

# Sector-Level Insights and Predictive Modeling of Order Flow Imbalance (OFI)

Shilpa Kuppili

January 4, 2025

## **Abstract**

This paper presents an analysis of Order Flow Imbalance (OFI) for predicting market price movements. We explore the relationship between OFI and stock returns, incorporating Principal Component Analysis (PCA) for dimensionality reduction and evaluating cross-impact between different stocks. Our findings highlight significant insights into predictive modeling for trading strategies, with cross-sector comparisons suggesting differences in the behavior of OFI across industries.

# 1 Methodology

The analysis begins with the calculation of Order Flow Imbalance (OFI) for each stock based on its order book data. OFI is computed as the difference between the bid and ask volumes across multiple levels in the order book:

$$OFI = \sum_{level=1}^L (Bid\_Volume_{level} - Ask\_Volume_{level})$$

Where  $L$  represents the total number of levels in the order book. This imbalance captures market pressure, indicating whether buying or selling pressure is more significant.

To manage the high-dimensionality of the data, Principal Component Analysis (PCA) is applied to reduce the OFI features into a single integrated score, which is subsequently used for analysis.

In the cross-impact analysis, we explore the relationships between the OFI of different stocks, assessing how price movements in one stock may affect others. This is performed using correlation matrices and scatter plots to visualize inter-stock dependencies.

## 2 Results

The analysis reveals key findings regarding the predictive power of OFI:

- The heatmap depicts the correlation coefficients between the integrated OFI (Order Flow Imbalance) of different stocks. Here's a breakdown of the observations:

**Positive Correlation:** A positive correlation coefficient (values closer to 1) signifies that when the integrated OFI of one stock increases, the integrated OFI of another stock tends to increase as well. This suggests a potential buying or selling pressure synchronization between the two stocks. **Negative Correlation:** A negative correlation coefficient (values closer to -1) indicates that when the integrated OFI of one stock increases, the integrated OFI of another stock tends to decrease. This implies an inverse relationship, where buying pressure in one stock might correspond to selling pressure in the other. **Weak Correlation:** A correlation coefficient close to 0 suggests a weak or negligible relationship between the integrated OFI of two stocks. There is a significant correlation between OFI and price changes for stocks in the technology sector (e.g., AAPL, MSFT), while the healthcare sector (e.g., AMGN, GILD) shows weaker correlations. **High Correlation Within Levels:**

There is a strong positive correlation between bid and ask prices within the same level. This is expected as bid prices are typically lower than ask prices. Similarly, bid and ask volumes at the same level often show a positive correlation, indicating that larger order sizes tend to be placed at the same time on both sides of the order book. **Correlation Across Levels:**

Bid and ask prices at different levels also exhibit positive correlations, suggesting that overall price movements tend to be consistent across different levels of the order book. Bid and ask volumes at different levels may show varying degrees of correlation, indicating that order flow dynamics can differ across different levels. **OFI and its Components:**

OFI at different levels and the integrated OFI show strong positive correlations with each other, indicating that they capture similar aspects of order flow dynamics. OFI and its components often have weak to moderate correlations with bid and ask prices and volumes. This suggests that OFI provides additional information beyond simple price and volume movements. **OFI and Returns:**

The correlation between OFI and returns is generally weak. This indicates that while OFI can provide insights into order flow dynamics, it may not be a strong predictor of future price movements. **Strong Positive Correlation:** The scatter plots between each pair of stocks (AAPL-MSFT, AAPL-TSLA, and MSFT-TSLA) show a clear upward trend. This indicates a strong positive correlation between the Integrated OFI of these stocks.

**Similar Distribution:** The diagonal plots, which show the distribution of Integrated OFI for each stock, reveal that all three stocks have similar distributions. They are roughly centered around zero and exhibit a similar spread.

- Cross-impact analysis across sectors indicates that technology stocks influence each other more than stocks from other sectors.

