Research Paper

Dynamic Pricing of Synthetic Collateralized Debt Obligations

Date:2008
Reference Number:8/1

Risk Control Limited
22 Romilly Street
London W1D 5AG
UK
T: +44 (0)20 7434 4007

Risk Control Limited inc
1705 Dixon Lane
Suite 200
St. Simons Island GA 31522
T: +1 404 418 5377
F: +1 404 437 6352

www.riskcontrollimited.com

DYNAMIC PRICING OF SYNTHETIC COLLATERALIZED DEBT OBLIGATIONS

Robert Lamb	William Perraudin	Astrid Van Landschoot
Imperial College	Imperial College	Standard & Poor's
London	London	London

March 2008*

Abstract

This paper applies a new class of dynamic credit loss rate models to the pricing of benchmark synthetic Collateralized Debt Obligations (CDOs). Our approach builds directly on the static, industry-standard, pricing approach to credit structured products based on Vasicek (1991). We generalize the Vasicek model by allowing risk factors to be driven by arbitrarily complex autoregressive processes. We show how to benchmark our model using CDX prices, and demonstrate that it can consistently and accurately fit the prices of multiple tranches with different subordination levels and tenors. Among other interesting results, we find that changes in tranche spreads are driven less by alterations in the market's estimate of default correlation (which is stable over time) and more by fluctuations in market perceptions of the persistence of credit shocks, i.e., the persistence of the credit cycle.

*The authors' may be contacted at rlamb@imperial.ac.uk, wperraudin@imperial.ac.uk or astrid_vanlandschoot@standardandpoors.com.

1 Introduction

The credit derivatives market has grown exponentially in recent years. At the end of 2006, its size exceeded \$30 trillion according to estimates by the British Bankers Association. One of the most actively traded categories of credit derivatives is synthetic CDOs. In a typical transaction, a protection seller agrees to bear losses incurred on a pool of credit exposures to a set of named borrowers with some notional or par amount. In return, a protection buyer pays the seller a premium proportional to the notional of the transaction. Losses in these deals are usually tranched in the sense that the protection seller promises to bear losses in some specified range such as from 3% to 6% of the notional.

Much of the trading in synthetic CDOs revolves around standardized contracts such as the iTraxx and the CDX. The names underlying these contracts are the debt issuers for which single-name Credit Default Swaps (CDS) are most widely traded. Because these names are central to the international debt markets, basket credit derivatives based on them like the iTraxx and CDX have come to now play a key role for market participants wishing to take on or hedge exposure to the credit market in general.

Researchers have developed a series of simple models for pricing synthetic CDOs. An important model widely used by market participants is based on a loss distribution originally derived by Vasicek (1991). The Vasicek model has been elaborated and extended by many studies, including Schonbucher (2002), Laurent and Gregory (2005) and Hull and White (2004). A comparative survey of such models is provided by Burtschell, Gregory, and Laurent (2005). The industry primarily uses a simple but robust version of the Vasicek model, namely the so-called base correlation approach described by McGinty and Ahluwalia (2004).

Instead of generating a loss distribution, in an influential contribution Li (2000) showed how one may simulate correlated default events using a Gaussian copula. Other copulas have then been suggested. Schonbucher and Schubert (2001) looks at these in detail including models with "infectious defaults" (i.e., models in which default probabilities for other names increase when a given obligor defaults). Giesecke and Goldberg (2005) also look at self-exciting processes where intensities respond to events as they occur. An early example of infectious defaults can be attributed to Davis and Lo (2001).

A major drawback of the Vasicek model and most of its generalizations is that these models are static. A loss distribution is formulated for a credit portfolio held over some fixed time such as the maturity of a synthetic CDO. A deal is valued by calculating the discounted, expected loss on a tranched exposure to this loss distribution. This approach does not yield consistent pricing of tranches with different maturities as risk is modeled from the standpoint of a single point in time and there is no attempt to develop a consistent set of distributions for losses over different horizons. Also, analysis of hedging is difficult within static models as there is no consistent framework for examining the behavior of price changes from one period to the next.

For these reasons, researchers have focussed on deriving dynamic models for pricing CDOs. Before reviewing recent research, it is worth noting that one of the earliest studies of CDO pricing, Duffie and Garleanu (2001), employed a fully dynamic model. These authors generated correlated intensities using affine processes for individual names and apply these to CDO valuation. The main problem with this approach is that it is known to exhibit limited correlated defaults even when using perfect correlation between two hazards, see Das, Duffie, Kapadia, and Saita (2007). Also practical difficulties due to Monte Carlo simulation and the complexities of calibration.

More recently, Chapovsky, Rennie, and Tavares (2006) propose a similar model. In their framework, individual defaults are driven by a hazard rate equal to the sum of a common random process with known dynamics, such as a CIR process, and a deterministic function calibrated to individual names. Giesecke and Goldberg (2005) develop an intensity based approach to modeling total portfolio losses, inferring single name default processes using 'thinning' techniques. Hull and White (2007) present a reduced form model in which the hazard rate for a company follows a deterministic process that is subject to periodic impulses. This leads to a jump process for the cumulative hazard rate. The model allows to value CDOs and options on CDOs analytically.

Recently, Sidenius, Piterbarg, and Anderson (2006), Schonbucher (2006) and Brigo, Pallavicini, and Torresetti (2007) amongst others have developed dynamic approaches modeling evolution of the losses on a portfolio. Sidenius, Piterbarg, and Anderson (2006) and Schonbucher (2006) are very similar in spirit. Both models are akin to the Heath-Jarrow-Morton term structure framework where they model the full forward distribution of the loss process. Sidenius, Piterbarg, and Anderson (2006) models the loss distribution in absence of information about default times. Calibration to the market is performed by conditioning upon a background process. The loss process then evolves as a Markov process based on the path of the background process.

