMSE is achieved by the Merton model for the 1 year CDS market, and the GARCH model for the 3 year and 5 year CDS markets. Turning to the second date of interest, on August 29, 2008 the lowest RMSE is obtained by the Merton model for the 1 year and 3 year maturities, and by the Duan model for the 5 year maturity, when Fannie Mae and Freddie Mac are included in the sample. After excluding Fannie and Freddie, however, the lowest RMSE is achieved by the GARCH model for all three maturities in August of 2008. The ranking of the models according to the lowest MAD is the same at that obtained using RMSE in all cases.

To summarize, the GARCH model achieves the lowest out-of-sample RMSE and MAD of the three models considered in six out of the ten cases considered, and achieves the lowest RMSE and MAD in five out of the six cases in which Fannie Mae and Freddie Mac were excluded from the sample. The high success rate of the GARCH model in the sample excluding Fannie Mae and Freddie Mac is probably the more relevant yardstick of success in absolute prediction, given that the market spreads of Fannie Mae and Freddie Mac, which are in the range of 30-38 basis points over the different CDS maturities, are far below the estimated model spreads, which are over 3000 basis points for the Duan and Merton models at the 1 year maturity in August of 2008, and reflect the high value of the government guarantee to these two entities.

Moving on to the OLS regression results, we find that the beta coefficient of actual on model spreads is significant at the 1% level for the Merton model in December of 2007, and insignificant for the Duan and GARCH models, over all five cases considered on that date. The regression of actual on predicted spreads for the Merton model also display significantly higher R^2 values than for the Duan and GARCH models in December of 2007. In August of 2008, however, the beta coefficients for the Duan and GARCH models are also significant at the 1% level at the 1 year maturity in the sample that includes Fannie Mae and Freddie Mac, and significant at either the 1% or the 5% level for all three maturities in the sample with Fannie Mae and Freddie Mac excluded. In addition, the R^2 values are higher for the Duan and GARCH models than the Merton model for the regressions at the 1 year maturity in August of 2008, and although they are lower at the 3 year and 5 year maturities, they are both near 26% and 22% at those maturities, respectively, compared to R^2 values for the Merton model of 42% and 45% at the 3 year and 5 year maturities, respectively.

The overall conclusions from our analysis of the model performance statistics can be summed up as follows. The GARCH model proposed in this paper appears to be the clear winner, compared to the Merton and Duan models, in terms of success measured by lowest absolute prediction errors, with either the RMSE or MAD criteria, and this is especially true in the samples that exclude Fannie Mae and Freddie Mac, for which model misspecification due to the existence of a government guarantee is obviously a serious issue. The calibrated Merton model, however, displays a surprising success at explaining the variation of CDS spreads in the cross section of banks at different maturities in the CDS market, although this success is more apparent in the comparison of the models in December of 2007 than in the more turbulent period of August 2008, in which the average levels and variation of CDS spreads was higher. The latter finding is consistent with the general conception among practitioners that the calibrated Merton model, while known to produce very low spreads, has the ability do adapt to diverse forms of model misspecification in practice. Finally, the relatively similar performance of the Duan and GARCH models along all indicators and on both dates of our sample, which appear to follow a different pattern than is the case with the

Merton model, should be compared to the results of our simulation study in the previous section. Those results indicated that the Duan and GARCH models tend to follow the pattern indicated here in scenarios of high business and financial risk, but constant volatility. The Merton model performs better on real data than in our simulation study in such conditions, and we mark that result up to the reason just stated, that the Merton model tends to adapt well to some types of unobserved model misspecification. In the samples excluding Fannie Mae, and Freddie Mac, however, the results seem to suggest a common scenario among many banks of high business and financial risk, but relatively stable (if not constant) volatility.

Although the analysis of prediction errors within maturity categories is important, also revealing is an examination of the information contained in the term structure of spreads, or the pattern of spreads across maturities. We now turn to an analysis of the spread term structure in the CDS data and our models during the 2007-2008 crisis period.

