Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

CONNECTING THE TOP-DOWN TO THE BOTTOM-UP: PRICING CDO UNDER A CONDITIONAL SURVIVAL (CS) MODEL

Xian Hua Peng
Steven S. G. Kou

312 S. W. Mudd Building
500 West 120th Street
Columbia University
New York, NY 10027, U.S.A.

ABSTRACT

In this paper, we use exact simulation to price CDO under a new dynamic model, the Conditional Survival (CS) model, which provided excellent calibration to both iTraxx tranches and underlying single name CDS spreads on March 14, 2008, the day before the collapse of Bear Sterns, when the market was highly volatile. The distinct features of the CS model include: (1) it is able to generate clustering of defaults occurring dynamically in time and strong cross-sectional correlation, i.e., the simultaneous defaults of many names, both of which have been evidenced in the current subprime mortgage crisis; (2) it incorporates idiosyncratic default risk of single names but does not specify concrete models for them; (3) it provides automatic calibration to underlying single name CDS; (4) it allows fast CDO tranche pricing and calculation of sensitivity of CDO tranches to underlying single name CDS.

1 INTRODUCTION

1.1 CDO and Current Subprime Mortgage Crisis

Collateralized debt obligation (CDO) is a derivative security constructed from a portfolio of fixed-income securities or credit derivatives. The stream of cash flow of a CDO is determined by the cumulative loss of the reference portfolio. A CDO is traded in terms of tranches, each of which corresponds to one specific portion of the portfolio loss. An investor of a CDO tranche receives periodic coupons and in return pays the portion of default losses that correspond to that tranche. For example, an investor of a $[3\%, 6\%]$ CDO tranche assumes the portfolio default losses that fall in the range from 3% to 6% of the portfolio notional.

CDO provides a way of creating securities with high credit quality out of a portfolio of securities with low credit quality. Senior CDO tranches rarely assume any losses unless a substantial number of names in the portfolio default. Therefore, senior tranches are usually safer and have higher credit ratings than the average names in the portfolio.

CDO has played an important role in the current subprime mortgage crisis started in 2007. In the last few years, financial institutions pooled and packaged mortgage-backed securities into CDO and sold them to third-party investors. This resulted in widespread dispersion of default risk. The declining of house prices and rising defaults of mortgage borrowers have caused huge loss to a wide range of investors, leading to credit crunch and turmoils in the financial markets.

However, as merely an instrument, CDO should not be held responsible for the financial crisis. The credit rating agencies have been blamed for their failure of evaluating CDO correctly. During the subprime mortgage crisis, senior tranches of CDO with AAA rating caused huge loss to investors. The difficulty of CDO pricing lies in the appropriate modeling of default correlation among the underlying names in the portfolio. Although the marginal default probability of single names can usually be implied from single name CDS spreads, the pricing of CDO requires the knowledge of the joint distribution of default times of underlying names, which can not be directly implied from the market.

1.2 Review of Extant CDO Pricing Models

In the rest of the paper, we assume that there are n names in the reference portfolio underlying the CDO and use τ_i to denote the default time of the i-th name, $i = 1, \ldots, n$.

The two main kinds of approaches in portfolio credit risk modeling are the bottom-up approach, which builds models for the correlated default times of single names in the portfolio; and the top-down approach, which builds models for the cumulative loss of the whole portfolio without referring to the underlying single names.

Bottom-up models can be separated into static models and dynamic models. Static models view the default times of the underlying names as static random variables and model their joint distribution directly. The Gaussian copula model proposed by Li (2000) used to be the most popular model in industry (see also Andersen et al. 2003 and Schönbucher 2003). A major problem is that the Gaussian copula model cannot generate tail dependence. More precisely, if the joint distribution of two default times τ_1 and τ_2 follows a Gaussian copula, it can be shown that (see Joe 1997 page 178) $\lim_{q\downarrow 0} P(\tau_2 < F_2^{-1}(q)|\tau_1 < F_1^{-1}(q)) = \lim_{q\uparrow 1} P(\tau_2 > F_2^{-1}(q)|\tau_1 > F_1^{-1}(q)) = 0$, where F_i is the distribution function of τ_i, $i = 1,2$. Therefore, the Gaussian copula model may not work during crisis, in which there should be strong dependence, but the Gaussian copula gives no dependence. In addition, Gaussian copula model cannot calibrate to the market data well. Extensions to the Gaussian copula model include the double-t copula (Hull and White 2004) and the random factor loading copula (Andersen and Sidenius 2004, 2005, Andersen 2006), among others.

The majority of dynamic bottom-up models are instantaneous intensity based models, where the default time of a single name is modeled as the first jump time of a Cox process, or doubly-stochastic Poisson process, which is characterized by its instantaneous default intensity. Single-name instantaneous intensity based credit risk models were introduced by Jarrow and Turnbull (1995), Lando (1994,1998), Schönbucher (1998) and Duffie and Singleton (1999). In instantaneous intensity based models, the default time τ of a single name can be represented as

$$\tau = \inf\left\{t \ge 0 : \int_0^t \lambda(s)ds \ge E\right\}, E \overset{d}{\sim} \exp(1),$$

where $\lambda(t)$ is the instantaneous default intensity of τ and $\exp(1)$ denotes exponential distribution with mean 1. The first instantaneous intensity based model for CDO pricing was proposed by Duffie and Gârleanu (2001) and was extended by Mortensen (2006), in which the correlation of default times τ_i are modeled through a factor structure of the default intensities. More precisely, Mortensen (2006) postulated that

$$\lambda_i(t) = a_i\lambda^m(t) + \lambda_i^{id}(t)$$
$$\tau_i = \inf\left\{t \ge 0 : \int_0^t \lambda_i(s)ds \ge E_i\right\}, i = 1,\dots,n, \quad (1)$$

where $\lambda^m(t)$ is the market factor intensity; $\lambda_i^{id}(t)$ is the idiosyncratic intensity of the i-th name; E_i, $i = 1,\dots,n$ are i.i.d. $\exp(1)$ random variables. Papageorgiou and Sircar (2007) used a square-root diffusion process with stochastic volatility to model the idiosyncratic intensity $\lambda_i^{id}(t)$.