Schonbucher (2006) looks at the transition rates of the loss process that are inferred from a Markov chain based on the transition probability distribution. Dynamics are then introduced by allowing the transition rates to be stochastic. Brigo, Pallavicini, and Torresetti (2007) assumes the loss process is a sum of independent Poisson processes that incorporates correlation into the model. He later builds dynamics into the model by allowing the intensities of the Poisson processes to be dynamic.

Lamb and Perraudin (2006) show how the dynamics may be introduced into the simple Vasicek (1991) by allowing the common factor to be an autoregressive time series process. They derive a closed form expression for a simple transformation of the losses on a credit portfolio and then apply this in modeling losses on aggregate loan portfolios of large US banks.

The contribution of the current paper is to generalize the Vasicek in a direct way to conditionally-evolving dynamic loss distributions and then to apply this approach to pricing synthetic CDOs. Though we focus here on synthetic CDOs, a type of structured product that has a very simple cash flow "waterfall" structure, our approach could be employed for pricing a much wider set of securitization-style exposures.

In Section 2 of the paper, we derive the dynamic process for the portfolio loss distribution when common factors possess an arbitrarily complex autoregressive form. We show how the distribution of losses at future dates is affected by conditioning information. In Section 3, we describe how the dynamic loss distribution may be employed in synthetic CDO valuation. In Section 4, we fit the model to data on CDX contract spreads. Section 5 concludes.

2 Dynamic Loss Model

2.1 Loss Rate Process

Suppose that time is discrete taking values $t = 0, 1, \ldots$ and that there are n obligors in an economy. Given survival until $t-1$, obligor i defaults at time t if:

$$Z_{i,t} \leq c_t \tag{1}$$

for a constant, c_t. As there are multiple obligors, default correlation is introduced into the model by defining $Z_{i,t}$ to be a latent random variable such that:

$$Z_{i,t} = \sqrt{\rho} X_t + \sqrt{1-\rho}\epsilon_{i,t}. \tag{2}$$

Here, the common factor, X_t, is a standard normal random variable. The obligor-specific idiosyncratic shock, $\epsilon_{i,t}$, has a distribution function H, a zero mean and unit variance, and is independent of X_t. This implies that $Z_{i,t}$ also has unit variance and zero mean and that the pairwise correlation between i and j for any i and j is ρ.

The distribution of $Z_{i,t}$ denoted G may be obtained as the convolution of H and a standard normal distribution function Φ. G depends on ρ and on a vector of parameters describing H denoted ν. G equals:

$$G(z) = \int_{-\infty}^{\infty} H\left(\frac{z-\sqrt{\rho}x}{\sqrt{1-\rho}}\right) d\Phi(x) \ . \tag{3}$$

Given this distribution, one may express the unconditional probability that default will occur at a future date t:

$$q_t = \text{Prob}\left(\text{default at } t\right) \ = \ G\left(c_t\right). \tag{4}$$

The model so far described resembles that of Vasicek (1991), in that it is static. To introduce dynamics, we follow Lamb and Perraudin (2006) by allowing $X_{i,t}$ to be a p^{th}-order autoregressive stochastic process:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sigma\eta_t. \tag{5}$$

Here, η_t is assumed to be standard normal and independent of $\epsilon_{i,t}$.

As a normalization, we require that $Z_{i,t}$ has a unit unconditional variance which, in turn, implies that X_t has unit unconditional variance. In the Appendix, we derive the unconditional standard deviation of X_t when σ is unity. Setting σ to be the inverse of this quantity ensures that $Z_{i,t}$ is appropriately normalized.

Given the above setup, a dynamic process can be derived for the loss rate of a pool of obligors. The derivation of this generalizes the model of Lamb and Perraudin (2006) to the multi-lag case and allows for non-Gaussian latent variable distributions. A sketched proof is provided.

Substitution of equation (2) into (1) shows that default occurs when :

$$\sqrt{\rho}X_t + \sqrt{1-\rho}\epsilon_{i,t} \leq c_t. \tag{6}$$

The probability of observing k defaults out of n obligors, conditional on X_{t-1}, denoted $P(k,n)$, may be expressed as:

$$\begin{aligned} P(k,n) &= \binom{n}{k} \int_{-\infty}^{\infty} H\left(\frac{c_t - \sqrt{\rho}\left(\sum_{i=1}^{p}\phi_i X_{t-i} + \sigma\eta_t\right)}{\sqrt{1-\rho}}\right)^k \\ &\times \left[1 - H\left(\frac{c_t - \sqrt{\rho}\left(\sum_{i=1}^{p}\phi_i X_{t-i} + \sigma\eta_t\right)}{\sqrt{1-\rho}}\right)\right]^{n-k} d\Phi(\eta_t)\ . \end{aligned} \tag{7}$$

Adopting the change of variables:

$$s(\eta) \equiv H\left(\frac{c_t - \sqrt{\rho}\left(\sum_{i=1}^{p}\phi_i X_{t-i} + \sigma\eta_t\right)}{\sqrt{1-\rho}}\right)\ , \tag{8}$$

one obtains:

$$P(k,n) = \binom{n}{k} \int_0^1 s^k (1-s)^{n-k}\, dW(s)\ , \tag{9}$$

where

$$W(s) \equiv \Phi\left(\frac{\sqrt{1-\rho}H^{-1}(s) - c_t + \sqrt{\rho}\sum_{i=1}^{p}\phi_i X_{t-i}}{\sigma\sqrt{\rho}}\right). \tag{10}$$

As the number of obligors increases to infinity, $n \to \infty$, one may derive an expression for the fraction of the pool that defaults denoted θ:

$$\lim_{n\to\infty} \sum_{i=0}^{[n\theta]} P(i,n) = \int_0^1 \left(\lim_{n\to\infty} \sum_{i=0}^{[n\theta]} \binom{n}{i} s^i (1-s)^{n-i}\right) dW(s) \tag{11}$$

$$= \int_0^1 1(s<\theta)dW(s) = W(\theta) - W(0) = W(\theta)\ . \tag{12}$$

Hence, the loss distribution conditional on X_{t-1} is:

$$W(\theta_t) \equiv \Phi\left(\frac{\sqrt{1-\rho}H^{-1}(\theta_t) - c_t + \sqrt{\rho}\sum_{i=1}^{p}\phi_i X_{t-i}}{\sigma\sqrt{\rho}}\right)\ . \tag{13}$$