A. **The CDS and Model Spread Term Structures in 2007 and 2008.** In order to better understand the patterns displayed in the shape of bank CDS term structures, we classified the shape of the term structure for each bank and financial company in our sample in December of 2007 and August of 2008. These statistics are reported in Tables V and VI, respectively. Divisions of the bank spread term structures are made into four groups: those that are upward sloping (U), in which the 5 year spread is greater than the 3 year spread, and the 3 year spread is greater than the 1 year spread; those that are humped (H), in which the three year spread is below both the 1 year and 5 year spreads, or above both the 1 year and 5 year spreads; those that are downward sloping (D), in which the spread decreases as a function of the maturity; and those that are flat (F), in which all spreads are equal. In practice, the classification of spread curves into the "F" category only applies to the model spread curves in cases where all three model spreads were equal to zero. There are no flat spread curves in the CDS market on either date.

In each table, we divide the group of banks within each category (market spreads and model spreads for the three models) into two groups, corresponding to the nonzero model spread subset and the zero model spread subset, as identified on the basis of whether the largest model spread at the 5 year maturity was greater than or less than 0.1 basis points on the relevant date. We report statistics for the full sample, before excluding banks based on zero model spreads, and for each of these two subgroups. In addition, although not reported in the text, we identified the banks that were present in both the December 14, 2007 sample and the August 29, 2008 sample, and summarized the categorization exercise among this subset of banks.

[INSERT TABLE V HERE]

[INSERT TABLE VI HERE]

Several important patterns emerge from the data and the model results. Perhaps the most striking fact we observe is that all three models, for both dates studied, produce downward sloping spread curves for every bank in the sample. This is quite remarkable, as it is known (see e.g. Rouah and Vainberg, 2007) that the only way for the Merton and Duan models in particular to generate downward sloping spread term structures is in the presence of leverage ratios $K/V > 1$. This suggests, at the least, high degrees of leverage among many banks in the sample. Two clarifications of this finding warrant mention. In the first place, the downward sloping spread term structure in the subsample of excluded banks is not of primary interest, since these spread curves are nearly flat in any case and very close to zero when they are not equal to zero. In the second place, for the sample of non-excluded banks, the inversion of the nontrivial spread term structures derived in this case for the models is actually quite consistent with the actual CDS market data, to which we will now turn.

In the CDS market data, we find that 83% of the banks in the non-excluded subsample have humped or downward sloping term structures in December of 2007, and this remains true for 57% of banks in the non-excluded subsample in August of 2008. The apparent drop in the degree of non-upward sloping term structures from 2007 to 2008, however, is masked by the fact that there were a significant number of new entrants into the CDS market for banks between the two dates, and these new entrants were much more likely to have standard, upward sloping spread curves than the banks that are common to both samples [1]. We infer that this is due to the fact that the market began demanding CDS contracts for several banks that, even in late 2007, were considered too remote as candidates for a default to merit a liquid market in CDS contracts. The higher credit quality of the new entrants is consistent with the upward sloping tendency of their spread term structures. In the full sample of CDS market spreads that includes both subgroups, we find that 61% of banks in December of 2007, and 33% of banks in August of 2008, had humped or downward sloping spread term structures. These figures are less than for the group of banks with nontrivial model spreads post-estimation, and are consistent with the fact that, besides having lower equity volatility and lower average spreads, the banks that were excluded according to our post estimation criterion of near-zero spreads were also less like to have inverted term structures.

[INSERT TABLE VII HERE]

To formally test the above observations, we performed a t-test of difference in proportions (and means where appropriate) of upward sloping spread term structures in the data across the categories of interest. These tests and the null hypotheses on which they are based are summarized in Table

[1] Only one bank, MBIA, was in the December 2007 sample but not in the August 2008 sample.

VII. The null hypothesis of equality in the proportion of banks with upward sloping term structures in the full sample in 2007 versus in 2008 was rejected at the 5% level, although we could not reject equality of this proportion in either subsample, possibly due to modest sample sizes. We reject equality of the proportion of banks with upward sloping spread term structures between the included and excluded groups (nonzero vs. zero spreads) at the 1% level during 2007, and again at the 1% level during 2008, with the excluded groups having significantly more upward sloping spread term structures. In order to verify that the increase in the proportion of banks with upward sloping term structures in the full sample was indeed due to the entry of several, low spread banks with upward sloping term structures, we performed a paired test of difference in means of the indicator variable equal to 1 if the term structure is upward sloping, and zero otherwise, among the group of banks common to both the December 2007 sample and the August 2008 sample. We cannot reject the null hypothesis of no difference in means, so we can make the natural conclusion that the difference arises from the additional banks that entered the CDS market in 2008. While the incidence spread term structure inversion among the incumbent banks in the CDS market was nearly constant between December 2007 and August 2008, the level and dispersion of spread levels at all three maturities increased noticeably for these banks between the two dates studied.