There are some other dynamic bottom-up models. Joshi and Stacey (2006) proposed the intensity Gamma model

$$\tau_i = \inf\left\{t \ge 0 : \int_0^t c_i(s)dI(s) \ge E_i\right\}, i = 1,\dots,n,$$

where $I(t)$ is a multi-Gamma process representing the market information that drives the default of all names and $c_i(t)$ is a piecewise constant function. Schönbucher (2007) proposed a time-changed instantaneous intensity model

$$\tau_i = \inf\left\{t \ge 0 : \int_0^{T(t)} \lambda_i(s)ds \ge E_i\right\}, i = 1,\dots,n,$$

where $T(t)$ is a stochastic time change process that introduces correlation among default times.

Top-down models are generally dynamic models that describe directly the dynamics of the cumulative loss process of the whole portfolio. Top-down models are investigated in Arnsdorf and Halperin (2007), Cont and Minca (2007), Errais, Giesecke, and Goldberg (2006), Giesecke and Kim (2007), Longstaff and Rajan (2007), among others.

Bottom-up models are consistent with single name default probabilities but have more difficulty calibrating CDO tranches. Top-down models can calibrate CDO tranches excellently but they make little connection to the underlying single names. In this paper, we propose a new dynamic model, the Conditional Survival (CS) model, which provides excellent calibration to both CDO tranches and single name CDS spreads and enables calculation of sensitivity of CDO to underlying single name CDS spreads.

The remainder of the paper is organized as follows. The next section proposes the CS model. Section 3 provides the exact simulation algorithm for CDO pricing, the sensitivity analysis, and the algorithm for calibration. Section 4 shows the numerical results of calibrating the model to the CDS and iTraxx tranche spreads on March 14, 2008, right before the collapse of Bear Sterns.

2 THE NEW MODEL: CONDITIONAL SURVIVAL (CS) MODEL

2.1 Motivation

Schönbucher (2003) mentioned that the range of default correlation that can be generated by traditional instantaneous intensity models as in (1) may be limited, due to the fact that the default times are indirectly correlated through a common market factor in the default intensities. He pointed out that the introduction of joint jumps in the instantaneous default intensities can enhance the level of correlation. Mortensen (2006) used mean-reverting affine jump diffusion processes to model $\lambda^m(t)$ and $\lambda_i^{id}(t)$ in (1).

However, a major drawback of model (1) is that it can-not generate simultaneous default of many names, which has been seen during the current financial crisis. In (1), default time τ_i is defined as the first passage time of the *cumulative default intensity* process $\int_0^t \lambda_i(s)ds$ across a random barrier E_i. Although the default intensity $\lambda_i(t)$ can jump, the cumulative default intensity $\int_0^t \lambda_i(s)ds$ is *always continuous*. When a jump in the market intensity $\lambda^m(t)$ occurs, all default intensities $\lambda_i(t)$ jump together, but the jump effect is smoothed by the integration, i.e., $\int_0^t \lambda_i(s)ds$ does not change at all at the time of jump, but only starts to increase continuously at a higher rate after the jump occurs. Hence, jumps in the market intensity $\lambda^m(t)$ can not lead to simultaneous default of many names.

Table 1 shows the mid-bid-ask tranche spreads of the iTraxx Europe Series 8 5 Year index on September 20, 2007 and March 14, 2008. The Federal Reserve provided an emergency loan to near-bankrupt Bear Sterns on March 14 and financed the purchase of Bear Sterns by JPMorgan Chase two days later. The extremely high spreads of senior tranches on March 14, which are around 10 times as high as they were half a year ago, imply that the default correlation was substantially high during the peak of the crisis. A good model should be able to generate such strong correlation, i.e., the concurrent default of many names.

Table 1: iTraxx Europe Series 8 5 Year index mid-bid-ask tranche spreads on 09/20/07 and 03/14/08, denoted in units of basis points.

Tranches(%)	0-3	3-6	6-9	9-12	12-22	22-100
09/20/07	1812	84	37	23	15	7
03/14/08	5150	649	401	255	143	70

2.2 Conditional Survival (CS) Model

In order to incorporate simultaneous defaults, we introduce jumps to the cumulative intensities of default times. More precisely, we postulate the following Conditional Survival (CS) model:

$$\Lambda_i(t) = \sum_{j=1}^{J} a_{i,j} M_j(t) + X_i(t),$$
$$\tau_i = \inf\{t \geq 0 : \Lambda_i(t) \geq E_i\}, i = 1, \ldots, n, \quad (2)$$

where
(1) $M_j(t)$ represents the j-th market factor in the cumulative intensities, which is a nonnegative and increasing stochastic process with $M_j(0) = 0$, $j = 1, \ldots, J$. These market factors may be dependent on each other, but they are all independent of E_i and $X_i(t)$, $\forall i$. Most importantly, $M_j(t)$ are allowed to have *jumps*, which enables the model to generate simultaneous default of many names.
(2) $a_{i,j} \geq 0$ is the constant loading of the i-th name on the j-th market factor, $i = 1, \ldots, n; j = 1, \ldots, J$.
(3) $X_i(t)$ represents the idiosyncratic part of the cumulative default intensity of the i-th name, which is a nonnegative and increasing process with $X_i(0) = 0$, $i = 1, \ldots, n$. $\{X_i(t) : i = 1, \ldots, n\}$ are mutually independent, and they are independent of market factors $M_j(t), j = 1, \ldots, J$.
(4) $\{E_i : i = 1, \ldots, n\}$ are i.i.d. exp(1) random variables; they are independent of stochastic processes $X_i(t)$ and $M_j(t)$, $\forall i, j$.