This implies that the transformed loss rate $\tilde{\theta}_t \equiv H^{-1}(\theta_t)$ conforms to the following Gaussian distribution:

$$\tilde{\theta}_t \equiv H^{-1}(\theta_t) \sim \mathrm{N}\left(\frac{c_t - \sqrt{\rho}\sum_{i=1}^{p}\phi_i X_{t-i}}{\sqrt{1-\rho}},\ \frac{\sigma^2\rho}{1-\rho}\right). \tag{14}$$

Hence, the transformed loss rate may be expressed as:

$$\tilde{\theta}_t = \frac{c_t - \sqrt{\rho}\sum_{i=1}^{p}\phi_i X_{t-i}}{\sqrt{1-\rho}} - \frac{\sigma\sqrt{\rho}}{\sqrt{1-\rho}}\eta_t. \tag{15}$$

where η_t is standard Gaussian. Alternatively, by substituting back in for the factor at time t, one may write the transformed loss rate as:

$$\tilde{\theta}_t = \frac{c_t - \sqrt{\rho}X_t}{\sqrt{1-\rho}}\ . \tag{16}$$

Rearranging equation (15), lagging and substituting, one may obtain:

$$\tilde{\theta}_t = \sum_{i=1}^{p}\phi_i\tilde{\theta}_{t-i} + \frac{1}{\sqrt{1-\rho}}\left(c_t - \sum_{i=1}^{p}\phi_i c_{t-i}\right) - \frac{\sigma\sqrt{\rho}}{\sqrt{1-\rho}}\eta_t. \tag{17}$$

2.2 Conditional Loss Distributions

To use the above model of loan losses in pricing applications, we must consider how the distribution of losses behaves conditional on recent factor realizations. At a given date, one may assume that the market observes a set of factor realizations and that the pricing of single name and multi-name credit derivatives is consistent with these realisations.

From a modeling viewpoint, this amounts to considering the process, X_t, at date 0 conditional on realizations before time 0, namely $(X_{p-1}, \ldots, X_{-1}, X_0)$. Conditioning on these realizations implies that current and future X_t's will have variances less than unity. The distribution of defaults for individual names at some date T will no longer be G but will instead will be a conditional distribution $G_{t,T}$. As T increases, the effect of the conditioning on the initial factors will become smaller, the variance of X_t will again approach unity and individual defaults will again be determined by G.

To see how to condition on past factor realizations, note that a pth-order AR process may be written in matrix form as a 1st-order AR process:

$$\overline{X}_t = F\overline{X}_{t-1} + \sigma\overline{\nu}_t. \tag{18}$$

where $\overline{X}_t \equiv (X_t, X_{t-1}, \ldots, X_{t-p+1})'$ and $\overline{\eta}_t \equiv (\eta_t, 0, \ldots, 0)'$ and where F is defined in equation (A9) in the Appendix.

Recursive substitution of this to time t leads to:

$$\overline{X}_t = F^t \overline{X}_0 + \sigma \sum_{j=0}^{t-1} F^j \overline{\nu}_{t-j}, \tag{19}$$

or in matrix form:

$$\begin{bmatrix} X_t \\ X_{t-1} \\ \vdots \\ \vdots \\ X_{t-p+1} \end{bmatrix} = F^t \begin{bmatrix} X_0 \\ X_{-1} \\ \vdots \\ \vdots \\ X_{-p+1} \end{bmatrix} + \sigma \sum_{j=0}^{t-1} F^j \begin{bmatrix} \eta_{t-j} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \tag{20}$$

The first row of this system gives X_t in terms of the factor values up to and including X_0:

$$X_t = \sum_{i=1}^{p} f_{1,i}^{(t)} X_{1-i} + \sigma \sum_{j=0}^{t-1} f_{1,1}^{(j)} \eta_{t-j}, \tag{21}$$

where $f_{1,i}^{(t)}$ is the $(1, i)$ element of F^t.

Now, substitution of this into (1) and (2) shows how the default of obligor i at time t is driven by the initial factor values and the compounded shocks:

$$\sqrt{\rho} \left(\sum_{i=1}^{p} f_{1,i}^{(t)} X_{1-i} + \sigma \sum_{j=0}^{t-1} f_{1,1}^{(j)} \eta_{t-j} \right) + \sqrt{1-\rho}\, \epsilon_{i,t} \leq c_t. \tag{22}$$

By conditioning on the information at time 0, we show how the distribution of the default quantile deviates from the unconditional case. Conditioning gives:

$$\sqrt{\rho}\sigma \sum_{j=0}^{t-1} f_{1,1}^{(j)} \eta_{t-j} + \sqrt{1-\rho}\, \epsilon_{i,t} \leq c_t - \sqrt{\rho} \sum_{i=1}^{p} f_{1,i}^{(t)} X_{1-i}. \tag{23}$$

Defining:

$$c_{0,t} = c_t - \sqrt{\rho} \sum_{i=1}^{p} f_{1,i}^{(t)} X_{1-i}, \tag{24}$$

gives:

$$\sqrt{\rho}\sigma \sum_{j=0}^{t-1} f_{1,1}^{(j)} \eta_{t-j} + \sqrt{1-\rho}\, \epsilon_{i,t} \leq c_{0,t}, \tag{25}$$

which, due to the shocks all being independent, is a $G_{0,T}$-distributed random variable where $G_{0,T}$ is the convolution of :

$$\Phi\left(\frac{x}{\sqrt{\rho\sigma^2\sum_{j=0}^{t-1}\left(f_{1,1}^{(j)}\right)^2}}\right) \quad \text{and} \quad H\left(\frac{\nu}{\sqrt{1-\rho}}\right) \ . \tag{26}$$

