Credit Derivatives: CDX Tranches

Goals:

- Examine CDX tranches as an example of correlation instruments
- Find a way to price them

Relevant literature:

- Dominic O'Kane, "Modelling Credit Derivatives": Chapters 12,13,16,18,19,20.

CDOs

- Collateralised debt obligations (CDOs) are securities backed by collections of debt instruments (the collateral).
- These debt instruments can be bonds, loans, mortgages, credit card debt, car loans, etc.
- The cashflows produced by the collateral are sliced into derivative instruments according to rules, which delineate how these cashflows will be distributed among the derivatives.
- Precise rules that guide the distribution of cashflows are known as waterfall.
- The result is that the CDO securities are designed to have varying risk profiles to appeal to different investor classes.

Synthetic CDOs

- The collateral can also be a collection of CDS or other, so called unfunded instruments.
- Then the CDOs are called synthetic, since the underlying CDS synthesize risk exposure similar to bonds without having to actually buy them.
- In particular, a CDS index or its basket can serve as the collateral for a CDO.
- Today we are going to look at one with a simple waterfall.

Synthetic CDOs

Recall that in the CDS index (we'll call it CDX, for short) there are cashflows in both directions.

- Premium leg of the index is essentially distributed to pay individual premia in the basket.
- Contingent leg in return pays out "insurance claims" when individual defaults occur and those cashflows are distributed evenly among long-protection holders.
- But these cashflows don't need to be distributed evenly: they can cut into slices of different priorities.

Tranches

- As time goes on and the CDS inside a CDX default, we can denote their cumulative losses up to time t by $L(t)$ with $L(t)$ scaled to be in $[0, 100]$.
- We define a tranche with attachment point K_1 and detachment point K_2 as a call spread on the index loss $L(t)$. More explicitly, define tranche loss $L(t, K_1, K_2)$ to be the function of $L(t)$ such that

$$L(t, K_1, K_2) = \begin{cases} 0 & L(t) \leq K_1 \\ 1 & L(t) \geq K_2 \end{cases} \tag{1}$$

and linear between K_1 and K_2. This can also we written as

$$L(t, K_1, K_2) = \frac{(L(t) - K_1)^+ - (L(t) - K_2)^+}{K_2 - K_1} \tag{2}$$

Tranches

- These tranches must cover the full range of possible losses, i.e. up the the max of L, i.e. 100.
- A typical collection is $[0,3]$, $[3,7]$, $[7,10]$, $[10,15]$, $[15,30]$, $[30,100]$.
- The bottom, in this case $[0,3]$ is called equity, the next one is mezzanine, next two are junior, then senior, and the top one is called super-senior.

Tranches

- For example, suppose we start with 125 CDS in an index and then some CDS that has recovery $R = 40\%$ defaults at time t_1. Then the contingent payment of the CDS is $1 - R = 0.6$, and since the CDS comprises 1/125th of the index, the contingent payment as a fraction of the index is

$$\frac{1}{125} 0.6 = 0.0048 \tag{3}$$

or 0.48%. Therefore, $L(t) = 0$ for $t < t_1$ and $L(t_1) = 0.48$.

- It follows that

$$L(t_1, 0, 3) = \frac{(0.48 - 0)^+ - (0.48 - 3)^+}{3 - 0} = \frac{0.48}{3} = 0.16 = 16\%. \tag{4}$$

Notice that $L(t_1, 3, 7) = 0$, and will continue to be zero until the equity tranche is exhausted.

CDX Tranches

- Thus, an investor who is long-protection on the $[0, 3]$ will receive 16% of the maximum after just one default. It stands to reason that the premium leg on this tranche is very high.
- Higher tranches, which are not activated until more losses accumulate have correspondingly lower premiums.
- But how to price them more precisely?
- First note that the equity tranche will suffer losses (i.e., holder of long-protection will benefit) if any of the 125 individual underlying CDS defaults. Therefore, the equity tranche should be more expensive if the underlyings are independent.
- Conversely, suppose all the underlyings are perfectly correlated. Then if one defaults, then they all do. This should do wonders to the price of the super-senior tranche.

Tranches are correlation instruments

- For this reason, tranches are referred to as *correlation* products.
- Therefore, the pricing model for tranches must hold some kind of correlation information.
- Moreover, this correlation cannot be historical, but should be computable from tranches prices and vice versa. Remember the difference between implied and physical densities? The same applies to correlations, since they are descriptors of densities.

CDX Tranches

Since tranches are call spreads on $L(t)$, in order to price them, one needs to somehow get at the distribution of $L(t)$. We drop the subscript t for simplicity; distribution of L still depends on time.

