

Evaluating Machine Learning Predictive Pricing Models

Gregg L. Bienstock *SVP*, *Group Head*, *Municipal Markets SOLVE* **Michael Pellerito** *SVP*, *Credit Strategy & Derived Data SOLVE*



Article featured in Bond Dealers of Americas FIXED INCOME INSIGHTS, Spring 2025

Muni Traders and Portfolio Managers (PMs) are facing increased competition from market participants who are using AI predictive pricing models to generate alpha. Understanding how these models work, where they perform best, and how to integrate them into decision-making is becoming essential for front-office professionals. The use of AI predictive models is increasingly allowing traders and PMs to make faster decisions, with less time spent on price discovery, and more automated insight generation. As a result, traders and PMs can dedicate more time to optimizing strategies, managing portfolios, and identifying relative value opportunities.

Some helpful advice as you explore this world is to start with education on AI predictive pricing models:

- Another lens into the liquidity of PRIV's constituents is to examine how widely they are held by mutual Review whitepapers and research reports
- Understand back-testing analysis and error distributions, being mindful of model carve-outs for better comparison
- Determine where these models excel for the subset of bonds that interest you from the universe and ask for back-testing on bonds you care about.

Market Applications

Al predictive pricing is used in a variety of ways, including:

- 1. Pre-trade pricing insights to enhance visibility into pricing before trades.
- 2. Independent pricing source that is unique from other sources and based on proprietary market data and AI models.
- 3. Relative value analysis facilitating value comparisons across bonds.
- 4. Client model input serving as an independent, external comparison pricing metric for internal client models and their output.
- 5. Size and side-specific pricing that offers clients precise pricing data based on their customized situations.

Introducing SOLVE Px

SOLVE has produced an artificial intelligence-driven predictive pricing platform that accurately predicts the next trade price, yield, and spread for municipal bonds. SOLVE Px[™] harnesses SOLVE's unique observable Quotes[™] data, along with reference and trade data, to power Al models that incorporate several hundred feature inputs to generate size- and side-specific predictive prices. SOLVE Px[™] is unique in its ability to leverage SOLVE Quotes[™] data to create an accurate predictive price of a hypothetical trade (with the size and side determined by the user) for any active municipal fixed-rate bond, regardless of its liquidity. This unique feature makes SOLVE Px[™] an invaluable tool for front-office municipal fixed-income professionals, enabling them to identify arbitrage opportunities and address the common lack of price transparency, even for illiquid bonds.

A recent addition to functionality is the introduction of SOLVE Px Confidence Score, an AI-powered metric that quantifies pricing certainty on a scale of 1-10, with higher scores indicating greater confidence in price predictions. This innovative feature is designed to provide transparency and enhance the usability of SOLVE Px[™] predictions.

Since the value of predictive pricing is intrinsically tied to its accuracy, SOLVE continuously examines the impact of various factors on the accuracy of the SOLVE Px output. This includes the AI and Machine Learning ("AIML") models themselves, as well as the data that is used by those models. Since one of the key differentiators is the addition of our proprietary Quotes data, we aim to measure that data's impact on the accuracy of the predictions.



Model and Methodology

SOLVE Px is designed to predict the next bid, mid, and offer trade for over 900,000 fixed coupon municipal bonds of any size. SOLVE Px uses Machine Learning ("ML") models that find patterns from a supervised set of input data. That dataset consists of:



Real-time SOLVE Quotes[™] Data: Comprising bids and offers from market participants¹

- **Trade Data:** All trades reported in real-time
- **Reference Data:** Bond-specific reference data
- Benchmark Curves: Real-time treasury and AAA benchmark curves

Back Testing and Benchmarking

Below, we present an overview of the model's efficacy by looking at benchmarking and back-testing. Access the full content and analyses from which this article is summarized by downloading these two whitepapers:

- The Power of Quotes: Increased Accuracy in Predictive Pricing
- "SOLVE Px™: SOLVE's Proprietary AI-Driven Municipal Bond Predictive Pricing

SOLVE is the largest supplier of indicative market quote data to the financial industry, parsing nearly 20 million quotes each day on 130k unique securities across all asset classes. SOLVE parses nearly 430k muni quotes on 55k unique securities each day. For comparison, that's about 8x more observations on 3x more unique securities relative to trades on a typical day.

Median of Absolute Errors

The SOLVE Px Prediction Model's primary goal is to minimize the absolute difference between the predicted price and the next trade level; this absolute difference is referred to as "error". Over the 146-business-day (7-month, 8.4m trades) back-testing period ending at YE 2024:

- 1. The median of all absolute yield errors was 5.7bps
- 2. The median of all absolute price errors was 21.8 cents (this includes all trades)
- **3. The bias was under 1 cent.** The bias measures the average raw error for each trade-Px pair during the back-testing period.

