CreditSpectrumCorp

Data Science Internship Project

Mapping CLO-Asset Credit Spreads to Credit Ratings

Introduction

Many modern CLO's contain loan-assets that do not carry a public credit rating by any rating agency. The WFE needs at least one credit rating for each loan-asset in order to assign ratings to the securities on the other side of the CLO's balance sheet. Without a rating, the loan's credit behavior cannot be simulated. As a result, when the time comes for us to rate CLO's with private asset-ratings only, and there are many such CLO's, we will have to find a way to map each loan's credit spread to its implied credit rating, if only because spreads are directly, albeit sluggishly related to credit ratings. In addition, the mapping exercise needs to be done in real time because WFE will be rerating CLO-tranches on a monthly basis and automatically.

The following notes describe a suitable algorithm based on Lebesgue integration. Once implemented, it will enable us to map the credit spread of any CLO loan-asset to a rating corresponding more or less to its credit spread based on a large training set of such loan-assets when both data elements were disclosed. The required statistical theory was presented in a lecture course at the University of Toulouse (France) and later published in a set of classroom notes we stumbled upon by sheer luck, for we have not been able to locate a published article with same technique. By all appearances however, this holds water. The European Union currently makes use of this technique.

Computation of the Correlation Between the Credit Spread and the Letter-Grade Credit Rating

Step 1

The first step consists of computing the implied correlation between credit spreads, which are always available, and alphanumeric credit ratings (e.g. Baa3, A2, etc.) To do so, we use modified Lebesgue integration.

Assume we are given a database of n CLO loan assets drawn randomly from the ensemble of US corporate CLO's. Assume further that these loan-assets are partitioned across $l \in [1, m]$ alphanumeric rating levels (notches) ranging from Aaa to Caa3. Note that in most cases, the credit rating of loan-assets in a CLO will not span anywhere near the entire credit rating distribution but will be concentrated in the sub-investment grade notches, i.e. below Baa3. This is usually referred to as SME lending.

Without loss of generality, we can assume that each rating notch contains a finite set $i \in [1, n_l]$ of loan-assets, each with its own credit spread s_i. The n loans in the whole database thus satisfy $n = \sum_{l=1}^{m} n_l$. As already highlighted, in any CLO credit spreads are given. In theory, all credit spreads at a given rating level ought to be either identical or nearly so, but that will commonly not happen for a variety of reasons.

The difference between the average asset-spread at a given credit quality, as revealed by agency ratings, and the actual spread of a single loan-asset is referred to as its *liquidity* spread. In theory, all liquidity spreads should thus be zero, but they're not by a long shot. Therefore, liquidity-spread volatility is an unbiased measure of market efficiency, or inefficiency perhaps.

To compute the implied spread-rating correlation, proceed as follows:

1. Compute notch-wise average credit spreads $\overline{s_l} = \frac{1}{n_l}\sum_{i=1}^{n_l} s_i \,, l \in [1, m]$
2. Compute notch-wise spread variances $\sigma_l^2 = \frac{1}{n_l}\sum_{i=1}^{n_l}[s_i - \overline{s_l}]^2 \,, l \in [1, m]$
3. Compute the global average credit spread $\bar{s} = \frac{1}{n}\sum_{i=1}^{m} n_l \overline{s_l}$
4. Compute the global credit spread variance $\sigma_S^2 = \frac{1}{n}\sum_{l=1}^{m} n_l[\overline{s_l} - \bar{s}]^2 + \frac{1}{n}\sum_{l=1}^{m} n_l \sigma_l^2$
5. Define $\sigma_E^2 \equiv \frac{1}{n}\sum_{l=1}^{m} n_l[\overline{s_l} - \bar{s}]^2$
6. Define the implied spread-rating correlation $\rho \equiv -\sqrt{\frac{\sigma_E^2}{\sigma_S^2}}$

<u>Step 2</u>

To compute each loan-asset's alphanumeric credit rating given its credit spread s_L, proceed as follows:

1. Compute the mean μ_c and standard deviation σ_c of the spread distribution used in Step 1.
2. Define standard normal deviate $y \equiv \frac{s_L - \mu_c}{\sigma_c}$.
3. Draw standard normal deviate z using a reliable machine-resident algorithm.
4. Solve for standard normal deviate x using $y = \rho z + \sqrt{1 - \rho^2}\, x$.
5. Solve for s_R using $x = \frac{s_R - \bar{s}}{\sigma_s}$, whereby $\bar{s}$ and σ_s are those computed in Step 1.
6. Using the m values of $\overline{s_l}$ already computed, locate s_R when the mean spreads $\overline{s_l}$ are rank ordered from lowest (highest rating) to highest (lowest rating). By definition, spread s_R will lie somewhere between two alphanumeric rating-notches corresponding to two consecutive ratings defined by mean spreads $\overline{s_J}$ and $\overline{s_{J+1}}$ with $\overline{s_{J+1}} \geq \overline{s_J}$. In general, there will be no way to narrow the choice down to either rating[1].
7. Define statistical distances $a \equiv \frac{s_R - \overline{s_J}}{\sigma_j}, a > 0$ and $b \equiv \frac{\overline{s_{J+1}} - s_R}{\sigma_{j+1}}, b > 0$.
8. Choose a suitable limbo margin ε. The value of ε is unknown *a priori* and needs to be computed from the optimization procedure detailed in the next sub-section.
9. When $\frac{a}{b} < 1 - \varepsilon$, select notch j and when $\frac{a}{b} > 1 + \varepsilon$, select notch $j + 1$.
10. In the rare cases whereby $1 - \varepsilon \leq \frac{a}{b} \leq 1 + \varepsilon$ holds, choose notch j when $z < 0$ and notch $j + 1$ when $z > 0$.
11. If $s_R < \overline{s_1}$, use the highest rating in the distribution and if $s_R > \overline{s_m}$, use its lowest rating.

Optimization of Limbo Parameter ε

As already highlighted, limbo parameter ε must be computed *via* an optimization procedure aimed at mis-rating as few loans as possible. In addition, recognizing that, even in the best case, some loans will

[1] Note that, even the market is inefficient and prices an entire rating-notch with a <u>lower</u> average spread than another, lower rating-notch, which would normally attract a <u>higher</u> average spread, the foregoing algorithm still works.

be mis-rated, the rating-error should be as small as possible without sacrificing the low or high side. Rating too high is no better than rating too low. Thus, balance is the key to the entire exercise.

Proceed as follows:

1. Using the database from Step 1, draw a random sample (with replacement) of $\alpha_i, i \in [1, n/10]$ CLO loan-assets using the VBA-resident[2], uniform random number generator Rnd *via* $c_i \equiv 1 + Rnd\ (n-1)$ and $\alpha_i \equiv Rounddown(c_i, 0)$.
2. Compute a new rating for each loan in the sample using the above mapping algorithm. By definition, each selected loan α_i already has a rating and a credit spread.
3. Define the loan-wise score k_i as the difference, in notches, between the given rating and the rating computed in item 2. Assign positive k_i values when the computed rating is higher (better) than the given rating and negative k_i values when the computed rating is lower (worse) than the given rating. Each k_i is equal to the number of notches by which the computed rating is either better (+) or worse (-) than the given rating.
4. For any value of limbo parameter ε, define the associated figure of merit $k_{FOM} \equiv \frac{10}{n} \sum_{i=1}^{n/10} k_i$.

For all values of limbo parameter ε in the basis-point range $\varepsilon \in [25{,}500]$ in intervals of 25 bps, go through items 1 to 4 above and produce the following output in tabular format, one row for each ε-value:

a) The limbo parameter ε.
b) k_{FOM}
c) The number of positive k_i values.
d) The number of negative k_i values.
e) The conditional notch-distribution for positive k_i values, i.e. the probability of $k_i = 1$, $k_i = 2$ and so on[3].
f) The exponent of the best exponential distribution fit to the probabilities in the last item.
g) The conditional notch-distribution for negative k_i values, i.e. the probability of $k_i = -1$, $k_i = -2$ and so on.
h) The exponent of the best exponential distribution fit to the probabilities in the last item.
i) The maximum k_i value.
j) The median of all positive k_i values. This is the value of k_i half-way up the sample.
k) The minimum k_i value.
l) The median of all negative k_i values. This is the value of k_i half-way down the sample.

Obviously, we are looking for $k_{FOM} \sim 0$ or thereabouts, and for the two conditional distributions to be as close as possible to exponential functions with the largest exponents possible. We are also looking for the maximum and minimum k_i-values to be as close to zero as possible. But, all of these desiderata are unlikely to be optimizable at the same ε-value. This is where analytical judgment comes in.

[2] Excel functions are all VBA-callable.

[3] The sum of all these probabilities is 1.0.