Note that in this model, when a market factor $M_j(t)$ jumps, all cumulative intensities $\Lambda_i(t), i = 1, \ldots, n$ jump together simultaneously, which might lead to the concurrent crossing of the barriers E_i for many names.

We call this model "conditional survival" because conditional survival probabilities are the building blocks for CDO pricing in our model, as will be shown in section 3.

2.3 Conditional Survival Probability in the CS model

The conditional survival probability in the CS model is very simple. Let $\mathbf{M}(t) \triangleq (M_1(t), \ldots, M_J(t))$ be the vector of market factor processes. Let

$$q_i^c(t) \triangleq P(\tau_i > t | \mathbf{M}(t)) \text{ and } q_i(t) \triangleq P(\tau_i > t) \quad (3)$$

be the conditional survival probability and marginal survival probability of the i-th name, respectively. Then we have the following lemma:

Lemma 1. *For the i-th name in model* (2)*, we have*

$$q_i^c(t) = E\left[e^{-X_i(t)}\right] e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)}, \quad (4)$$
$$q_i(t) = E\left[e^{-X_i(t)}\right] E\left[e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)}\right], \quad (5)$$

And $q_i^c(t)$ can be represented by $q_i(t)$ and market factors as

$$q_i^c(t) = q_i(t) \cdot \frac{e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)}}{E\left[e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)}\right]}. \quad (6)$$

Proof. By (2), we have

$$\begin{aligned} q_i^c(t) &= E\left[1_{\{X_i(t)+\sum_{j=1}^{J} a_{i,j} M_j(t) < E_i\}} \middle| \mathbf{M}(t)\right] \\ &= E\left[E\left[1_{\{X_i(t)+\sum_{j=1}^{J} a_{i,j} M_j(t) < E_i\}} \middle| X_i(t), \mathbf{M}(t)\right] \middle| \mathbf{M}(t)\right] \\ &= E\left[e^{-X_i(t)-\sum_{j=1}^{J} a_{i,j} M_j(t)} \middle| \mathbf{M}(t)\right] \\ &= E\left[e^{-X_i(t)} \middle| \mathbf{M}(t)\right] e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)} \\ &= E\left[e^{-X_i(t)}\right] e^{-\sum_{j=1}^{J} a_{i,j} M_j(t)}, \end{aligned} \quad (7)$$

where the last equality follows from the independence of $X_i(t)$ and $\mathbf{M}(t)$. Then by taking expectation on both sides of (7), we obtain (5). Dividing equation (4) by (5), we obtain (6). □

Remark 1. *Note that marginal survival probability $q_i(t)$ can be implied from the market quotes of single name CDS spreads by standard bootstrap method. Using marginal survival probability $q_i(t)$ as input to the model allows automatic calibration to single name CDS.*

Remark 2. *Equation (6), though simple, has an important implication that there is no need to specify concrete models for idiosyncratic cumulative intensities $X_i(t)$. This is a distinctive feature of the CS model, in which we do not specify any dynamics for $X_i(t)$, because in the CS model, the portfolio loss distribution only depends on the market factor $\mathbf{M}(t)$ and marginal survival probability $q_i(t)$. The contribution of idiosyncratic risk $X_i(t)$ to the CDO pricing has been incorporated in $q_i(t)$.*

2.4 Constraints on Factor Loading Coefficients in the CS Model

In the CS model, the idiosyncratic cumulative default intensity $X_i(t)$ is nonnegative and increasing. Therefore, for a sequence of coupon payment dates $0 = T_0 < T_1 < \cdots < T_m = T$, it must hold that

$$E\left[e^{-X_i(T_k)}\right] \le E\left[e^{-X_i(T_{k-1})}\right], 1 \le k \le m. \tag{8}$$

By equation (5) in Lemma 1, we have $E\left[e^{-X_i(t)}\right] = \frac{q_i(t)}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(t)}\right]}$, which in combination with equation (8) gives the constraints on loading coefficients $a_{i,j}$:

$$\frac{q_i(T_k)}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(T_k)}\right]} \le \frac{q_i(T_{k-1})}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(T_{k-1})}\right]}, 1 \le k \le m. \tag{9}$$

2.5 Specifying Market Factors in the Model

As we will see in section 3, in order to price CDO using exact simulation, we have two requirements on the dynamics of market factors: (1) the market factors can be exactly simulated, (2) the Laplace transform of the market factors should have closed-form.

We use Pólya process to model the clustering of defaults that occurs both cross-sectionally and dynamically in time when the market is in crisis, and use the discrete integral of CIR process to model the fluctuation in the market during normal time.

2.5.1 Pólya Process: Modeling Clustering of Defaults

A Pólya process $M(t)$ is a generalized Poisson process. More precisely, it is a Poisson process with random rate Λ, where Λ is a Gamma random variable. Suppose the shape and scale parameters of Λ are α and β respectively, then the Laplace transform of $M(t)$ is given by

$$E\left[e^{-uM(t)}\right] = \left(\frac{p}{1-(1-p)e^{-u}}\right)^{\alpha}, \ p = \frac{1}{1+\beta t}. \tag{10}$$

A Pólya process $M(t)$ has a negative binomial distribution and has stationary but *positively correlated* increments. Indeed, it can be shown that $\text{cov}(M(t), M(t+h) - M(t)) = ht\alpha\beta^2 > 0$, so the arrival of one event tends to trigger the arrival of more events, which makes Pólya process suitable for the modeling of clustering of defaults occurring dynamically in time. As we have pointed out, the jumps in the Pólya process cause jumps in the cumulative intensities of all names, which produces the cross-sectional strong correlation.