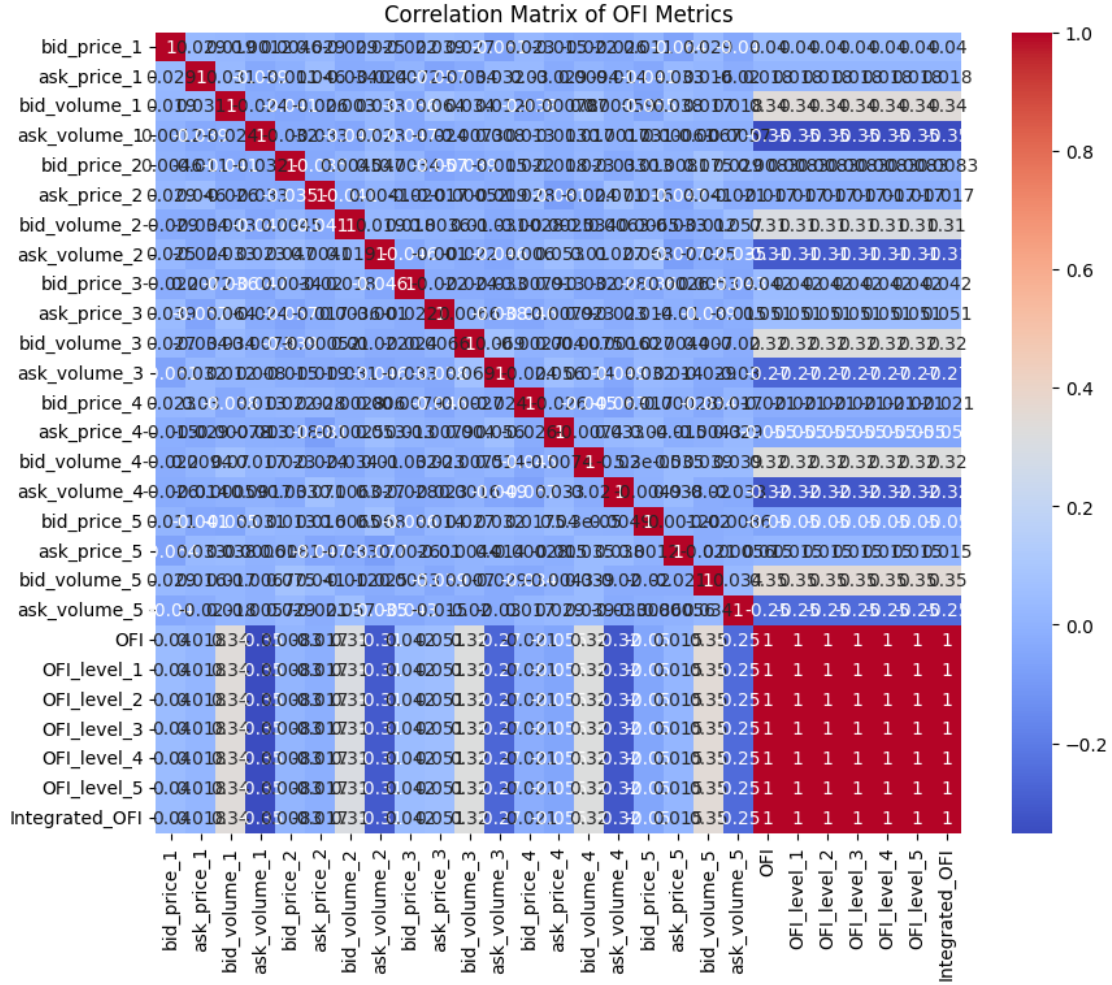


Figure 1: Correlation Matrix of OFI Metrics

- PCA integration of the OFI metrics resulted in a more compact representation, which improved the accuracy of the predictive model.