The probability conditional on information at date 0 of default at a future date t by a single name is now:

$$q_{0,t} = G_{0,t}\left(c_{0,t}\right). \tag{27}$$

Rearranging this equation shows how the default quantile c_t is altered by the conditioning information:

$$c_t = G_{0,t}^{-1}\left(q_{0,t}\right) + \sqrt{\rho}\sum_{i=1}^{p} f_{1,i}^{(t)} X_{1-i}. \tag{28}$$

When t is close to 0, the conditioning factor values are still dominant and these perturb the default quantile from its unconditional case. As $t \to \infty$ the term containing the initial factors becomes negligible and:

$$\lim_{t\to\infty}\sigma^2\sum_{j=0}^{t-1}\left(f_{1,1}^{(j)}\right)^2 \to 1 \qquad \lim_{t\to\infty} G_{0,t}(z) \to G(z)\ . \tag{29}$$

3 CDO Valuation

3.1 Tranche Valuation

To value tranches of a CDO, one may simulate the transformed loss rate process in equation (17), and then calculate the cumulative loss to the pool in each future period. Suppose that a structure pool has total exposure of unity and the loss rate in any future period is assumed to be θ_t. The cumulative loss rate is then defined as:

$$L_t = 1 - \prod_{i=1}^{t}\left(1-\theta_i\right)\ . \tag{30}$$

If the pool has been tranched in a particular way to create levels of subordination, then the loss to a specific tranche, denoted j, can be calculated using:

$$L_{t,j}^{tr} = \min\left(\max\left(\left(1-\gamma\right)L_t - A_{1,j}, 0\right), A_{2,j} - A_{1,j}\right), \tag{31}$$

where $A_{1,j}$ and $A_{2,j}$ are the attachment and detachment points respectively and γ is the recovery rate.

To value a tranche, one must consider two sets of cash flows. The tranche holder offers protection against losses in a given range defined by the attachment and detachment points. Payments to cover these losses are termed the default leg payments. On the other hand, the tranche holder receives from the purchaser of protection premiums proportional to the un-defaulted principal at any given moment. These payments are termed the premium leg payments. The value of the tranche is then the difference between the expected discounted cash flows of the premium and default legs.

To make this more precise, suppose the time horizon of the tranche is split into k discrete periods starting from the time of valuation, $t = 0$, until the maturity of the CDO, $T = t_n$, and that default can occur in any one of these time intervals. The expected discounted value of the default leg cash flows is:

$$D_{0,j} \;=\; E\left[\sum_{k=1}^{n} B_{0,t_k}\left(L^{tr}_{t_k,j} - L^{tr}_{t_{k-1},j}\right)\right] \;=\; \sum_{k=1}^{n} B_{0,t_k}\left[E\left(L^{tr}_{t_k}\right) - E\left(L^{tr}_{t_{k-1}}\right)\right] \;. \quad (32)$$

Here, $B_{0,t}$ price at date 0 of a pure discount bond paying \$1 for sure at date t.

Assume the premium leg of the tranche is paid discretely in each of the m periods. If the tranche premium is ω, then the premium leg is given by:

$$P_{0,j} \;=\; E\left[\omega \sum_{k\in m} B_{0,t_k}\left(A_{2,j} - A_{1,j} - L^{tr}_{t_k}\right)\right] \;=\; \omega \sum_{k\in m} B_{0,t_k}\left[A_{2,j} - A_{1,j} - E\left(L^{tr}_{t_k}\right)\right]. \quad (33)$$

Knowing the expected losses on a tranche at the different future dates is, therefore, enough to value the tranche. The expected losses may be estimated by simulating the dynamic loss process of the pool, (17), and then taking expectations of (31). In our analysis, we assume a constant recovery rate of 40%.

3.2 Market Calibration

In the market, for standard tradable synthetic structures, quotes are available for each tranche within a structure. There are also multiple maturities, or tenors. As our model is a dynamic one, it can be used to fit consistently prices or spreads for tranches of different tenors.

The first step in calibrating the model is to infer the $q_{0,j}$ from Credit Default Swap (CDS) spreads at date 0. In our pricing, we assume that the pool consists of identical borrowers and hence we wish to extract a single set of default probabilities. We could proceed by extracting default probabilities for each of the individual names in the pool using CDS spread quotes for those same names and then take a value-weighted average taken to obtain a single set of probabilities. An alternative is to regard the index spreads i.e., the spreads for a non-tranched vanilla synthetic CDO as comparable to the spreads on an individual name CDS contract and then to infer the default probabilities using pricing formula for a single-name CDS. In what follows, we take the latter approach.

To see how one may extract default probabilities from CDS spreads, suppose that, conditional on information at date 0, we suppose as before that $q_{0,t}$ denotes the probability that an individual obligor defaults in period t having survived until $t-1$. The probability that the obligor defaults at some time between dates 0 and t is:

$$Q_{0,t} = 1 - \prod_{j=1}^{t} (1 - q_{0,j}) \tag{34}$$

For a CDS contract with a notional value of unity, the fair spread on un-defaulted notional denoted ζ satisfies:

$$\zeta \sum_{j=1}^{t} (1 - Q_{0,j})\, B_{0,j} \quad = \quad (1-\gamma) \sum_{j=1}^{t} (Q_{0,j} - Q_{0,j-1})\, B_{0,j} \ . \tag{35}$$

Implicitly, this equation depends on the $q_{0,j}$. Given a set of CDS spreads, one may infer the $q_{0,j}$ by minimizing the squared difference between the actual quotes and the quotes implied by equation (35). In doing this, we assume that the default probabilities $q_{0,t}$ are constant for dates t between the maturity dates of the synthetic CDO contracts we ultimately wish to price, namely 5, 7 and 10-year maturities.

