Order Flow Imbalances (OFI)

Godfred Antwi Koduah

May 2025

Answers to Conceptual Questions

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

The multi-level Order Flow Imbalance (OFI) measurement is motivated by:

- **Information content in deeper levels**: The best bid/ask levels (level 0) represent only a fraction of total market liquidity. Deeper levels contain additional information about:

$$\text{Market Depth} = \sum_{m=1}^{10} (\text{BidSize}_m + \text{AskSize}_m) \tag{1}$$

- **Strategic order placement**: Traders use varied strategies:

Order Type	Typical Level
Immediate execution	Best (level 0)
Patient trading	Deeper levels (2-5)
Hidden liquidity	Deepest levels (6-10)

- **Improved explanatory power**: As shown in Cont et al. (2023), integrated multi-level OFI explains 87.14% of price movements versus 71.16% for best-level OFI alone.

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

The authors prefer LASSO regression due to:

1. **Dimensionality**: For N assets, cross-impact requires estimating N^2 parameters. With typical $N \approx 100$, OLS becomes infeasible:

$$\text{Parameters} = N^2 = 10,000 \gg \text{Sample Size} \tag{2}$$

2. **Sparsity**: The true cross-impact matrix β is sparse. LASSO enforces this through:

$$\min_{\beta} \left\{ \frac{1}{2n} \|\mathbf{r} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1 \right\} \tag{3}$$

3. **Multicollinearity**: High correlation between asset OFIs ($\rho \geq 0.3$ for 10% of pairs) makes OLS unstable.

3. Why is OFI considered a better predictor of short-term returns than trade volume?

OFI performs better than trade volume in short-term prediction because:

- **Directional sensitivity**:

$$\text{OFI} = \text{Buy Pressure} - \text{Sell Pressure} \tag{4}$$

$$\text{Volume} = \text{Buy Pressure} + \text{Sell Pressure} \tag{5}$$

- **Microstructure foundations**: The Kyle (1985) model shows prices respond to *net* order flow:

$$\Delta p_t = \lambda \cdot \text{OFI}_t + \epsilon_t \tag{6}$$

- **Empirical results**: Cont et al. report OFI explains 71-87% of contemporaneous returns versus ∼40% for volume-based measures.