BEHIND THE MARKETS PODCAST: USING CHATGPT TO FORECAST EQUITY PRICE MOVEMENTS, WITH ALEJANDRO LOPEZ LIRA

Christopher Gannatti – Global Head of Research, WisdomTree.
10 Aug 2023

In late July 2023, we had the pleasure of speaking with Alejandro Lopez Lira, Assistant Professor at the University of Florida's Warrington College of Business. It was great to join him back at the University of Pennsylvania's campus where he earned his PhD.

The focus of the discussion was on asset pricing, and this is where the bulk of Alejandro's research and work has been. Alejandro took us back to the line of thinking he had during his research at The University of Pennsylvania, where he was motivated to see what drives different assets to have different expected returns. Machine learning was an interesting tool that could be used for text analysis, and it was clear that individual companies were disclosing lots of risks within their 10-K documentation. The concept was to capture these risks from within the unstructured, text data and then to figure out whether the market was treating them in a way that allowed investors to be compensated for bearing them.

Looking at academic factors

At WisdomTree, we spend a lot of time looking at different factors to both help in explaining the market returns that we are seeing, as well as helping to build new strategies for investors. There is a massive amount of attention paid to the group of so-called 'academic factors'. To give people a baseline for what an academic factor is, 'book-to-market' is a great example. It is referring to the ratio of a company's book value of equity, relative to its market value of equity, with a higher figure indicating that the company is more of a 'value-stock.'

Alejandro, along with his co-authors Andrew Y. Chen and Tom Zimmerman, wrote a paper titled 'Peer-Reviewed Theory Does Not Help Predict the Cross-section of Stock Returns'. While our conversation didn't cover the paper in full, we did look at how for every published factor there are numerous other factors that also have a similar explanatory power over the observed variation in returns. The gist is that published, academic factors are not the be all and end all, and there seems to be a high degree of correlation between both published and unpublished factors. It is looking like the explanatory power in general is declining, indicating that it is possible that more and more information is disseminating into asset prices more and more quickly.

The cost of research and development

We considered the ongoing debate as to whether companies should capitalise their research and development (R&D) expenditures. The logic is that when a company buys a physical asset, they hold that asset on their balance sheet and, over time, depreciation

expenses take the carrying value lower and lower until the useful life of that asset has ended. Think of the example of Bard, the large language model (LLM) developed by Alphabet's Google. The company clearly had to invest significantly to build this model, and we should assume it brings the company some value. Yet, the Bard LLM does not live anywhere directly on Alphabet's balance sheet, so researchers get to debate whether it ends up captured in the firm's book value of equity or somewhere else.

Jeremy Schwartz, WisdomTree's Global CIO, asked Alejandro if he had a 'favorite factor'. Alejandro noted that he prefers to stay away from aggregate market prediction and to always mix up different factors and not restrict himself to a single signal. He appreciates the capability that machine learning brings to his research because it allows for patterns and signals to be seen within large datasets.

Machines aren't infallible

Now, an important part of the discussion regarded some of the limitations of machine learning. Simply put, it works better with more data. We must also recognise that, if we are using ChatGPT to apply machine learning and artificial intelligence (AI), then this model is designed to predict text well—importantly it is not designed to perform any specific finance-oriented functions. Interestingly, Alejandro noted that if one tried to have ChatGPT multiply two large numbers together, there is a high likelihood that it would come to an incorrect answer and, without proper plug-ins, techniques like linear regression are beyond its capabilities. But, if the remit is instead to ingest massive amounts of company news and headlines, arriving very quickly, it could be the ideal tool.

Anyone seeking to repeat these results or to backtest strategies using LLMs should keep the following in mind:

When models are trained, essentially on the entire public internet, it's possible that it will not be possible to truly see the predictive power of a given approach to equity price forecasting because it could be the case that the systems are simply regurgitating the actual price movements which could very well have been included in the training data.

Since ChatGPT was cut off at September 2021, Alejandro has only been able to truly test the system's predictive power since that point. This is why the work focuses on daily returns—there is the highest frequency and thereby more data of daily returns. It might be possible to look at some weekly studies but, with only about 18 months of data, monthly return forecasting is not yet possible. Therefore, any strategy seeking to employ these models would be based on higher frequency, higher turnover approaches as opposed to anything more long term or more 'buy-and-hold'.

It was notable to take the new toolkit, large language models, and apply it to basically a better, more comprehensive look at sentiment analysis. Alejandro hypothesised that a critical reason why it could be working well is that, on a daily horizon, it is difficult for market participants to trade small stocks, particularly on the short side, since negative headlines would then require a short position. Transaction costs could also be higher, particular for those investors with more assets under management, making it impossible for them to exploit these opportunities without moving the market.

Those interested in the full discussion can access it here.

Related blogs

+ Behind the Markets Podcast: a conversation about deep decarbonisation

View the online version of this article here.

Important Information

Marketing communications issued in the European Economic Area ("EEA"): This document has been issued and approved by WisdomTree Ireland Limited, which is authorised and regulated by the Central Bank of Ireland.

Marketing communications issued in jurisdictions outside of the EEA: This document has been issued and approved by WisdomTree UK Limited, which is authorised and regulated by the United Kingdom Financial Conduct Authority.

WisdomTree Ireland Limited and WisdomTree UK Limited are each referred to as "WisdomTree" (as applicable). Our Conflicts of Interest Policy and Inventory are available on request.

For professional clients only. The information contained in this document is for your general information only and is neither an offer for sale nor a solicitation of an offer to buy securities or shares. This document should not be used as the basis for any investment decision. Investments may go up or down in value and you may lose some or all of the amount invested. Past performance is not necessarily a guide to future performance. Any decision to invest should be based on the information contained in the appropriate prospectus and after seeking independent investment, tax and legal advice.

The application of regulations and tax laws can often lead to a number of different interpretations. Any views or opinions expressed in this communication represent the views of WisdomTree and should not be construed as regulatory, tax or legal advice. WisdomTree makes no warranty or representation as to the accuracy of any of the views or opinions expressed in this communication. Any decision to invest should be based on the information contained in the appropriate prospectus and after seeking independent investment, tax and legal advice.

This document is not, and under no circumstances is to be construed as, an advertisement or any other step in furtherance of a public offering of shares or securities in the United States or any province or territory thereof. Neither this document nor any copy hereof should be taken, transmitted or distributed (directly or indirectly) into the United States.

Although WisdomTree endeavours to ensure the accuracy of the content in this document, WisdomTree does not warrant or guarantee its accuracy or correctness. Where WisdomTree has expressed its own opinions related to product or market activity, these views may change. Neither WisdomTree, nor any affiliate, nor any of their respective officers, directors, partners, or employees accepts any liability whatsoever for any direct or consequential loss arising from any use of this document or its contents.