2.5.2 Discrete Integral of CIR Process

Let $\lambda(t)$ be a square-root diffusion that has dynamics

$$d\lambda(t) = \kappa(\theta - \lambda(t))dt + \sigma\sqrt{\lambda(t)}dW(t), \tag{11}$$

where $W(t)$ is a standard Brownian motion. This process was studied in Feller (1951) and was proposed by Cox, Ingersoll, and Ross (1985) as a short rate model, generally referred to as the CIR model. The transition law of $\lambda(t)$ given $\lambda(s)$ for $s < t$ is $\lambda(t)|\lambda(s) \overset{d}{\sim} \alpha(s,t)\chi_d'^2(\beta(s,t)\lambda(s))$, where $\chi_v'^2(\lambda)$ denotes the noncentral chi-square distribution with v degrees of freedom and noncentrality parameter λ, $\alpha(s,t) = \frac{\sigma^2(1-e^{-\kappa(t-s)})}{4\kappa}$, $\beta(s,t) = \frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2(1-e^{-\kappa(t-s)})}$, and $d = \frac{4\kappa\theta}{\sigma^2}$. The transition law allows the exact simulation of the CIR process (see Glasserman 2004).

The discrete integral of CIR process is defined as

$$M(t) = \frac{h}{2}\lambda(t_0) + h\sum_{i=1}^{m-1}\lambda(t_i) + \frac{h}{2}\lambda(t_m), \tag{12}$$

where m is the number of discretization steps in the time interval $[0,t]$, $h = t/m$, $t_i = \frac{it}{m}$. The discrete integral $M(t)$ defined in (12) is the trapezoidal approximation to the exact integral $\int_0^t \lambda(s)ds$. The advantage of using (12) is that its exact simulation only involves the simulation of a CIR process, which may be easier and faster than the exact simulation of $\int_0^t \lambda(s)ds$ (see Broadie and Kaya 2006 for the exact simulation of $\int_0^t \lambda(s)ds$).

The transition law of the CIR process implies that the conditional Laplace transform of $\lambda(t)$ is given by

$$E\left[e^{-u\lambda(t)}\middle|\lambda(s)\right] = (1+2\alpha(s,t)u)^{-\frac{d}{2}} e^{-\frac{\alpha(s,t)\beta(s,t)u}{1+2\alpha(s,t)u}\lambda(s)}. \tag{13}$$

Let $\alpha_i = \alpha(t_i, t_{i+1})$, $\beta_i = \beta(t_i, t_{i+1})$, $i = 0, 1, \ldots, m-1$. Then the Laplace transform of the discrete integral process (12) is given in the following lemma:

Lemma 2. *The Laplace transform of $M(t)$ in (12) is*

$$E\left[e^{-uM(t)}\right] = e^{f_0(u)\lambda(t_0)} \left(\prod_{i=0}^{m-1}(1-2\alpha_i f_{i+1}(u))\right)^{-d/2},$$

where $d = \frac{4\kappa\theta}{\sigma^2}$ and $f_m(u)$, $f_{m-1}(u)$, …, $f_0(u)$ are defined by the recursion

$$f_m(u) = -\frac{h}{2}u; f_{i-1}(u) = -hu + \frac{\alpha_{i-1}\beta_{i-1}f_i(u)}{1-2\alpha_{i-1}f_i(u)}, m \geq i \geq 2;$$
$$f_0(u) = -\frac{h}{2}u + \frac{\alpha_0\beta_0 f_1(u)}{1-2\alpha_0 f_1(u)}. \tag{14}$$

Proof. See Appendix A. □

3 CDO PRICING, SENSITIVITY ANALYSIS and CALIBRATION IN THE CS MODEL

3.1 Notation and Setting of CDO Pricing

Let T be the maturity date of the CDO, and $T_1, T_2, \ldots, T_m = T$ be the coupon payment dates. Let $[a,b]$ be the CDO tranche loss window. Let N_i and R_i be the notional principle and recovery rate of the i-th name in the portfolio, respectively. The cumulative loss process of the portfolio is defined as

$$L_t = \sum_{i=1}^{n}(1-R_i)N_i 1_{\{\tau_i \leq t\}}, 0 \leq t \leq T. \tag{15}$$

The tranche cumulative loss process is defined as

$$L_t^{[a,b]} = (L_t - a)^+ - (L_t - b)^+, 0 \leq t \leq T, \tag{16}$$

which denotes the cumulative loss assumed by the investor of the CDO tranche up to time t. The outstanding notional of the tranche at time t is $O_t^{[a,b]} = b - a - L_t^{[a,b]}$. Let $D(0,t)$ be the risk free discounter factor from time t to time 0.

If a default happens at time τ, the investor of the tranche would make a payment equal to $L_\tau^{[a,b]} - L_{\tau-}^{[a,b]}$, where $L_{\tau-}^{[a,b]} \triangleq \lim_{t\uparrow\tau} L_t^{[a,b]}$. Therefore, the present value of the default leg of the CDO tranche can be represented as $E[\int_0^T D(0,t)dL_t^{[a,b]}]$, under the risk-neutral probability measure.

For simplicity, it is usually assumed in the literature that defaults only occur in the middle of coupon payment dates (see e.g. Mortensen 2006, Andersen et al. 2003, Papageorgiou and Sircar 2007). Under this simplification, the present value of the default leg is given by

$$E\left[\int_0^T D(0,t)dL_t^{[a,b]}\right] = E\left[\sum_{i=1}^{m}\int_{T_{i-1}}^{T_i} D(0,t)dL_t^{[a,b]}\right]$$
$$= E\left[\sum_{i=1}^{m} D(0, \frac{T_{i-1}+T_i}{2})(L_{T_i}^{[a,b]} - L_{T_{i-1}}^{[a,b]})\right] \tag{17}$$

Let $S^{[a,b]}$ be the CDO tranche spread. The coupon payment at time $T_i, i = 1, \ldots, m$ is specified as

$$S^{[a,b]}(T_i - T_{i-1})\int_{T_{i-1}}^{T_i} \frac{O_t^{[a,b]}}{T_i - T_{i-1}}dt = S^{[a,b]}\int_{T_{i-1}}^{T_i} O_t^{[a,b]}dt.$$