Let B_j be a Bernoulli random variable that describes the default of the jth CDS in a CDX up to time t, incurring the loss l_j with probability p_j.

$$B_j = \begin{cases} 0 & 1 - p_j \\ l_j & p_j \end{cases} \tag{5}$$

The probability p_j can be extracted from the individual CDS survival curve: it is the probability that the jth issuer defaults before t. Then

$$L = \sum_{j=1}^{N} B_j \tag{6}$$

where N is 100 or 125.

CDX Tranches

Note that if you are only interested in expected losses of CDX

$$E[L] = \sum_{j=1}^{N} E[B_j] = \sum_{j=1}^{N} l_j p_j \tag{7}$$

you already have all you need to know. This information is sufficient for pricing the CDX itself.

In order to price the tranches, you need the whole distribution of L, and for that you need to know something about interdependence between the B_js.

CDX Tranches

One can recast each B_j in terms of a normally random variable with mean 0 and variance 1, X_j, by defining a critical value x_j^{crit} so that:

$$P(B_j = l_j) = p_j = P(X_j \leq x_j^{crit}) \tag{8}$$

This formulation is inspired by the idea that as soon as a value of a corporation drops to a certain level, namely, x_j^{crit}, it will exercise its bankruptcy option and trigger the default.

CDX Tranches

To model dependence among the B_js, or, equivalently, X_js, we can model each X_j as a sum of two normal $(\mu = 0, \sigma = 1)$ random variables M and ϵ_j:

$$X_j = \sqrt{\rho}M + (\sqrt{1-\rho})\epsilon_j \tag{9}$$

The parameter ρ is called correlation (even though it isn't really). The variable M, the common part, represents the market and ϵ_j represents the idiosyncratic part. All ϵ_js are assumed to be independent of each other and of M.

CDX Tranches

- Note that if $\rho = 1$, all X_js are the same and if $\rho = 0$, all X_js are independent.
- Regardless of the value of ρ, when conditioned on M, X_j are independent.
- When X_js are independent, the corresponding B_js are also independent. (why?)

Now we are ready to tackle the distribution of L:

$$
\begin{aligned}
P(L \leq a) &= \int P\left((L\,|_{M=m}) \leq a\right) \phi_M(m)\, dm && (10) \\
&= \int P\left(\sum_{j=1}^{N} (B_j\,|_{M=m}) \leq a\right) \phi_M(m)\, dm && (11)
\end{aligned}
$$

CDX Tranches

Taking derivative of (10) wrt to a, we get

$$\phi_L(a) = \int \phi_{(L|M=m)}(a)\, \phi_M(m)\, dm \tag{12}$$

$$= \int \phi_{(\sum B_j|_{M=m})}(a)\, \phi_M(m)\, dm \tag{13}$$

So we need to compute distribution of $\sum B_j \;|_{M=m}$.

First, on the next slide, we compute distribution of each individual $B_j \;|_{M=m}$.

CDX Tranches

Each $B_j \mid_{M=m}$ is Bernoulli, and the default probability conditioned on $M = m$ is

$$\begin{aligned}
p_j(m) &:= P((B_j \mid_{M=m}) = l_j) = P((X_j \mid_{M=m}) \leq x_j^{crit}) \\
&= P((\sqrt{\rho}M + (\sqrt{1-\rho})\epsilon_j \mid_{M=m}) \leq x_j^{crit}) \\
&= P(\sqrt{\rho}m + (\sqrt{1-\rho})\epsilon_j \leq x_j^{crit}) \\
&= P\left(\epsilon_j \leq \frac{x_j^{crit} - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) \\
&= \Phi\left(\frac{x_j^{crit} - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) \qquad (14)
\end{aligned}$$

CDX Tranches

Therefore, distribution of each $B_j \mid_{M=m}$ is

$$B_j \mid_{M=m} \quad = \quad \begin{cases} 0 & 1 - p_j(m) \\ l_j & p_j(m) \end{cases} \tag{15}$$

and they are independent of each other for every value of m.

Therefore, the distribution of their sum is the convolution of their N individual distributions.

CDX Tranches

This convolution can be computed with $(N+1)$ FFTs and N multiplications of vectors of length N_{mesh} for a total of

$$O((N+1)N_{mesh}\log(N_{mesh})+N_{mesh}N)=O(NN_{mesh}\log(N_{mesh}))$$

operations, where N_{mesh} is the number of discretization points used for integration, which is quite doable.

Integration against $\phi_M(m)\,dm$ presents no challenges, giving us the total cost of

$$O(NN_{mesh}^2\log(N_{mesh})) \tag{16}$$

if we use N_{mesh} points for the $\phi_M(m)\,dm$ integral. This is a lot better than the direct method, which would cost

$$O((N_{mesh})^{N-1}). \tag{17}$$