Comparable Bond Groupings

Development of a numerical methodology for grouping comparable municipal bonds

Municipal Bond Information Services

The municipal bond information services (MBIS) is a group of eleven firms, including ten that specialize in municipal inter-dealer brokerage, have assembled to create a consortium for the purpose of aggregating the market data generated by businesses of these entities. This consortium has formed a new business entity, Municipal Bond Information Services, LLC ("MBIS"). Triangle Park, an information provider that services the financial community, acts as operating manager in the development and distribution of MBIS services.

For information about MBIS products or services please contact:
Web: **www.mbis.com** Phone: 1-(888)-319-3755 Email: **sales@mbis.com**

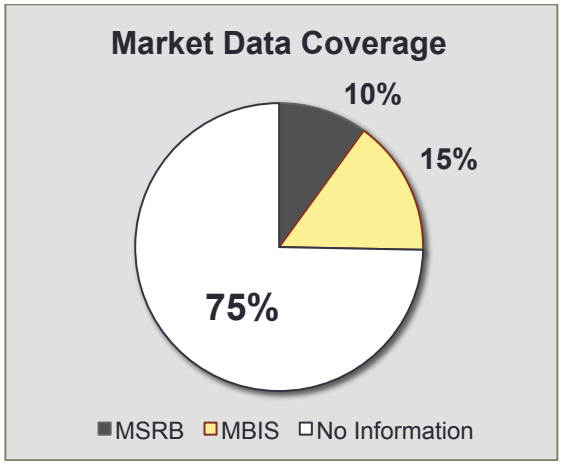
Municipal Bond Information Services
953 Islington St.
Portsmouth, NH 03801

A Lack of Information

With the availability of all of the trade data for the municipal bond market through the MSRB Transaction Reporting Program, transparency of the market is better than for many other fixed-income markets globally. This would make price discovery a relatively easy task if it were not for the issues of activity and scale. On a normal business day, counting every trade from institutional size blocks down to 5 bonds trades, there are approximately 30,000-50,000 trades. Many liquid bonds have multiple trades, so only about 16,000 individual bonds have any trading activity each day. With over 1 million actively traded bonds, this leaves the majority of them with no recent market activity.

The MBIS Data Advantage

In addition to the trade information, MBIS makes available a wealth of pre-trade market information from major municipal brokers and alternative trading systems. This expands the transparency of the municipal market, adding market information on over 20,000 bonds a day.



The above chart shows the overage of market data sampled from a typical 30-day period.

Leveraging the Data

Once all of the available market information has been captured, a way of leveraging the data to cover the illiquid securities is still needed.

Our goal is to organize the universe of municipal securities into groups of similar bonds. If these groups are properly designed we can provide a range of potential contemporaneous market prices for a bond, even if that bond has no recent observed market price.

Our Methodology

To develop these groups we will follow three steps. The first is based on our intuitive understanding of the market, and the second is based on data mining techniques.

Step1: Known Factors

In the first step we will create high level groups based upon basic characteristics: tax status, bank qualification, and refunding status.

<table>
<tr><td rowspan="2">Taxable</td><td colspan="2">Pre-Refunded/ETM</td></tr>
<tr><td colspan="2">Not Refunded</td></tr>
<tr><td rowspan="3">Tax Exempt</td><td colspan="2">Re-Refunded/ETM</td></tr>
<tr><td rowspan="2">Not Refunded</td><td>Bank Qualified</td></tr>
<tr><td>Not Bank Qualified</td></tr>
</table>

In addition, bonds with a very short duration will be grouped separately, as will variable rate demand notes, convertible CABs, floating rate notes, and other specialized market segments. In addition to the categories above, the bonds will be broken down into groups by similar duration.

Further Classification
In a traditional approach, other key factors known to drive market pricing (see below) would be used to further classify the bonds into smaller subsets in order to obtain groups of bonds that will trade at similar yield levels.

Key Market Pricing Factors:
Rating, Revenue vs. General Obligation Pledge, Purpose of the Proceeds, State, Insurance, Issue Size, AMT, Sinking Funds, Coupon, Call Protection, etc.