Overall, the empirical evidence from the CDS market of a significant incidence of spread term structure inversion, combined with our model results of spread term structure inversion for all banks in the sample, point strongly toward high degrees of leverage and asset volatility for the banks and financial companies active in the CDS market, and this is true in particular for the incumbent banks common to the 2007 and 2008 samples. The latter list of banks includes American Express, American International Group, Bank of America, Capital One, Citigroup, Goldman Sachs, JP Morgan Chase, Lehman Brothers, Merrill Lynch, Morgan Stanley, and Washington Mutual, among names that have appeared frequently in the news during the 2007-2009 period. The models tested are able to capture, for the most part, the shape of the spread term structure for these banks, and in most cases, especially in August of 2008, our GARCH model is able to capture the levels of the spreads across the maturity structure better than the Merton or Duan models tested.

V. CONCLUSION

In this paper, we develop a structural credit risk model in which the underlying asset follows a GARCH process, as in Heston and Nandi (2000). We show how to estimate the parameters of the model using observed equity time series data for the firm, using an Expectation Maximization algorithm, as in the work of Duan (1994). We perform a simulation study to benchmark the relative performance of the calibrated Merton model, the Duan model, and our GARCH model under different scenarios for leverage and asset volatility, both in a setting with and a setting without stochastic volatility in the data generating process. We find that the GARCH model significantly outperforms the Duan and Merton models in a setting with stochastic asset volatility of the GARCH variety, and that in a setting with constant asset volatility, the out-of-sample performance of Duan and GARCH are similar, with Duan slightly outperforming GARCH in out-of-sample prediction of the true asset level, asset volatility, and the spread, and both significantly outperforming Merton.

We then conduct an empirical application of our model in the context of the CDS market for banks and financial companies in the US on two dates, December 14, 2007, and August 29, 2008. The former date corresponds roughly to the high point of the US equity markets before they began a period of sustained negative returns, and the latter date corresponds to two weeks before the collapse of the investment bank Lehman Brothers, when it formally declared bankruptcy. We find evidence of a significant government guarantee to the national mortgage agencies Freddie Mac and Fannie Mae, in the form of a government promise to cover a large portion of the expected loss of those institutions, along the lines described in Gray and Malone (2008), as GARCH and Duan model spreads are much greater than the CDS market spreads for these institutions. Moreover, we find that the GARCH model achieves the lowest out-of-sample prediction errors of the three models in the various samples considered, and that this is overwhelmingly true in the cases where we exclude Fannie Mae and Freddie Mac from the sample, both for December 2007 and August 2008. There is moderate evidence of stochastic volatility, but not strong evidence, as the performance of the Duan model in the latter cases does not lag significantly behind the performance of the GARCH model in terms of absolute prediction errors.

We find that all three of our models predict downward sloping spread term structures for all banks in our sample on both dates. We examine the empirical evidence from the CDS market on spread term structures across the banks, and find that a high proportion of banks and financial companies display downward sloping spread term structures in both December of 2007 and August of 2008. This is especially true of the incumbent banks that are present in both samples. The latter list includes the American International Group, Lehman Brothers, Morgan Stanley, Goldman Sachs, and a range of other banks that have featured prominently in the financial press during the financial crisis that began in 2007. The group of banks for which the structural credit risk models tested are able to generate nontrivial spread values across the maturity spectrum display a significantly higher average annualized equity volatility, higher average spreads for all maturities, and a higher incidence of spread term structure inversion than the banks for which the models are unable to generate nontrivial (different from zero) spreads. Duan, Gauthier, and Simonato (2004) have shown that the method used by Moody's-KMV to estimate their structural credit risk model is equivalent to the maximum likelihood method that we use in this paper under the "Duan" model specification. If those results are true, then our results of estimating the Duan model on December 14, 2007 indicate that the credit risk models used by Moody's KMV should, by all indications, have given overwhelming evidence of significant weakness in the credit quality of most major investment banks already at that time. This calls into question the decision of the rating agencies to delay the downgrading of many investment banks during the period studied. Our results for the Merton and GARCH models tell a similar story. The models are also generally able to match the rise in spreads from 2007 to 2008, as the credit crisis worsened, and the Duan and GARCH model spreads for Lehman Brothers two weeks before its collapse are noticeably above CDS market spreads, as consistent with the expectation by the market of a government bailout that never materialized.