Given estimates of the $q_{0,t}$, one may infer the parameters of the loss rate process from the spreads on the synthetic CDO tranches. The parameters of the loss rate process to be inferred are (i) the common factor weight ρ, (ii) the unobserved common factor autoregressive parameters, ϕ_i, for $i = 1, \ldots, p$, where p is the number of lags and (iii) the parameters, if any, of the idiosyncratic distribution H denoted ν. (In the case of a Gaussian, H is the standard normal distribution function and has no parameters. In the case of other distributions we consider below, ν will include one or more parameters.)

To infer the loss rate parameters, we evaluate the tranche spreads implied by a set of parameters and then iterate using an optimization routine to minimize the sum of squared differences between the observed and model-implied tranche spreads. These differences are expressed in the quadratic objective function of the optimization as a ratio to the observed spread.

Note that our approach of first extracting the $q_{0,t}$ ensures that the index spread is precisely fitted. An alternate approach would be to fit the index spread as part of the more general fitting of the loss rate parameters. This would give more flexibility in the fitting procedure and improve the accuracy of the implied tranche spread.

3.3 Calibration of a Static Model

Before discussing the fit of our model to data, we present results for a static loss distribution model similar to current market practice. Vasicek (1991) proposed a simple closed form loss distribution for a pool of credit exposures. His loss distribution is a function of a factor correlation parameter and the default probability over the given horizon.

It is common practice to infer the default probabilities from the CDS index spread assuming a constant hazard rate. The correlation parameter is then extracted from spread data for a given tranche with a particular tenor. In theory, if the model were correct, the same correlation parameter would accurately fit the prices of tranches with different levels of subordination. When correlations are extracted from spreads on different tranches, however, one generally finds a "correlation smile", with the correlation parameter appearing higher for junior and senior tranches and lower for mezzanine tranches.

Figure 1 shows the correlation smile implied by 5 and 7-year tenor CDX tranche spreads, averaged over weekly observations from June 2006 until December 2007, based on a simple Vasicek loss distribution. This plot shows that to fit each market quote perfectly requires that one associate different factor correlations with the different tranches. Based on this, one may argue that a more richly parameterized loss distribution that can fit multiple tranches in a consistent fashion is called for.

The first row of Panels A and B of Table 1 show the average absolute and percentage differences between the implied and observed CDX spreads for different subordination levels or tranches. The implied spreads are obtained by the factor correlation that optimizes the spread fit for the different tranches. The averages are accumulation of all dates and maturities (5, 7, and 10-year) at the given subordination. The spreads implied by the model fit poorly. This again points to the need for a model that can be calibrated across a whole structure or even across multiple maturities.

3.4 Calibration of Dynamic Models

Up to now, we have presented our dynamic loan loss distribution in general terms without specifying the distribution, H, for the idiosyncratic shocks, $\epsilon_{i,t}$. In what follows, we shall calibrate our model for three different H distributions.

One might think that H would have little influence on aggregate losses as idiosyncratic randomness should be diversified away in a large portfolio. However, the form of H affects the model in two ways. First, the loss rate distribution depends on default cut-off points or quantiles $c_{0,t}$ which are obtained from the $q_{0,t}$ by inverting the convolution of H and the Gaussian common factor.

Second, H determines the relation between the "transformed loss rate", $\tilde{\theta}_t$, that follows a given autocorrelated stochastic process in our framework, and actual losses, θ_t, in that

$$\theta_t = H(\tilde{\theta}_t). \tag{36}$$

We employ three specifications of H, namely:

1. Using a standard Gaussian random variable for the idiosyncratic shock, $\epsilon_{i,t}$, H is then a standard normal distribution. The default driver, $Z_{i,t}$ is then the convolution of two Gaussians which is simply another Gaussian. By definition a standard normal random variable has unit variance. Hence, no scaling of the idiosyncratic term is required to enforce this condition.

2. If the idiosyncratic shocks follow a student-t distribution, then the default driver $Z_{i,t}$ will exhibit more fat-tailed behavior. The distribution of $Z_{i,t}$ will be the convolution of a Gaussian distribution $N(0, \rho)$ and of a random variable which once re-scaled has a t-distribution with ν degrees of freedom. By "re-scaled", we mean that the random variable is scaled by $\sqrt{1-\rho}\sqrt{(\nu-2)/\nu}$ so that it has a variance of $1-\rho$. (Note here that the variance of a random variable that is t-distributed with ν degrees of freedom is $\nu/(\nu-2)$.)

In this case, the inverse of the G distribution does not have a closed form expression so a numerical solution must be obtained. We employed a root finding algorithm to perform the inversion.

Note that with this H distribution, an additional parameter, namely ν, is available for fitting the tranche spreads.

3. Finally, we consider a case in which H is a Gaussian mixture. In this case, the idiosyncratic shock may be viewed a random draw from one of two different Gaussian distributions. Like the Student-t distribution, this form of H implies fat tailed behavior for the default driver $Z_{i,t}$.

   We assume that $\epsilon_{i,t}$ has the following distribution:

$$H\left(\epsilon_{i,t}\right) = \lambda\Phi\left(\frac{\epsilon_{i,t}}{\sigma_\epsilon}\right) + \left(1-\lambda\right)\Phi\left(\frac{\epsilon_{i,t}}{\sigma_m\sigma_\epsilon}\right). \quad (37)$$

   If σ_ϵ equals unity, this distribution would have the variance:

$$E\left(\epsilon_{i,t}^2\right) = \lambda + \left(1-\lambda\right)\sigma_m^2 \ . \quad (38)$$

   Hence, we must set:

$$\sigma_\epsilon \ = \ \sqrt{\lambda + \left(1-\lambda\right)\sigma_m^2} \ . \quad (39)$$

   Then G is given by the convolution of $N\left(0,\rho\right)$ and of a Gaussian mixture random variable scaled by $\sqrt{1-\rho}$. Again, as with the t distribution, to find G^{-1}, we use a numerical inversion procedure. With this H distribution, additional parameter λ and σ_m are available for improving the fit of the tranche spreads.

4 Results

We calibrate the model for CDX tranches with 5, 7 and 10-year tenors and with attachment-detachment points 0-3%, 3-7%, 7-10%, 10-15%, and 15-30% and the CDX index spread. Our sample includes observations of the contracts weekly from June 2006 to December 2007.