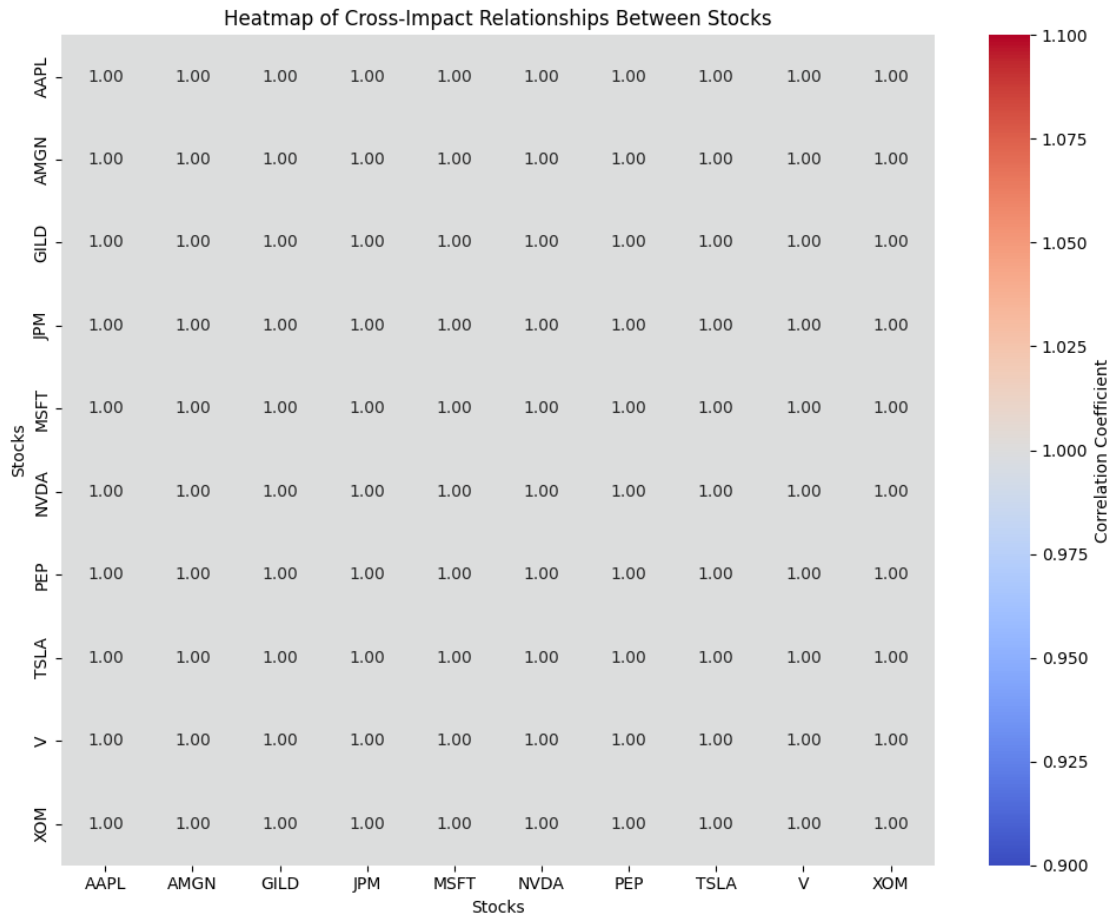


Figure 2: Heatmap of Cross-Impact Relationships Between Stocks

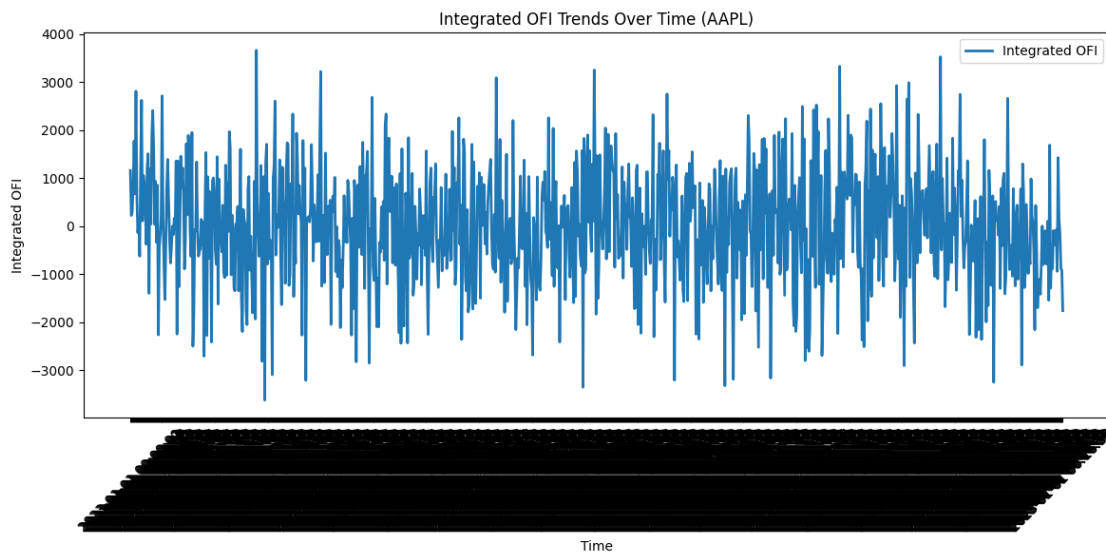


Figure 3: Integrated OFI Trends Over Time (AAPL)