Assuming defaults only occur in the middle of premium periods, the present value of the premium leg is given by

$$E\left[\sum_{i=1}^{m} D(0,T_i)S^{[a,b]}\int_{T_{i-1}}^{T_i} O_t^{[a,b]}dt\right]$$
$$= E\left[\sum_{i=1}^{m} D(0,T_i)S^{[a,b]}(T_i - T_{i-1})\frac{1}{2}(O_{T_{i-1}}^{[a,b]} + O_{T_i}^{[a,b]})\right]$$
$$= S^{[a,b]}E\left[\sum_{i=1}^{m} D(0,T_i)(T_i - T_{i-1})\left(b - a - \frac{L_{T_i}^{[a,b]} + L_{T_{i-1}}^{[a,b]}}{2}\right)\right] \tag{18}$$

Then the fair spread for the CDO tranche is

$$S^{[a,b]} = \frac{E\left[\sum_{i=1}^{m} D(0,\frac{T_{i-1}+T_i}{2})(L_{T_i}^{[a,b]} - L_{T_{i-1}}^{[a,b]})\right]}{E\left[\sum_{i=1}^{m} D(0,T_i)(T_i - T_{i-1})\left(b - a - \frac{L_{T_i}^{[a,b]} + L_{T_{i-1}}^{[a,b]}}{2}\right)\right]}. \tag{19}$$

The cash flow specification for the equity tranche is different from the other tranches. The seller of the equity tranche pays an upfront fee at time 0 to the investor and pays coupons at a fixed running spread of 500 basis points. The equity tranche spread is defined as the ratio of the upfront fee to the notional of equity tranche, which is given by

$$S^{[0,b]} = \frac{1}{b}\left\{E\left[\sum_{i=1}^{m} D(0,\frac{T_{i-1}+T_i}{2})(L_{T_i}^{[0,b]} - L_{T_{i-1}}^{[0,b]})\right]\right.$$
$$\left. -0.05E\left[\sum_{i=1}^{m} D(0,T_i)(T_i - T_{i-1})\left(b - \frac{L_{T_i}^{[0,b]} + L_{T_{i-1}}^{[0,b]}}{2}\right)\right]\right\}. \tag{20}$$

3.2 Pricing CDO in the CS Model by Exact Simulation

It is clear from equations (19) and (20) that the CDO tranche spreads are determined by the marginal distribution of the cumulative loss process L_t at coupon payment dates T_k. Hence, to price CDO, we only need to simulate L_{T_k} exactly.

The key observation is that conditional on $\mathbf{M}(t)$, the random variables $1_{\{\tau_i \le t\}}$, $i=1,\dots,n$ are conditionally independent, and $1_{\{\tau_i \le t\}}$ has a Bernoulli$(1-q_i^c(t))$ distribution, $i=1,\dots,n$. Therefore, by equation (15), conditional on the market factor $\mathbf{M}(t)$, the cumulative loss L_t is equal to the weighted sum of n independent Bernoulli random variables, which can be easily simulated.

More precisely, suppose $\mathbf{M}(t)$ can be easily simulated and $E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(t)}\right]$ can be calculated in closed form, then we have the following algorithm to simulate the cumulative loss L_{T_k}, $k=1,\dots,m$:

Step 1: Generate sample path of market factors $\mathbf{M}(T_k)$, $k=1,\dots,m$.
Step 2: For each $i=1,\dots,n$, calculate the conditional survival probability $q_i^c(T_k)$, $k=1,\dots,m$, according to (6), using the samples of market factors generated in Step 1.
Step 3: Generate independent Bernoulli random variables $I_{i,k} \overset{d}{\sim}$ Bernoulli$(1-q_i^c(T_k))$, $i=1,\dots,n; k=1,\dots,m$.
Step 4: Calculate $L_{T_k} = \sum_{i=1}^n (1-R_i)N_i I_{i,k}$, $k=1,\dots,m$.

Using simulated samples of L_{T_k}, we can estimate the CDO tranche spreads $S^{[a,b]}$ given in equations (19) and (20).

By equation (15), the expectation of the cumulative loss L_t is given by $E[L_t] = \sum_{i=1}^n (1-R_i)N_i[1-q_i(t)]$. So L_{T_k} can be used as the control variates for variance reduction.

3.3 Sensitivity Analysis in the CS Model

Since the present value of the default leg and the premium leg of the CDO tranches are both determined by $E[L_t^{[a,b]}]$(cf. (17) and (18)), calculating CDO sensitivities amounts to calculating sensitivities of $E[L_t^{[a,b]}]$. One of the major advantages of the CS model is that the sensitivities of CDO to single name CDS can be obtained concurrently with CDO pricing. More precisely, we have the following lemma:

Lemma 3. *The sensitivities of $E[(L_t - a)^+]$ to the survival probability of the i-th name is*

$$\frac{\partial E[(L_t-a)^+]}{\partial q_i(t)} = E\left\{\frac{q_i^c(t)}{q_i(t)}[(L_t^{(-i)}-a)^+ - (L_t^{(-i)} + (1-R_i)N_i - a)^+]\right\}, \quad (21)$$

where $L_t^{(-i)} \triangleq \sum_{j\neq i}(1-R_j)N_j 1_{\{\tau_j \le t\}}$ represents the cumulative loss of the portfolio excluding the i-th name.

Proof. See Appendix B. □

Remark 3. *By* (6)*, we have* $\frac{q_i^c(t)}{q_i(t)} = \frac{e^{-\sum_{j=1}^J a_{i,j}M_j(t)}}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(t)}\right]}$. *In addition, $L_{T_k}^{(-i)} = \sum_{j\neq i}(1-R_j)N_j I_{j,k}$, where $\{I_{j,k}\}$ are the Bernoulli random variables generated in the pricing algorithm in Section 3.2. Hence, the sensitivities can be calculated at the same time as the CDO is priced with little extra computational cost. This is one of the major advantages of the CS model.*

3.4 Calibrate the Model to Market CDS and CDO Spreads

As we mentioned in section 2.3, we use $q_i(t)$ implied from market CDS spreads as input to our model, so the CS model provides automatic calibration to the CDS market data.