We fit the data for several different model specifications. Specifically, we fit six specifications with different H functions (idiosyncratic shock distributions): (i) Gaussian, (ii) Student's t, and (iii) Gaussian mixture, and with different numbers of factor lags: (a) one and (b) two lags.

Table 1 summarizes the accuracy of the fits by reporting the average absolute differences in basis points and the average absolute percentage differences between the actual and model-implied spreads for the five different attachment-detachment point ranges. (The model fit is performed recursively in that it matches the index spread perfectly before we choose model parameters to fit the other spreads. So we do not report accuracy measures for the index spreads.)

The model fits the spreads with reasonable accuracy for the full range of tenors and attachment-detachment points. The 0-3% spreads are least well fitted as measured by absolute differences but this clearly reflects the substantial size of these spreads and the fit as measured by percentage differences is reasonably accurate.

The best specification appears to be the mixture model with one factor lag. The absolute errors and the percentage errors all seem reasonable in this case.

Figure 2 provides graphical summaries of the accuracy of the fit in the case of the one-lag, normal mixture model. Each of the six panels in the figure shows the actual spreads for three different tenors (5, 7 and 10-years) with solid lines, and the spreads implied by the fitted model with dotted lines.

In the case of the index spread, the solid and dotted lines coincide for the reasons already explained and hence only a single line for each of the three tenors is visible. In the case of the other attachment-detachment point ranges, one may see that the dotted line tracks the solid lines quite well throughout the sample period.

Figure 3 shows, for each week in the sample period, the model parameters that the algorithm has come up with in fitting the tranche spreads. The four parameters in question are (i) the factor correlation, ρ, (ii) the common factor autoregressive parameter, ϕ, (iii) the weight between the two Gaussian distributions in the mixture, λ, and (iv) the volatility of the second distribution in the mixture, σ_m.

An interesting point to note is the fact that the correlation parameter, ρ is highly stable through the sample period, ranging from 60 to 70 percent for almost all the dates. This is in contrast to what one finds in the case of the simple static, single risk factor model. On the other hand, the autoregressive factor parameter moves about considerably over the sample period, ranging from negative values up to 80 percent at the end of the sample period.

The intuition this suggests is that, in valuing CDX tranches, market participants frequently revise not the correlation between individual defaults but the degree to which they think credit cycle shocks are persistent. A high level of the autoregressive parameter implies that when a shock occurs, it's impact is felt for multiple periods and so will affect cumulative pool losses to a greater degree.

5 Conclusion

This paper generalizes in a simple and transparent manner the most standard and widely employed valuation model for synthetic CDOs in such a way that it consistently prices tranche spreads for multiple subordination levels and maturities. The resulting model is fully dynamic and hence may be used for hedging portfolios of synthetic CDO exposures over time in a rigorous fashion.

Our model sheds interesting light on the loss distribution implicit in market spreads and how this changes over time. It is commonly thought that market values are driven by fluctuations in the market's perceptions of default correlations. In our richer parameterization, correlation parameters appear relatively stable over time while the implied parameter that measures the persistence of credit shocks moves around substantially as tranche spreads evolve over time.

Future research ideas suggested by our study include (i) generalizations in which factors driving defaults display GARCH-type properties, and (ii) simultaneous empirical investigation of tranche pricing and of the stochastic evolution of individual CDS spreads.

Appendix

A Unit Variance Scaling

First, using lag operators, setting $\sigma = 1$, equation (5), can be written as:

$$X_t B(L) = \eta_t \ , \tag{A1}$$

for a lag operator $B(L)$ defined as:

$$B(L) = 1 - \phi_1 L - \cdots - \phi_p L^p. \tag{A2}$$

The lag polynomial can be factorized into the form:

$$B(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) \ldots (1 - \lambda_p L) \ . \tag{A3}$$

We assume that the roots of the lag operator, λ_i lie outside the unit circle. This implies that the lag operator can be inverted and the factor process may therefore be written as a weighted sum of lagged innovations:

$$X_t = B(L)^{-1} \eta_t \ . \tag{A4}$$

To derive the coefficients on lagged innovations, start by expanding the lag polynomial through partial fractions to obtain:

$$\begin{aligned} X_t &= \frac{c_1}{1 - \lambda_1 L}\eta_t + \ldots + \frac{c_p}{1 - \lambda_p L}\eta_t && \text{(A5)} \\ &= c_1 \sum_{j=0}^{\infty} \lambda_1^j \eta_{t-j} + \ldots + c_p \sum_{j=0}^{\infty} \lambda_p^j \eta_{t-j} && \text{(A6)} \\ &= \sum_{i=1}^{p} c_i \sum_{j=0}^{\infty} \lambda_i^j \eta_{t-j} \ , && \text{(A7)} \end{aligned}$$

where:

$$c_i = \frac{\lambda_i^{p-1}}{(\lambda_i - \lambda_1) \ldots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \ldots (\lambda_i - \lambda_p)} \ . \tag{A8}$$

To obtain the λ_i's, note that, by definition, the eigenvalues of the matrix:

$$F \equiv \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{A9}$$

are the solutions of the polynomial:

$$z^p - \phi_1 z^{p-1} - \cdots - \phi_{p-1} z - \phi_p = 0 \ . \tag{A10}$$

Setting $z = 1/\lambda$, we obtain the

$$1 - \phi_1 \lambda - \cdots - \phi_{p-1} \lambda^{p-1} - \phi_p \lambda^p = 0 \ . \tag{A11}$$

Hence, the λ_i in the factorization of the lag operator in equation (A3) are simply the inverses of the eigenvalues of F. Our assumption above that the roots of the lag operator lie outside the unit circle is equivalent to assuming that the eigenvalues of F lie inside the unit circle.

