Order Flow Imbalance (OFI) Analysis - Conceptual Questions

Joseph Diggins

June 15, 2025

Motivation Behind Measuring OFI at Multiple Depth Levels of the Order Book

The motivation behind measuring OFI at multiple depth levels of the order book according to the paper is that multi-level OFI includes more information. Using best-level OFI ignores useful information from the deeper levels, and this can help us understand price movements and explain them more accurately. Best-level OFI also shows the smallest correlation with any of the other 9 levels across many different stocks. According to Table 2, Multi-level OFIs explain over 89% of the variance, which shows there is a lot more understood through these deeper levels. This information can also be compressed using PCA and is another motivation behind measuring OFI at multiple levels.

Why Do the Authors Use Lasso Regression Rather than OLS for Estimating Cross-Impact?

Authors use Lasso regression rather than OLS for estimating cross-impact because it can select the most important cross-impact terms. "LASSO is a regression method that performs both variable selection and regularization, in order to enhance the prediction accuracy and interpretability of regression models." It is also used instead of OLS as OLS struggles with cross-asset OFI data. The number of predictors and strong correlation between them make it unsuitable for OLS. "OLS regression becomes ill-posed when there are fewer observations than parameters." Moreover, the multicollinearity issue of features contradicts the necessary condition for a well-posed OLS.

Why is OFI Considered a Better Predictor of Short-Term Returns than Trade Volume?

OFI is considered a better predictor of short-term returns than trade volume because it distinguishes between buying and selling, whereas trade volume does not. According to the paper, price changes over short time intervals are mainly due to OFI, and deep learning models trained on OFIs outperform most models trained on order books or returns. This shows OFIs have more predictive power than just looking at volume or raw data. Models using OFI outperform others in predicting short-term returns. "We also implement the same hypothesis test described in Section 3... We observe that the cross-impact models exhibit significantly superior performance than the price impact models across all stocks, at the 1% confidence level."