To calibrate to market CDO tranche spreads, we need to determine two sets of parameters: (1) the parameters that specify the dynamics of $\mathbf{M}(t)$, which we denote by Θ; (2) the factor loading coefficients $(a_{i,1},\dots,a_{i,J})$ for each name i, $i=1,\dots,n$.

3.4.1 Determine Factor Loading Coefficients by Regression

Suppose the parameters Θ for market factors have been fixed. By equation (2), the factor loading coefficients $(a_{i,1},\dots,a_{i,J})$ can be viewed as the regression coefficients of $\Lambda_i(t)$ on the regressors $(M_1(t),\dots,M_J(t))$, with $X_i(t)$ being the regression error term. When the regression error term $X_i(t)$ is small, $E\left[e^{-X_i(t)}\right]$ is close to 1. Therefore, the regression error can be measured by $\left|1-E\left[e^{-X_i(t)}\right]\right|$. Expecting that good market factors are able to explain substantial part of the default risk, we determine $(a_{i,1},\dots,a_{i,J})$ by minimizing the regression error $\left|1-E\left[e^{-X_i(t)}\right]\right|$. By equation (5) and the fact that $E\left[e^{-X_i(t)}\right] \le 1$ since $X_i(t)\ge 0$, minimizing $\left|1-E\left[e^{-X_i(t)}\right]\right|$ is equivalent to minimizing $E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(t)}\right] - q_i(t)$. Recalling the loading coefficients must satisfy the constraints in (9), $(a_{i,1},\dots,a_{i,J})$ can be obtained by solving the following optimization problem:

$$\begin{aligned} &\min E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(T)}\right] - q_i(T) \\ &\text{s.t.} \ \frac{q_i(T_k)}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(T_k)}\right]} \le \frac{q_i(T_{k-1})}{E\left[e^{-\sum_{j=1}^J a_{i,j}M_j(T_{k-1})}\right]}, 1\le k\le m \\ &\quad 0 \le a_{i,j},\ j=1,\dots,J \end{aligned} \quad (22)$$

3.4.2 Objective Function in Calibration

Let $s_k^{o,b}$ and $s_k^{o,a}$ be the bid and ask of the k-th CDO tranche spread observed in the market, respectively, $k = 1,\ldots,K$, where K is the number of CDO tranches. Let $s_k^o = (s_k^{o,a} + s_k^{o,b})/2$ be the mid-bid-ask tranche spread. Let s_k be the tranche spread computed by the CS model. The chi-square of model fitting is given by

$$\text{CHISQ} = \sum_{k=1}^{K} \frac{(s_k - s_k^o)^2}{s_k}. \tag{23}$$

The root-mean-square error of model fitting is given by

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\frac{s_k - s_k^o}{s_k^{o,a} - s_k^{o,b}}\right)^2}. \tag{24}$$

We use the chi-square as the objective function in the calibration, i.e, we search for the market factor parameters Θ which minimize the chi-square. As a reference, we also report the root-mean-square error.

3.4.3 The Calibration Algorithm

The calibration algorithm for market factor parameters Θ is as follows:

Step 1: We start from an initial guess of the market factor parameters Θ.
Step 2: Given the market factor parameters Θ, for each $i = 1,2,\ldots,n$, we determine the factor loading coefficients $(a_{i,1},\ldots,a_{i,J})$ by solving the problem formulated in (22).
Step 3: Calculate CDO tranche spread $s_k, k = 1,\ldots,K$ by the CS model using the market factor parameters Θ and the factor loading coefficients obtained in the previous step.
Step 4: Calculate the chi-square that corresponds to the current market factor parameters Θ.
Step 5: If the chi-square is small enough, stop; otherwise, update the market factor parameters Θ using certain unconstrained optimization algorithm (e.g., Powell's direction-set method, see Press et al. 2002), and go to Step 2.

Remark 4. *In step 2, we need to solve n optimization problems* (22) *to find the loading coefficients for all n names. The dimension of problem* (22) *is equal to the number of market factors, which is typically not grater than 3. The objective function and the constraints all have closed-form derivatives, and the objective function is monotonic. Therefore, problem* (22) *can be solved quickly by common methods such as sequential quadratic programming. We used the CFSQP package developed by Lawrence et al. (1997).*

Table 2: The second and the fourth rows showed the iTraxx Europe Series 8 5 Year Index mid-bid-ask tranche spreads and the bid-ask spreads observed in the market on 03/14/08, respectively. The third and the fifth rows showed the tranche spreads calculated from the CS model and their standard error in the Monte Carlo simulation, respectively. We used 50,000 replications in the Monte Carlo simulation without using variance reduction. All the spreads are denoted in units of basis points. The chi-square of the fitting is 6.7. The root-mean-square error is 1.01.

Tranches(%)	0-3	3-6	6-9	9-12	12-22	22-100
Market	5150	649	401	255	143	70
Model	5231	665	374	250	165	67
B-A spread	158	24	25	20	12	3
MC error	15	5	4	3	3	1

4 NUMERICAL RESULTS OF CALIBRATION

In this section, we show the numerical result of calibrating the CS model to the CDS and iTraxx Europe Series 8 5 Year index data on March 14, 2008.

4.1 Description of the Data

The iTraxx Europe Series 8 5 Year index is composed of the most liquid 124 CDS referencing European investment grade companies. Each company has a notional of 8/3 million, except UniCredit SpA which has a notional of 16/3 million. The index has 6 trances, which correspond to the 0%-3%, 3%-6%, 6%-9%, 9%-12%, 12%-22% and 22% to 100% of the portfolio notional, respectively. The 0%-3% is quoted in terms of percentage upfront fee, and the other tranches are quoted in terms of running spread. The index was launched on September 20, 2007 and matures in 5 years. It pays quarterly coupons. The first coupon date after March 14, 2008 is June 20, 2008. The last coupon date is December 20, 2012. The market quotes of the mid-bid-ask tranche spreads and the bid-ask spreads are showed in the second and fourth rows of Table 2, respectively.