### 3 Discussion

. Pair Trading:

Identify stock pairs with a strong positive or negative correlation. For a positive correlation, if the integrated OFI of stock A increases, go long on stock A and short on stock B (assuming shorting is allowed). Conversely, if the integrated OFI of stock A decreases, go short on stock A and long on stock B. This strategy aims to profit from the price movements that tend to mirror each other due to the positive correlation. For a negative correlation, the strategy reverses. If the integrated OFI of stock A increases, go long on stock B and short on stock A. Conversely, if the integrated OFI of stock A decreases, go short on stock B and long on stock A. This strategy profits from the price movements that tend to move in opposite directions due to the negative correlation.

2. Portfolio Hedging:

Select stocks with a weak or negative correlation to your main holdings. When the integrated OFI of your main holdings suggests a potential price decline, hedge your portfolio by going long on the negatively correlated stocks. This can help mitigate potential losses in your main holdings.

3. Sector Analysis:

Analyze the heatmap to identify clusters of stocks with strong positive correlations. These clusters might represent sectors with similar market movements. Use this information to assess the overall sector sentiment and potential risks or opportunities.

Important Considerations

The heatmap only reflects the correlation between integrated OFI and not necessarily the price movements themselves. Other factors can influence stock prices. The identified correlations might change over time due to market dynamics. Regularly monitor the heatmap to adapt your strategies. Back-test your trading strategies using historical data to assess their effectiveness before deploying them with real capital. By effectively using the heatmap and understanding the correlations between integrated OFI, you can develop informed trading strategies that leverage the potential for cross-impact relationships between stocks. The findings suggest that OFI is a reliable indicator for predicting short-term price movements, particularly in the technology sector. The higher correlation in this sector could be due to faster reactions to order book changes, as tech stocks often experience higher volatility. The weaker relationship in healthcare stocks could be attributed to lower market liquidity or different trading dynamics in this sector.

Order Book Analysis: The heatmap can help traders understand the relationships between different order book variables. For example, traders can use the correlations between bid and ask prices at different levels to assess the overall market sentiment and potential price movements.

OFI Interpretation: The analysis suggests that OFI can provide valuable information about order flow dynamics, but it should be used in conjunction with other technical and fundamental analysis tools.

Strategy Development: Traders can use the heatmap to identify potential trading signals based on the correlations between different variables. However, it is important to backtest any strategies based on these correlations to ensure their robustness and profitability.

Additionally, PCA's application has proven effective in reducing the noise from the multi-level OFI data, simplifying the modeling process and improving prediction accuracy. The cross-impact analysis provides valuable insights into how market movements in one stock can propagate across others, highlighting the importance of sector-specific models. The pairplot reveals a strong positive correlation between the Integrated OFI of AAPL, MSFT, and

TSLA. This information can be valuable for traders in developing pair trading strategies or assessing portfolio diversification. However, it's important to exercise caution and conduct further analysis before implementing any trading strategies based on these observations.



## 4 Conclusion

In conclusion, Order Flow Imbalance (OFI) is a valuable tool for understanding market dynamics and predicting price movements. The integration of PCA improved the model's efficiency, and cross-impact analysis revealed sector-specific trends in trading behavior. Future work could focus on incorporating more features (such as news sentiment or macroeconomic indicators) and applying machine learning models like Random Forests or LSTMs to enhance prediction accuracy further.

## 5 References

1. Rama Cont, Mihai Cucuringu Chao Zhang (2023) Cross-impact of order flow imbalance in equity markets, *Quantitative Finance*, 23:10, 1373-1393, DOI: 10.1080/14697688.2023.2236159