

Submission

June 15, 2025

1 GitHub Link

<https://github.com/pradeep-muthaiya/orderFlowImbalance/tree/main>

2 Conceptual Question Answers

1. What's the motivation behind measuring OFI at multiple depth levels of the order book?

The primary motivation for measuring OFI at multiple depth levels of the order book is to gain a better more comprehensive understanding of the overall market pressure and overflow imbalance. While best-level OFI (level '00') captures immediate market pressure, multi-level OFI provides several key advantages. At different price levels (as shown by the different depth levels), and not just the best one, you are then able to see how order flow pressure propagates through different price levels, revealing the true depth of market interest.

As different levels of the order book have varying degrees of price impact, multi-level OFI helps predict price movements more accurately by considering the total pressure across all levels. This is because in many cases, larger orders are often split across multiple price levels to minimize market impact, and multi level OFI can detect these hidden liquidity pools that would not be shown at just the best (00) price.

Lastly, as with many other data analysis, having more information helps filter out what is true movement, versus what is just noise. By having more information, multi-level OFI can reduce the noise inherent in best-level data, providing a more stable and reliable indicator of market pressure. This approach allows a better understanding of the true market pressure and make more informed trading decisions based on the complete order book structure rather than just the top level.

2. Why do the authors use Lasso regression rather than OLS for estimating cross-impact?

The authors use Lasso regression instead of OLS for estimating cross-impact because it provides a more wholistic and practical approach to handling the complex relationships between different assets. While OLS would give us a basic understanding of cross-impact, Lasso offers several key advantages that make it more suitable for this analysis.

Fundamentally, Lasso helps us identify which cross-asset relationships actually matter, rather than just showing us all possible relationships. This is crucial because in financial markets, many assets might appear to have some relationship, but not all of those relationships are meaningful or useful for prediction. Lasso automatically filters out these weaker relationships by setting their coefficients to zero, giving us a clearer picture of the true cross-impact effects.

Another important aspect is how Lasso handles the high dimensional nature of cross-impact analysis. When we're looking at relationships between multiple assets, where we consider the various levels of the limit book for both assets, and potentially analyze their cross product, the number of potential interactions can grow very large. Lasso is specifically designed to handle this kind of data, whereas OLS might struggle or give us unreliable results when dealing with so many variables.

Perhaps most importantly, Lasso helps us deal with the noise in financial data. Similar to multi-level OFI, Lasso helps filter out noise in cross-impact analysis by focusing on the most significant relationships. This gives us a more stable and reliable model for understanding how order flow imbalances in one asset affect others.

This approach allows us to make more informed decisions about cross-asset relationships, focusing on the most important connections rather than getting lost in the noise of less important relationships that OLS would include in its analysis.

3. 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 trade volume doesn't consider the direction in which an asset is moving. Furthermore, it doesn't account for the weight of movement that can be quite beneficial to the movement as compared with higher volume and lower pressure which can have overall less impact for a trade's material value.

OFI is considered a better predictor of short term returns than trade volume because it captures the directional pressure of the market, which trade volume alone simply can't do. While trade volume tells us how much activity is happening, it doesn't tell us which way the market is actually moving or how strong that movement is.

Even if you had a high volume of trades, but they're evenly split between buyers and sellers, the actual price impact might be minimal. OFI, on the other hand, shows us the imbalance between buying and selling pressure, giving us a much clearer picture of where the price is likely to go.

What makes OFI particularly powerful is how it accounts for the weight of market pressure. A large order at the best price level can have a much bigger impact on price movement than multiple smaller trades that add up to the same volume amount. This is because OFI considers not just the size of orders, but also their placement in the order book and their potential to move the market.

Furthermore, OFI helps us understand the true market pressure by looking at the imbalance between bid and ask sizes. A high volume of trades might look impactful, but if the underlying book shows a strong imbalance in favor of either buyers or sellers, that's where the real price movement potential lies. This is especially important for the short term, where the immediate pressure on the order book can be a better indicator of price movement than the historical volume of trades.

In essence, while trade volume gives us a broad picture of market activity, order flow imbalance gives us the specific directional pressure that's actually driving price movements, making it a more precise tool for predicting shorter term traders and returns.