Liquid vs Illiquid Performance and the Power of Quotes

When looking at prediction/trade pairs that had a preceding trade and quote on the same day, the error was lower than the universe at 4.3 bps compared to all bonds at 5.7 bps. **Interestingly, when Quotes are included in SOLVE Px, regardless of the existence of recent trade observations, error rates significantly decrease, demonstrating that the inclusion of Quotes in the AI Model materially improves prediction accuracy.**

When Quotes are included in SOLVE Px, regardless of the existence of recent trade observations, error rates significantly decrease.² :

- Where trades do not exist, prediction accuracy improves by 34%.
- Where trades do exist, **prediction accuracy improves by 25%**, thus demonstrating a symbiotic and synergistic benefit leading to the most accurate predictions.

² Quotes are available leading up to the majority of trades
67% of back-testing trades had Quotes within 20 days before the trade.
58% of back-testing trades had Quotes within 5 days before the trade.
53% of back-testing trades had Quotes within 1 business day of the trade.

SHAP Analysis Summary

- Although quote-related features constitute 27% of all feature inputs, they account for approximately 40% of the prediction accuracy for SOLVE Px.
- Trade-related features represent 29% of all feature inputs, yet they represent approximately 45% of the prediction accuracy for SOLVE Px.
- Combined quotes and trades contribute approximately 85% of the predicted accuracy for SOLVE Px.

Trade Size Analysis

Our analysis reveals that error rates decrease consistently as trade size increases, with the lowest MAE found in trades greater than \$1m face value. The smallest retail trades (< \$2,000) exhibited the highest error rates.



Coupon Analysis

Our analysis of MAE across bond coupons revealed that the lowest error rates were found in bonds with coupons ranging from 4% to 5.99%, which accounted for 81.0% of trades during the back-testing period. Bonds with coupons between 2.99 and 3.99% (3% category) represented 11.8% of all trades. The following chart illustrates error rates by Coupon:



Maturity Analysis

Our analysis of maturity revealed that error rates decrease as bond maturities increase. During the back-testing period, 58% of trades were for bonds with maturities within 10 years. The following chart highlights error rates by Maturity:



Trade Time of Day Analysis

Our analysis of hourly trades reveals that 99.6% of trades occur between 8 am and 4:59 pm. Generally, we observe higher error rates in the afternoon, while the morning hours, between 8 am and 11:59 am EST, exhibit lower error rates. This morning period is particularly significant as over 75% of daily Quotes data are received during this time, with 26% of daily Quote volume received during the 9 am hour, significantly influencing Solve Px. The following chart displays error rates by the time of day for each back-testing trade:



Closing Remarks

Al-driven predictive pricing models are transforming municipal trading by enhancing decision-making, improving efficiency, and unlocking new opportunities for traders and portfolio managers. As these models continue to evolve, understanding their capabilities, limitations, and optimal use cases will be essential for maximizing their value.

SOLVE's commitment to innovation, exemplified by the SOLVE Px[™] platform, underscores the power of proprietary data and AI in refining trade price predictions. By leveraging these advanced tools, front-office professionals can gain a competitive edge, optimize strategies, and navigate the complexities of the municipal bond market with greater confidence.

Through rigorous back-testing and benchmarking analysis, SOLVE continuously evaluates model performance, ensuring high accuracy and reliability across various bond subsets and market conditions. Our extensive datasets provide powerful insights into model behavior, error distribution, and market dynamics, enabling traders and PMs to make more informed decisions. This comprehensive approach allows us to refine predictions, minimize errors, and deliver a deeper understanding of price movements, giving market participants a critical advantage.

About

SOLVE is the leading market data platform provider for fixed-income securities, trusted by sophisticated buy-side and sell-side firms worldwide. Founded in 2011, SOLVE leverages its Al-driven technology and deep industry expertise to offer unparalleled transparency into markets, reduce risk, and save hundreds of hours across front-office workflows. With the largest real-time datasets for Securitized Products, Municipal Bonds, Corporate Bonds, Syndicated Bank Loans, Convertible Bonds, CDS, and Private Credit, SOLVE empowers clients to transform the way they bring new securities to market, trade on secondary markets, and value highly illiquid securities. Headquartered in New York, with offices across the globe, SOLVE is the definitive source for market pricing in fixed-income markets. For more information, visit https://solvefixedincome.com

Contact SOLVE

600 Summer Street, Suite 503 Stamford, CT 06901 USA +1 (646) 699-5041 info@solvefixedincome.com www.solvefixedincome.com