Given the representation (A5), it is a straight forward step to calculate the unconditional variance of the process. This is:

$$\text{Variance}\,(X_t) = E\left[\left(\sum_{i=1}^{p} c_i \sum_{j=0}^{\infty} \lambda_i^j \eta_{t-j}\right)^2\right] = \sum_{j=0}^{\infty}\left(\sum_{j=1}^{p} c_i \lambda_i^j\right)^2 \ , \tag{A12}$$

where we have reversed the order of the summations and used the temporal independence of the shocks.

To ensure that X_t has a unit unconditional variance, each period, one may set σ to the inverse of $\sqrt{\text{Variance}\,(X_t)}$ to obtain:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \frac{1}{\sqrt{\sum_{j=0}^{\infty}\left(\sum_{j=1}^{p} c_i \lambda_i^j\right)^2}}\eta_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sigma \eta_t \ . \tag{A13}$$

References

Brigo, D., A. Pallavicini, and R. Torresetti (2007): "Default Correlation, Cluster Dynamics and Single Names: The GPCL Dynamical Loss Model," Working paper, Banca IMI, Milan, Italy.

Burtschell, X., J. Gregory, and J.-P. Laurent (2005): "A Comparative Analysis of CDO Pricing Models," Working paper, BNP Paribas.

Chapovsky, A., A. Rennie, and A. Tavares (2006): "Stochastic Intensity Modelling for Structured Credit Exotics," Discussion paper, Merrill Lynch International.

Das, S. R., D. Duffie, N. Kapadia, and L. Saita (2007): "Common Failings: How Corporate Defaults are Correlated," *Journal of Finance*, 62(1), 93–117.

Davis, M., and V. Lo (2001): "Infectious Defaults," *Quantitative Finance*, 01(4), 382–387.

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

Giesecke, K., and L. Goldberg (2005): "A Top Down Approach to Multi-Name Credit," Working paper, Stanford University and MSCI Barra, Inc, Stanford, CA.

Hull, J., and A. White (2004): "Valuation of a CDO and nth to Default CDS Without Monte Carlo Simulation," *Journal of Derivatives*, 12(2), 8–23.

——— (2007): "Dynamic Models of Portfolio Credit Risk: A Simplied Approach," Working paper, University of Toronto, Toronto.

Lamb, R., and W. Perraudin (2006): "Dynamic Default Rate Distributions," Working paper, Imperial College, London.

Laurent, J., and J. Gregory (2005): "Basket Default Swaps, CDOs and Factor Copulas," *Journal of Risk*, 7(4).

Li, D. X. (2000): "On Default Correlation: A Copula Approach," *Journal of Fixed Income*, 9, 43–54.

McGinty, L., and R. Ahluwalia (2004): "Introducing Base Correlation," Working paper, J.P. Morgan.

Schonbucher, P., and D. Schubert (2001): "Copula-Dependent Default Risk in Intensity Models," Discussion paper, Department of Statistics, Bonn University.

Schonbucher, P. J. (2002): "Taken to the Limit: Simple and not so simple Loan Loss Distributions," Working paper, Bonn University.

Schonbucher, P. J. (2006): "Portfolio Losses and the Term Structure of Loss Transition Rates: A New Methodology for the Pricing of Portfolio Credit Derivatives," Working paper, Eth University of Zurich, Zurich.

Sidenius, J., V. Piterbarg, and L. Anderson (2006): "A New Framework for Dynamic Credit Portfolio Loss Modelling," Working paper, Royal Bank of Scotland and Barclays Capital and Bank of America.

Vasicek, O. (1991): "Limiting Loan Loss Probability Distribution," Unpublished mimeo, KMV.

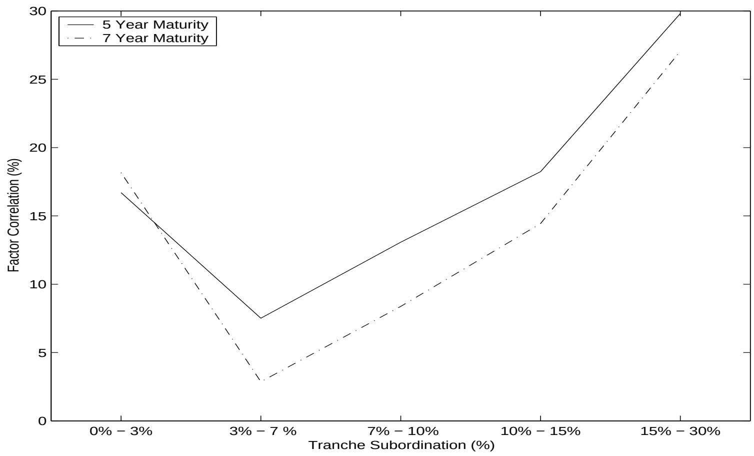

Figure 1: Plot of the average implied correlation estimates for different subordination levels for CDX for 5 and 7-year maturities. The implied correlation estimates are obtained using a static model over the period of June 2006 to December 2007.