A Challenge: Failure of Intuition
The potential problem with this intuition-based grouping is that bonds might trade in ways that are not consistent with these groupings. For instance, while it would make sense to separate general obligation credits from revenue credits, many revenue credits for utilities and other essential services might trade in ways more similar to unlimited general obligation credits. By contrast, many limited general obligation credits might trade more similar to revenue credits for more speculative projects. We can envision many situations where this intuitive grouping does not necessarily capture actual patterns in yields and bid-ask spreads.

Step 2: Cluster Analysis
For that reason our analysis will include a second step. In this second step we will create a different set of 100 groups based on a data mining algorithm known as "agglomerative cluster analysis." Agglomerative cluster analysis answers a simple question: How do we create groups of bonds where yields and bid-ask spreads of bonds within the groups are as similar as possible, while the differences in yields and bid-ask spreads of bonds across the groups are as different as possible? In other words, how do we minimize the variation within groups while maximizing the variation across groups? We will cluster the universe of municipal bonds on yield and bid-ask spreads, and other relevant characteristics. Bid-ask spreads are one of the unique features of the MBIS data and a crucial part of our analysis. From bid-ask spreads we can learn about market depth, liquidity, and volatility and other features of bonds that are difficult to discern from observed trades alone. The cluster analysis will be particularly useful in identifying instances where bonds don't trade as expected.

Cluster Analysis
Cluster analysis is a statistical classification technique for discovering whether the individuals of a population fall into different groups by making quantitative comparisons of multiple characteristics

From those groupings we can identify a range of likely yields for a given bond at any point in time based on the observed yields and bid-ask spreads of other bonds in that particular credit's "peer group."

Shortcomings of Cluster Analysis
There are two problems with the clustering approach described above when applied to this application:

1. The clustering algorithm as described requires market data on each bond in order to perform the clustering.
2. It is difficult to arrive at an intuitive description of each cluster beyond its immediate price dynamics.

The first problem is the most significant because it means that we cannot rely on clustering alone to group bonds with no recent market data. Using the last known market data for each bond is problematic because of changes in the interest rate environment, changes in the spread relationships in the market, and because of the natural "roll-down" on the yield curve that will happen for each bond over time. Another approach is required. The second problem must be overcome because our goal is to arrive at a set of groupings that is both fairly captures actual price dynamics and is intuitive.

Step 3: Statistical Classification

In order to group bonds with no recent market data, a grouping model must be developed that can operate entirely on reference data. If a way can be developed to identify the reference data factors driving the clusters, these factors could be used to further classify all of the bonds, including the majority of the bonds that are not traded often. The approach we will use to accomplish this is knows as logistic regression.

Logistic Regression

Logistic regression ("logit regression") is a probabilistic statistical classification model that is used to predict a categorical dependent variable (class) based upon one or more predictor variables (factors). In the case where there are more than two classes, it is referred to as a multinomial logistic regression model.

Simply stated, the logistic regression will determine how to use the reference data characteristics to map all bonds, both with market data and without, into the clusters identified in step 2. It will also measure the probability that each bond has been classified correctly.

From a technical perspective, a multinomial logistic regression model will be used, with the categorical dependent variable being the clusters identified in step 2 and the predictor variables being the reference data characteristics. The regression will be performed on the bonds with recent market activity for which the clusters are already known. It will then be used to reclassify all of the bonds, regardless of market activity.

The Final Product

The analysis described above will result in grouping that will be captured by assigning a grouping ID to each group and to each municipal bond. The bond-group ID mapping will be provided to customers and each piece of market data, both MSRB trades and MBIS quotes, will be labeled with a group ID so that it will be easy to relate the market data to the relevant group.

This data can be used to easily select market data that is likely to be relevant for any bonds with little or no recent market activity.

Many Applications

The groupings can be used to aid in price discovery to support many different functions: mark-to-market of assets, pricing surveillance, pre and post trade compliance, and many others.

More Information

For more information contact sales@mbis.com.