VI. APPENDIX: PROOF OF CONVERGENCE OF EM ALGORITHM

Given initial values for $\mathbf{S^{(0)}}$, $\eta^{\mathbf{(0)}}$ and $\eta^{*\mathbf{(0)}}$, the following two-step iterative algorithm converges to the maximum likelihood estimator of η and $\mathbf{S}$.

(1) Update the values of the underlying asset $\mathbf{S}$ from the observed prices of the equity as $\mathbf{S^{(n+1)}} = \mathbf{g^{-1}}(\mathbf{C}|\eta^{*\mathbf{(n)}})$.
(2) Update the values of the model parameters by setting $\eta^{\mathbf{(n+1)}} = \hat{\eta}(\mathbf{S^{(n+1)}})$ and construct the new set of risk-neutralized parameters $\eta^{*\mathbf{(n+1)}}$ from them.

Proof. First we note that, since g is invertible and the value of assets is obtained deterministically from the value of the equity, therefore,

$$p(\mathbf{S}, \mathbf{C}|\eta) = p(\mathbf{S}|\eta)\delta_{\mathbf{g}(\mathbf{S}|\eta^*)}(\mathbf{C})$$

$$p(\mathbf{S}|\mathbf{C}, \eta) = \delta_{g^{-1}(\mathbf{C}|\eta^*)}(\mathbf{S})$$

where $\delta_{\boldsymbol{\theta}}(\cdot)$ denotes the degenerate measure putting all mass on $\boldsymbol{\theta}$. These functions satisfy the requirements in **?**. Therefore the E step of the EM algorithm reduces to

$$\mathbb{E}_{\mathbf{S}|\mathbf{C},\eta^{(\mathbf{n})}}(\log[p(\mathbf{S}, \mathbf{C}|\eta)]) = \log(\mathbf{p}(\mathbf{g^{-1}}(\mathbf{C}|\eta^{*\mathbf{(n)}})|\eta))$$

while the maximization step solves,

$$\eta^{(\mathbf{n+1})} = \arg\max_{\eta} \log(\mathbf{p}(\mathbf{g}^{-\mathbf{1}}(\mathbf{C}|\eta^{*(\mathbf{n})})|\eta))$$

Denoting $g^{-1}(\mathbf{C}|\eta^{*(\mathbf{n})}) = \mathbf{S}^{(\mathbf{n+1})}$, we see that the expectation step correspond to step 1 above, while the maximization step corresponds to step 2. □

REFERENCES

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*, 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **31**, 307–327.

Christoffersen, P. and K. Jacobs (1995). Which garch model for option valuation? *Management Science 50*, 1204–1221.

Duan, J. (1994). Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance 4*, 155–167.

Duan, J., G. Gauthier, and J. Simonato (2004). On the equivalence of the kmv and maximum likelihood methods for structural credit risk models.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical Finance 5*, 13–32.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica* **50**, 987–1008.

Eom, Y., J. Helwege, and J.-Z. Huang (2004). Structural models of corporate bond pricing: An empirical analysis. *The Review of Financial Studies 17*(2), 499–544.

Ericsson, J. and J. Reneby (2005). Estimating structural bond pricing models. *Journal of Business 78*(2).

Gray, D. and S. Malone (2008). *Macrofinancial Risk Analysis*. John Wiley and Sons.

Heston, S. L. and S. Nandi (2000). A closed-form garch option valuation model. *The Review of Financial Studies* **13**, 585625.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of business and Economic Statistics* **12**, 371–389.

Merton, R. (1973). The theory of rational option pricing. *Bell Journal of Economics and Management Science 4*, 141–183.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance 29*, 449–470.