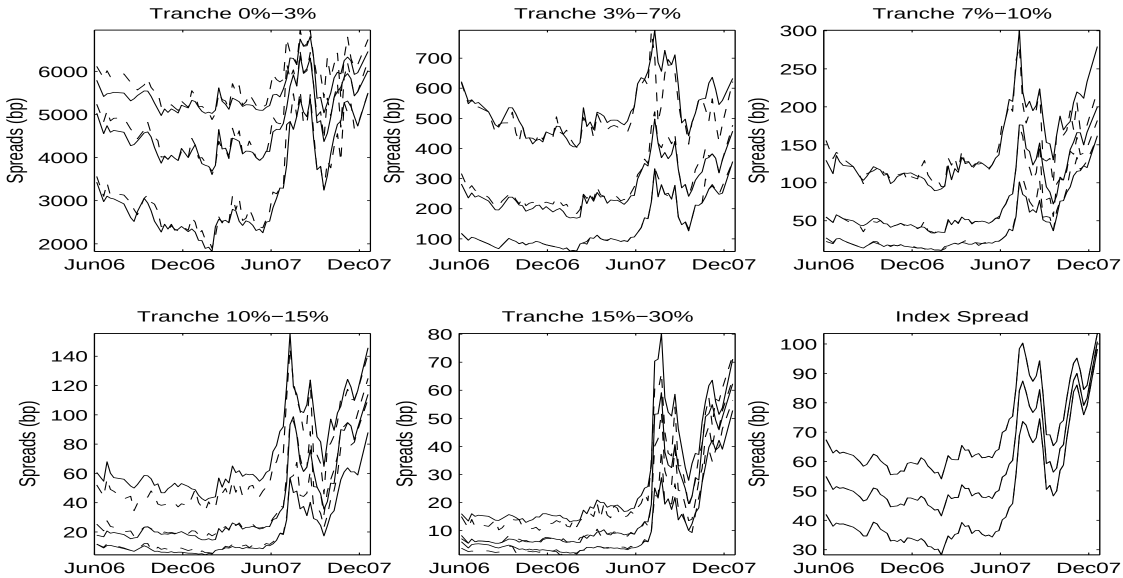

Figure 2: Plots of market and model implied tranche spreads for the CDX tranches for 5, 7 and 10-year maturities over the period of June 2006 to December 2007. A Gaussian mixture distribution was assumed for the idiosyncratic distribution H with one autoregressive lag in the common factor. Each panel shows the spreads for all maturities at a particular tranche subordination. The solid lines are the market quoted spreads. The dashed lines are the model implied spreads.

Table 1: Absolute and percentage differences of market to model implied tranche spreads for the CDX synthetic CDO over the period of June 2006 to December 2007. The first row of panels A and B present the results for the static model with a Gaussian distribution. The other rows present the results for the dynamic model with different distributions for the idiosyncratic distribution H: Normal, Student-t and Gaussian mixture distribution. For each model and distribution, the results for one and two autoregressive lags are presented. Each entry in the table is the accumulation of all dates and maturities at the given subordination. Thus, each entry is the average of the 5, 7, and 10-year maturities for all data points at the given level of protection.

Subordination	0% - 3%	3% - 7%	7% - 10%	10% - 15%	15% - 30%
Panel A: Absolute Differences (bp)					
Static	427.16	119.63	55.97	23.30	10.03
Normal 1 Lag	299.67	14.43	12.23	6.48	11.53
Normal 2 Lags	288.74	14.32	12.30	6.47	11.55
Student-t 1 Lag	428.56	19.29	8.85	4.89	3.90
Student-t 2 Lags	381.88	19.77	9.63	3.28	4.44
Mixed 1 Lag	184.33	16.89	6.59	5.72	2.52
Mixed 2 Lags	225.99	19.58	7.74	7.24	2.94
Panel B: Percentage Differences (%)					
Static	9.48	41.83	39.39	50.21	71.81
Normal 1 Lag	7.40	4.33	11.61	20.04	69.52
Normal 2 Lags	7.13	4.34	11.71	19.87	69.57
Student-t 1 Lag	9.51	5.40	11.33	10.67	30.93
Student-t 2 Lags	8.66	5.13	11.83	7.79	32.70
Mixed 1 Lag	4.34	4.68	6.26	14.85	13.06
Mixed 2 Lags	5.05	5.53	6.63	15.97	13.55

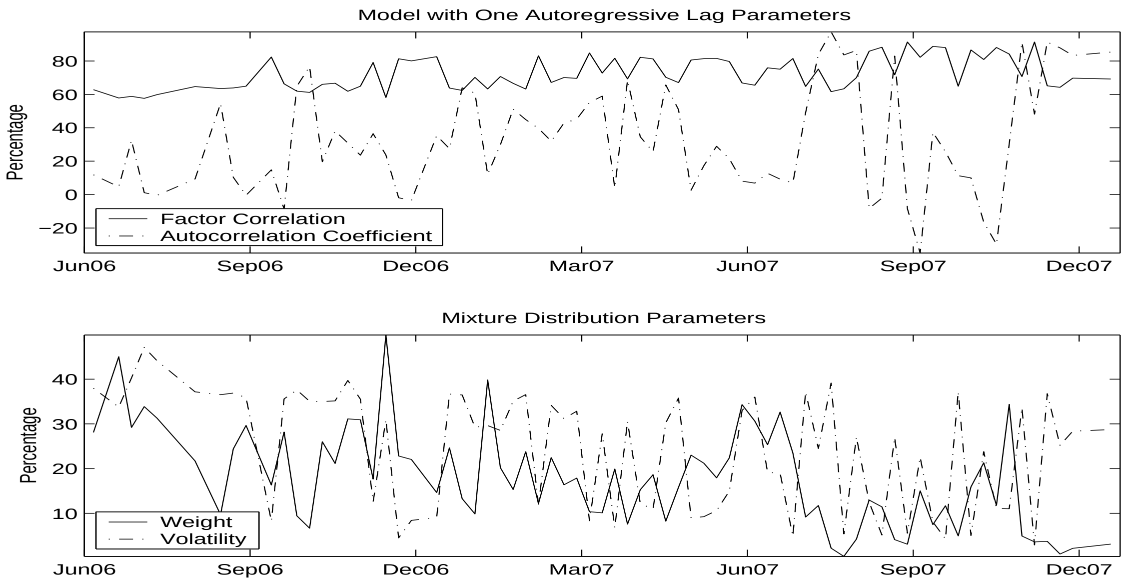

Figure 3: Parameter estimations using a Gaussian Mixture H distribution and one autoregressive lag in the common factor. The data calibrated to is the CDX tranches for 5, 7 and 10-year maturities over the period of June 2006 to December 2007. The top panel shows the parameters independent of the choice of idiosyncratic distribution H, so the factor correlation, ρ, and the common factor autoregressive coefficient, ϕ. The bottom panel shows a time-series of the parameters specific to a Gaussian mixture model as choice for H. These are the weight between the two Gaussian distributions, λ, and the volatility of the second distribution, σ_m.