Are bond price predictible?

Saurav Mawandia[1] & Sougandh Kohli[1]

[1] Harrisburg University of Science and Technology

Author Note

The authors made the following contributions. Saurav Mawandia: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Sougandh Kohli: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Saurav Mawandia, 326 Market St, Harrisburg, PA 17101. E-mail: smawandia@my.harrisburgu.edu

Abstract

The capital market plays a crucial role in the development of a country. The overall size of the global bond market as of 2021 is over 128.3tn. In contrast, the overall size of the equities market is 44 trillion as of 2022. However, the bond market is not much transparent compared to stock, and only a few firms try to predict the price of bonds. Different brokers can give different prices for the same bond, making it challenging for anyone to predict the bond's price based on these data. A bond is usually based on the last traded price and indicative bids and asks. Moreover, predicting the price of a bond thus becomes challenging. That is why it is essential to understand the overall market for bonds and consider the prices. For humans, it is tough to keep an eye on the fixed income market and make a decision. That is where Artificial Intelligent systems help us. Machine learning algorithms use the data and find hidden patterns to predict the correct price of a given bond. Given the data related to the fixed income market, news related to the market, and so on, we train an AI algorithm to predict the price of the given bond. It considers the details like rating, last traded price, type, and news around the bond to predict the right price for a trader to trade.

Keywords: keywords

Word count: X

Are bond price predictible?

Introduction

A bond valuation uses a yield to maturity or a zero yield curve using a few points on the par bond or Nelson Siegel parameters. The bond is valued using the settlement date and maturity date using an algorithm that can calculate the time to maturity and each coupon date. Maturity can be in years; if unavailable, a zero maturity date is used. This mode is particularly convenient for price par bonds or price other bonds on issue dates or coupon dates. Using this mode between coupon dates is difficult as the user has to compute the day count and year fraction and provide the maturity as, say, 5.3 years. Traditionally bond price was calculated using YTM(Yield to maturity) or zero yield curve. However, this approach does not predict an accurate price of a bond because it does not consider other parameters such as inflation. We have used the LSTM model to predict trade price based on the last trade price, which can be improved further to consider external factors such as inflation. Long short-term memory (LSTM) is an artificial neural network in artificial intelligence. Unlike standard feedforward neural networks, LSTM has feedback connections. Such a recurrent neural network (RNN) can process not only single data points but also entire data sequences and is also suitable for analyzing time series data.

Background

As far as price transparency is concerned, there has historically been a huge gap between the amount of reference information available to those trading equities versus those trading corporate bonds. Stock exchanges report trades, bids and offers at all times. Free access is available online with a 15 minute delay while traders who demand more information can pay for ultra efficient real time data and information about size of current bids and offers. By contrast, bond trades are required to be reported within 15 minutes and only those who pay for the TRACE feed can access this information. No quotes are

publicly available and the best way to get a quote is to solicit multiple brokers and wait for a reply. Alternatively there are data companies that provide end of day prices, published after the market has closed and with no guarantee that the specific information sought will be included. Accurate bond pricing is also hindered by lack of liquidity. Only a fraction of TRACE eligible bonds trade on a given day, so the most recent trade price is often multiple days old. Pricing bonds based on other more liquid bonds that have similar features is common, but again limited by the presence of such bonds. Thus, in the context of bond price predictions, machine learning should be fast and accurate. They have evaluated the performance of various supervised learning algorithms for regression followed by ensemble methods, with feature and model selection considerations being treated in detail(Ganguli and Dunnmon (2017)). (Götze, Gürtler, and Witowski (2020)) compares the forecasting performance of linear regression models and enhanced machine learning methods in the catastrophe (CAT) bonds market using linear regression with variable selection, penalization methods, random forests, and neural networks to forecast CAT bond premia. Random forests exhibit the highest forecasting performance among the considered models, followed by linear regression models and neural networks.

Benchmark Solutions is the first provider of realtime corporate bond prices. Every 10 seconds they provide accurate prices that incorporate interest rate data, trades or quotes of the bond in question, trades or quotes of other bonds or CDS of the issuer of the bond in question as well as other input sources. Pricing bonds accurately requires an exacting knowledge of payment schedules, trading calenders and reference data for each bond. This, as well as synthesizing all of the bonds and CDS quotes and trades of a given issuer into implied hazard and funding curves, is something that they feel is beyond the scope of this challenge. (Golbayani, Florescu, and Chatterjee (2020)) Rather, they provide you with a reference price which is an intermediate result of our calculations and is labeled 'curve_based_price' in the dataset. Thus the competition focuses on trading dynamics and microstructure of individual bonds, rather than all bonds from a given issuer. (Ganguli and

Dunnmon (2017)) proposes a novel hybrid time-series aided machine learning method that could be applied to such datasets in future wor.

Benchmark solutions are the first provider of real-time corporate bond prices discussed in (Bacas (2018)). Every 10 seconds, they provide accurate prices that incorporate interest rate data, trades or quotes of the bond in question, trades or quotes of other bonds or CDS of the issuer of the bond in question, and other input sources. Pricing bonds accurately requires exacting knowledge of payment schedules, trading calendars, and reference data for each bond. This, as well as synthesizing all of the bonds and CDS quotes and trades of a given issuer into the implied hazard and funding curves, is something that we feel is beyond the scope of this challenge. Rather, we provide you with a reference price which is an intermediate result of our calculations and is labeled 'curve_based_price' in the dataset. Thus the competition focuses on trading dynamics and microstructure of individual bonds rather than all bonds from a given issuer.

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

Material

Procedure

Data analysis

We used R (Version 4.1.0; R Core Team, 2021) and the R-packages *papaja* (Version 0.1.0.9999; Aust & Barth, 2020), and *tinylabels* (Version 0.2.3; Barth, 2022) for all our analyses.

Results

Discussion

References

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. Retrieved from https://github.com/crsh/papaja

Bacas, T. (2018). *Predicting Accurate Bond Prices with Neural Networks*. https://medium.com/@thomasbacas/predicting-accurate-bond-prices-with-neural-networks-ff46c051c25c/.

Barth, M. (2022). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Ganguli, S., & Dunnmon, J. (2017). Machine learning for better models for predicting bond prices. *arXiv Preprint arXiv:1705.01142*.

Golbayani, P., Florescu, I., & Chatterjee, R. (2020). A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. *The North American Journal of Economics and Finance*, *54*, 101251.

Götze, T., Gürtler, M., & Witowski, E. (2020). Improving CAT bond pricing models via machine learning. *Journal of Asset Management*, *21*(5), 428–446.

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/