Nicolato, E. and E. Venardos (2003). Option pricing in stochastic volatility models of the ornstein-uhlenbeck type. *Mathematical Finance* **13**(4), 445–466.

Rouah, F. and G. Vainberg (2007). *Option Pricing Models and Volatility Using Excel-VBA*. London: John Wiley & Sons.

TABLE I. Simulation Results: Out-of-Sample Mean Average Estimation Errors and Standard Deviations for the Asset, Asset Volatility, and Spread by DGP and Scenario

Data Generating Process: Constant Volatility

Scenario	Business Risk	Low		Low		High		High	
	Financial Risk	Low		High		Low		High	
	Model	MAE	St. Dev.	MAE	St. Dev.	MAE	St. Dev.	MAE	St. Dev.
Asset	Merton	0.01	0.04	5.51	7.10	0.67	1.40	9.24	8.85
	Duan	0.01	0.02	2.09	4.15	0.27	0.51	3.89	3.91
	GARCH	0.01	0.02	2.08	4.25	0.29	0.51	3.93	3.91
Asset Volatility	Merton	0.02	0.02	0.11	0.08	0.07	0.06	0.17	0.12
	Duan	0.02	0.01	0.04	0.04	0.04	0.03	0.07	0.05
	GARCH	0.02	0.02	0.06	0.06	0.05	0.04	0.10	0.08
Spread	Merton	1.65	8.96	666.63	985.67	150.97	339.31	1208.60	1286.50
	Duan	1.01	4.64	275.35	700.97	61.45	123.56	634.76	602.86
	GARCH	1.04	4.97	274.44	706.97	69.64	138.09	652.23	701.13

Data Generating Process: Stochastic Volatility (GARCH)

Scenario	Business Risk	Low		Low		High		High	
	Financial Risk	Low		High		Low		High	
	Model	MAE	St. Dev.	MAE	St. Dev.	MAE	St. Dev.	MAE	St. Dev.
Asset	Merton	0.03	0.19	20.04	25.43	0.96	2.56	22.66	21.80
	Duan	0.08	0.65	15.56	25.29	0.89	2.16	18.84	21.44
	GARCH	0.02	0.14	7.13	17.84	0.70	1.76	13.34	15.54
Asset Volatility	Merton	0.05	0.05	0.53	0.95	0.11	0.12	0.53	0.71
	Duan	0.04	0.03	0.45	1.10	0.06	0.07	0.42	0.69
	GARCH	0.03	0.03	0.12	0.09	0.03	0.07	0.13	0.11
Spread	Merton	6.94	42.84	5272.00	13173.00	232.78	702.59	5265.60	9256.70
	Duan	18.58	165.16	5836.00	18039.00	216.09	566.61	5736.20	11038.00
	GARCH	5.19	36.51	1120.00	1765.00	177.15	310.12	1231.00	8379.00

TABLE II. Sample Sizes by Date, CDS Maturity, and Minimum Model Spread Criterion

Date	CDS Maturity	Total Firms	Firms with Largest Model Spread > 0.1bps	Remaining Firms
December 14, 2007	1 YR	18	13	5
	3 YR	20	12	8
	5 YR	38	18	20
August 29, 2008	1 YR	44	25	19
	3 YR	43	22	21
	5 YR	43	21	22

TABLE III. Out-of-Sample Model Performance Statistics for December 14, 2007

Sample Including Fannie and Freddie

Maturity	Model	RMSE	MAD	Linear Fit: Actual vs. Predicted α	β	R^2
1 YR	NO CDS DATA FOR FANNIE AND FREDDIE					
3 YR	NO CDS DATA FOR FANNIE AND FREDDIE					
5 YR	Merton	257.026931	206.7639745	174.7436 (31.38266)***	18.62909 (5.103533)***	0.4876
	Duan	207.6202076	155.9718759	174.5067 (53.48686)***	0.4545662 (0.4656228)	0.0637
	GARCH	198.6523108	147.0013663	156.5545 (49.31121)***	0.709845 (0.4247118)	0.1663