The discount factors $D(0,t)$ in the pricing equation (19) are extracted from the Euro fixing swap curve on March 14, 2008. We use linear extrapolation to obtain the zero curve.

We use fixed recovery rate $R_i = 40\%$, $i = 1,\ldots,n$ in the calibration.

We use the 5 year CDS spread on March 14, 2008 to extract the marginal survival probabilities $q_i(t)$ for each name i, assuming that the hazard rate function is flat during the 5 year period.

All the market data were obtained from Bloomberg business electronic resources.

4.2 Parameters to be Calibrated

We use three independent market factors in the calibration: one Pólya process and two discrete integral of CIR processes defined in (12). For a CIR process $\lambda(t)$ with parameters $(\kappa, \theta, \sigma, \lambda(0))$ which has dynamics given in (11), its multiple $c\lambda(t)$ is still a CIR process with parameters $(\kappa, c\theta, \sqrt{c}\sigma, c\lambda(0))$. Therefore, one can always scale the CIR process by a factor $c > 0$ to make θ equal to a constant and divide the corresponding loading coefficients $a_{i,j}$ by c at the same time. This way model (2) does not change. In other words, only three free parameters need to be determined for the CIR process. We choose to fix $\theta = 0.1$ and estimate $(\kappa, \sigma, \lambda(0))$. To summarize, there are 8 parameters $\Theta = (\alpha, \beta, \kappa_1, \sigma_1, \lambda_1(0), \kappa_2, \sigma_2, \lambda_2(0))$ in total, where α and β are parameters for the Pólya process and $\kappa_i, \sigma_i, \lambda_i(0)$ are parameters for the i-th CIR process, $i = 1, 2$. For the discrete integral of CIR processes, we used 10 discretization steps for the first premium period (from 03/14/08 to 06/20/08) and 8 steps for each of the remaining premium periods.

4.3 Calibration Results

The iTraxx tranche spreads calculated from the CS model and their standard error in the Monte Carlo simulation are showed in the third and fifth rows of Table 2, respectively. The chi-square of the fitting is 6.7, corresponding to a p-value 0.24. The root-mean-square error is 1.01. It is clear that the CS model provided excellent fitting to all the tranche spreads.

The calibrated parameters are $\alpha = 2.64515812$, $\beta = 0.00583919$, $\kappa_1 = 0.12792013$, $\sigma_1 = 1.34700857$, $\lambda_1(0) = 1.14110000$, $\kappa_2 = 0.00100369$, $\sigma_2 = 7.40483282$, and $\lambda_2(0) = 0.10650913$.

Figure 1 shows the implied copula correlation of the market quotes on March 14, 2008 and the model generated tranche spreads. The skew of the implied copula correlation, which was dramatic during the financial crisis, is excellently reproduced by the CS model.

ACKNOWLEDGMENTS

We would like to thank Zaiwen Wen for his kind help with the optimization issues in calibration.

A PROOF OF LEMMA 2

Proof. Let $\mathscr{F}_t \triangleq \sigma(\lambda(s), s \le t)$ and $f_k(u)$ be defined in (14), $k = m, \ldots, 0$. Then

$$E\left[e^{-uM(t)}\right] = E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-1}\lambda(t_i)+\frac{h}{2}\lambda(t_m))}\right]$$
$$=E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-1}\lambda(t_i))+f_m(u)\lambda(t_m)}\right]$$
$$=E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-1}\lambda(t_i))}\,E\left[e^{f_m(u)\lambda(t_m)}\middle|\mathscr{F}_{t_{m-1}}\right]\right]$$
$$=E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-1}\lambda(t_i))}\,E\left[e^{f_m(u)\lambda(t_m)}\middle|\lambda(t_{m-1})\right]\right]$$
$$=E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-1}\lambda(t_i))}\frac{e^{\frac{\alpha_{m-1}\beta_{m-1}f_m(u)}{1-2\alpha_{m-1}f_m(u)}\lambda(t_{m-1})}}{(1-2\alpha_{m-1}f_m(u))^{\frac{d}{2}}}\right] \quad \text{(by (13))}$$
$$=(1-2\alpha_{m-1}f_m(u))^{-\frac{d}{2}}E\left[e^{-u(\frac{h}{2}\lambda(t_0)+h\sum_{i=1}^{m-2}\lambda(t_i))+f_{m-1}(u)\lambda(t_{m-1})}\right].$$

Repeating the above argument by successively conditioning on $\mathscr{F}_{t_{m-2}}$, $\mathscr{F}_{t_{m-3}}$, ..., and $\mathscr{F}_{t_1}$, we obtain $E\left[e^{-uM(t)}\right] = e^{f_0(u)\lambda(t_0)}\left(\prod_{i=0}^{m-1}(1-2\alpha_i f_{i+1}(u))\right)^{-d/2}$. □

B PROOF OF LEMMA 3

Proof. Let $c_i \triangleq (1-R_i)N_i$. Noting that conditional on $\mathbf{M}(t)$, the n default indicators $1_{\{\tau_i \le t\}}$, $i = 1, \ldots, n$ are conditionally independent, we have

$$E[(L_t - a)^+|\mathbf{M}(t)] = E[(L_t^{(-i)} + c_i 1_{\{\tau_i \le t\}} - a)^+|\mathbf{M}(t)]$$
$$=E[E[(L_t^{(-i)} + c_i 1_{\{\tau_i \le t\}} - a)^+|\mathbf{M}(t), L_t^{(-i)}]|\mathbf{M}(t)]$$
$$=E[(L_t^{(-i)} + c_i - a)^+(1 - q_i^c(t)) + (L_t^{(-i)} - a)^+ q_i^c(t)|\mathbf{M}(t)]$$
$$=E[(L_t^{(-i)} + c_i - a)^+|\mathbf{M}(t)]$$
$$+E\{q_i^c(t)[(L_t^{(-i)} - a)^+ - (L_t^{(-i)} + c_i - a)^+]|\mathbf{M}(t)\}.$$