Sample Excluding Fannie and Freddie

Maturity	Model	RMSE	MAD	Linear Fit: Actual vs. Predicted α	β	R^2
1YR	Merton	281.8805605	202.6135752	142.5147 (41.6422)***	59.71004 (15.27127)***	0.5816
	Duan	409.0171236	317.9051205	160.6056 (78.72125)*	0.1239035 (0.1533487)	0.056
	GARCH	398.0652861	286.6306006	154.3658 (73.52279)*	0.1537942 (0.1455263)	0.0922
3 YR	Merton	254.365733	196.9580918	148.2823 (38.8047)***	121.363 (37.49189)***	0.5117
	Duan	207.0962884	146.1214077	175.5613 (71.28697)**	0.1583166 (0.3643509)	0.0185
	GARCH	200.1160847	137.2951725	161.0189 (65.72933)**	0.2860469 (0.341177)	0.0657
5 YR	Merton	242.6390312	187.8885624	156.905 (30.41203)***	19.08448 (5.238632)***	0.4534
	Duan	203.1174645	156.6709989	179.4102 (54.78813)***	0.1146979 (0.4305881)	0.0044
	GARCH	199.5533095	153.5283752	169.1292 (52.78617)***	0.2220318 (0.3869252)	0.0202

TABLE IV. Out-of-Sample Model Performance Statistics for August 29, 2008

Sample Including Fannie and Freddie

				Linear Fit: Actual vs. Predicted		
Maturity	Model	RMSE	MAD	α	β	R^2
1YR	Merton	1368.741289	626.274176	426.8755 (101.3533)***	0.1992522 (0.0613489)***	0.3144
	Duan	1475.492397	889.7348733	306.6601 (120.659)**	0.1910055 (0.063326)***	0.2834
	GARCH	1439.531243	875.4666627	307.6148 (123.4855)**	0.1907165 (0.0664109)***	0.2639
3 YR	Merton	493.2605128	366.8081911	289.487 (70.49514)***	3.3154 (0.8954331)***	0.4067
	Duan	521.3732462	412.0604218	339.4894 (100.1858)***	0.1829685 (0.1785674)	0.0499
	GARCH	526.6544187	413.7504248	342.1013 (100.893)***	0.1718464 (0.1778754)	0.0446
5 YR	Merton	447.4713548	345.6810508	270.1861 (60.35925)***	4.36686 (1.168367)***	0.4237
	Duan	416.7490419	353.3113108	326.9397 (88.53717)***	0.1831354 (0.2380441)	0.0302
	GARCH	419.4579428	354.519464	329.6925 (89.14734)***	0.1677114 (0.2370648)	0.0257

Sample Excluding Fannie and Freddie

				Linear Fit: Actual vs. Predicted		
Maturity	Model	RMSE	MAD	α	β	R^2
1 YR	Merton	1426.95758	677.4801022	464.3475 (107.1711)***	0.1942818 (0.06225)***	0.317
	Duan	1231.938184	695.4688065	317.1863 (102.5303)***	0.2662314 (0.058419)***	0.4972
	GARCH	1160.722433	671.0410527	308.5794 (103.4689)***	0.2801066 (0.0614485)***	0.4974
3 YR	Merton	517.294756	401.9878225	324.1077 (72.9148)***	3.228923 (0.8862642)***	0.4244
	Duan	421.4431171	343.099729	316.8389 (89.01308)***	0.4774088 (0.1908653)**	0.2579
	GARCH	418.7563149	341.2201094	314.5062 (89.61333)***	0.4806344 (0.1925748)**	0.2571
5 YR	Merton	470.3434772	379.2148069	303.0883 (61.84275)***	4.229056 (1.142761)***	0.4462
	Duan	375.3198153	317.1718033	307.0698 (79.24523)***	0.5596253 (0.2557823)**	0.2197
	GARCH	373.7791812	315.918875	305.4637 (79.87288)***	0.5620542 (0.258427)**	0.2177

TABLE V. Summary Statistics for the Shape of the Spread Curve: CDS Market and Models on December 14, 2007.

			U	H	D	F	Total
CDS Market	FULL SAMPLE	Count	7	7	4	0	18
		Percentage	39 %	39 %	22 %	0 %	100 %
	NON ZERO SPREAD SUBSET	Count	2	7	3	0	12
		Percentage	17 %	58 %	25 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	5	0	1	0	6
		Percentage	83 %	0 %	17 %	0 %	100 %
Merton	FULL SAMPLE	Count	0	0	18	0	18
		Percentage	0 %	0 %	100 %	0 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	3	0	3
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	0	15	0	15
		Percentage	0 %	0 %	100 %	0 %	100 %
Duan	FULL SAMPLE	Count	0	0	18	0	18
		Percentage	0 %	0 %	100 %	0 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	10	0	10
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	0	8	0	8
		Percentage	0 %	0 %	100 %	0 %	100 %
GARCH	FULL SAMPLE	Count	0	0	14	4	18
		Percentage	0 %	0 %	78 %	22 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	12	0	12
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	0	2	4	6
		Percentage	0 %	0 %	33 %	67 %	100 %

Key: U = Upward Sloping. H = Humped Upward or Humped Downward. D = Downward Sloping. F = Flat (all zeros)

*The "nonzero spread" subsamples are constructed using the post-estimation criterion that 5 YR spreads in the relevant category (CDS Market, Merton, Duan, and GARCH, resp.) are greater than 0.1bps.

TABLE VI. Summary Statistics for the Shape of the Spread Curve: CDS Market and Models on August 29, 2008

			U	H	D	F	Total
CDS Market	FULL SAMPLE	Count	29	2	12	0	43
		Percentage	67 %	5 %	28 %	0 %	100 %
	NON ZERO SPREAD SUBSET	Count	9	1	11	0	21
		Percentage	43 %	5 %	52 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	20	1	1	0	22
		Percentage	91 %	4.5 %	4.5 %	0 %	100 %
Merton	FULL SAMPLE	Count	0	2	35	6	43
		Percentage	0 %	5 %	81 %	14 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	17	0	17
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	2	18	6	26
		Percentage	0 %	8 %	69 %	23 %	100 %
Duan	FULL SAMPLE	Count	0	0	37	6	43
		Percentage	0 %	0 %	86 %	14 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	19	0	19
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	0	18	6	24
		Percentage	0 %	0 %	75 %	25 %	100 %
GARCH	FULL SAMPLE	Count	0	0	26	17	43
		Percentage	0 %	0 %	60 %	40 %	100 %
	NON ZERO SPREAD SUBSET	Count	0	0	21	0	21
		Percentage	0 %	0 %	100 %	0 %	100 %
	ZERO SPREAD SUBSET	Count	0	0	5	17	22
		Percentage	0 %	0 %	23 %	77 %	100 %

Key: U = Upward Sloping. H = Humped Upward or Humped Downward. D = Downward Sloping. F = Flat (all zeros)

*The "nonzero spread" subsamples are constructed using the post-estimation criterion that 5 YR spreads in the relevant category (CDS Market, Merton, Duan, and GARCH, resp.) are greater than 0.1bps.

TABLE VII. Classical Hypothesis Tests of Differences in Shapes of Spread Term Structures Among Dates and Subsamples

Null Hypothesis	Test Statistic	Test Statistic Value	p-value
H_0: $\theta^{full}_{U2007} = \theta^{full}_{U2008}$	z	-2.03	0.0428 **
H_0: $\theta^{nonzero}_{U2007} = \theta^{nonzero}_{U2008}$	z	-1.52	0.1281
H_0: $\theta^{zero}_{U2007} = \theta^{zero}_{U2008}$	z	-0.56	0.5744
H_0: $\theta^{nonzero}_{U2007} = \theta^{zero}_{U2007}$	z	-2.71	0.0068 ***
H_0: $\theta^{nonzero}_{U2008} = \theta^{zero}_{U2008}$	z	-3.36	0.001 ***
H_0: $I_{U2008} = I_{U2007}$	t	-0.5657	0.5795

Key: θ^{SAMPLE}_{UYEAR} = The proportion of upward sloping spread term structures in subsample SAMPLE in year YEAR.

The "nonzero" and "zero" subsamples are as defined in the text, according to the postestimation criterion.

I_{UYEAR} = An indicator variable equal to 1 if the term structure is upward sloping, zero otherwise. These indicator variables were used in the t-test of a difference in means in a paired sample in the banks common to both years.