Therefore,

$$E[(L_t - a)^+] = E[(L_t^{(-i)} + c_i - a)^+] + q_i(t)E\left\{\frac{q_i^c(t)}{q_i(t)}[(L_t^{(-i)} - a)^+ - (L_t^{(-i)} + c_i - a)^+]\right\}. \quad (25)$$

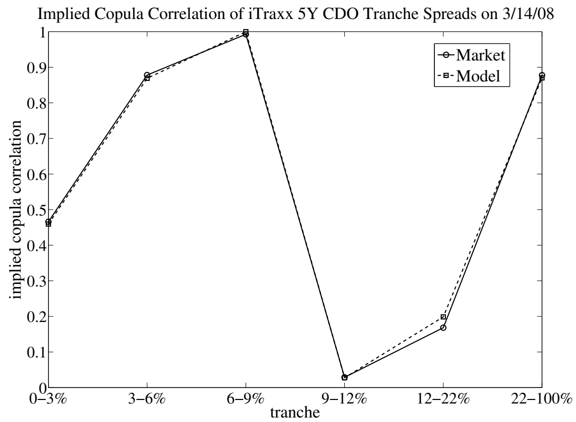

Figure 1: The implied copula correlation of the market data on 03/14/08 and the model generated spreads.

By (6), $\frac{q_i^c(t)}{q_i(t)}$ does not depend on $q_i(t)$, so (21) is obtained through differentiating the above equation by $q_i(t)$. □

REFERENCES

Andersen, L. 2006. Portfolio losses in factor models: term structures and intertemporal loss dependence. Working paper, Bank of America.

Andersen, L., and J. Sidenius. 2004. Extensions to the gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk* 1 (1): 29–70.

Andersen, L., and J. Sidenius. 2005. Cdo pricing with factor models: survey and comments. *Journal of Credit Risk* 1 (3): 71–88.

Andersen, L., J. Sidenius, and S. Basu. 2003. All your hedges in one basket. *Risk* 16 (11): 62–72.

Arnsdorf, M., and I. Halperin. 2007. Bslp: Markovian bivariate spread-loss model for portfolio credit derivatives. Working paper, Quantitative Research J.P. Morgan.

Broadie, M., and O. Kaya. 2006. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54 (2): 217–231.

Cont, R., and A. Minca. 2007. Reconstructing portfolio default rates from cdo tranche spreads. Working paper, Columbia University.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. A theory of the term structure of interest rates. *Econometrica* 53:129–151.

Duffie, D., and N. Gârleanu. 2001. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal* 57:41–59.

Duffie, D., and K. Singleton. 1999. Modeling term structure of defaultable bonds. *Review of Financial Studies* 12:687–720.

Errais, E., K. Giesecke, and L. Goldberg. 2006. Pricing credit from the top down with affine point processes. Working paper, Stanford University.

Feller, W. 1951. Two singular diffusion problems. *Annals of Mathematics* 54:173–182.

Giesecke, K., and B. Kim. 2007. Estimating tranche spreads by loss process simulation. In *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton: IEEE Press.

Glasserman, P. 2004. *Monte carlo methods in financial engineering*. New York: Springer-Verlag.

Hull, J., and A. White. 2004. Valuation of a cdo and an n-th to default cds without monte carlo simulation. *Journal of Derivatives* 12 (2): 8–23.

Jarrow, R., and S. Turnbull. 1995. Pricing options on financial securities subject to credit risk. *Journal of Finance* 50:53–85.

Joe, H. 1997. *Multivariate models and dependence concepts*. New York: Chapman & Hall/CRC.

Joshi, M. S., and A. M. Stacey. 2006, July. Intensity gamma: a new approach to pricing portfolio credit derivatives. *Risk* 19 (7): 78–83.

Lando, D. 1994. *Three essays on contingent claims pricing*. Ph.D. thesis, Cornell University.

Lando, D. 1998. On cox processes and credit risky securities. *Review of Derivatives Research* 2:99–120.

Lawrence, C. T., J. L. Zhou, and A. L. Tits. 1997. User's guide for cfsqp version 2.5: a c code for solving (large scale) constrained nonlinear (minimax) optimization problems. Document, University of Maryland.

Li, D. X. 2000. On default correlation: a copula function approach. *Journal of Fixed Income* 9:43–54.

Longstaff, F., and A. Rajan. 2007. An empirical analysis of collateralized debt obligations. Working paper, University of California, Los Angeles.

Mortensen, A. 2006. Semi-analytical valuation of basket credit derivatives in intensity-based models. *Journal of Derivatives* 13:8–26.

Papageorgiou, E., and R. Sircar. 2007. Multiscale intensity models and name grouping for valuation of multi-name credit derivatives. Working paper, Department of Operations Research & Financial Engineering, Princeton University.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 2002. *Numerical recipes in c++: the art of scientific computing*. 2nd ed. New York: Cambridge University Press.

Schönbucher, P. J. 1998. Term structure modelling of defaultable bonds. *Review of Derivatives Research* 2:161–192.

Schönbucher, P. J. 2003. *Credit derivatives pricing models: Models, pricing and implementation*. England: John Wiley & Sons.

Schönbucher, P. J. 2007. Time for a time-change: A new approach to multivariate intensity models of credit risk. Working paper.

AUTHOR BIOGRAPHIES

XIAN HUA PENG is a Ph.D. candidate in the Department of Industrial Engineering and Operations Research at Columbia University. His research interests include stochastic modeling, credit derivatives and risk management. His email address is <xp2102@columbia.edu> and web page is <www.columbia.edu/~xp2102>.

STEVEN S. G. KOU is a professor in the Department of Industrial Engineering and Operations Research at Columbia University. He received a Ph.D. degree in statistics from Columbia University. His research interests include financial engineering, applied probability and stochastic models in medicine. His email address is <sk75@columbia.edu> and web page is <www.columbia.edu/~sk75>.