Modeling Temporary Market Impact

Project Repository

The code, data preprocessing, and figures associated with this report are available at:
https://github.com/oowenn/blockhouse_OA

Problem Setup

We are given high-frequency limit order book data for three tickers: `CRWV`, `SOUN`, and `FROG`. Our objective is not to optimize a specific trading strategy, but to **characterize and model temporary price impact** (slippage) as a function of trade size and market conditions.

We define the **temporary slippage function** for a trade of size x executed at time t as:

$$g_t(x) := \text{average slippage (in dollars or \%) incurred when trading } x \text{ shares at time } t.$$

Our goal is to:

- Understand how $g_t(x)$ behaves across different tickers and times,
- Evaluate how well different models fit the observed behavior,
- Discuss implications for volume allocation over time.

Data Normalization

To ensure comparability across tickers, we normalize slippage by the **midprice** at each time t. The normalized slippage becomes:

$$\tilde{g}_t(x) := \frac{g_t(x)}{P_{\text{mid},t}}$$

expressed as a percentage. This standardization helps account for differences in stock price magnitude across tickers.

Empirical Observations

We visualize $\tilde{g}_t(x)$ across random 1-minute intervals for each ticker and day. Key findings include:

- Slippage increases nearly monotonically with trade size, as expected.
- Some minutes exhibit near-flat curves (indicating deep liquidity), while others show rapid slippage growth (indicating shallow books).
- There is significant intra-day and inter-day variability, even after normalization.

Modeling Approach

We propose a **ReLU-style piecewise-linear model** to describe the normalized temporary slippage function:

$$\tilde{g}_t(x) = \begin{cases} 0, & x \leq x_0 \\ \eta_{t,s} \cdot (x - x_0), & x > x_0 \end{cases}$$

where:

- $\tilde{g}_t(x)$ is the **normalized slippage** at time t,
- x is the trade size (in shares),
- x_0 is a **liquidity threshold**: the trade size that can be absorbed without significant impact,
- $\eta_{t,s}$ is a **slope parameter** encoding market impact per unit of excess volume beyond x_0. It is time-varying and ticker-specific.

Motivation

This model reflects two key stylized facts observed in our exploratory analysis:

1. **Flat region for small trades**: Small orders—often the first executed within each minute—exhibit near-zero slippage.
2. **Linear growth afterward**: Once available liquidity near the midprice is consumed, slippage grows approximately linearly with trade size.

The ReLU function captures this behavior: flat response until a threshold is reached, followed by linear increase. It provides a simple and interpretable form aligned with order book dynamics.

Advantages

- **Interpretability**: Parameters x_0 and η_t have clear economic meaning.
- **Robustness**: Avoids overfitting and generalizes across tickers and times.
- **Market Microstructure Alignment**: Matches observed liquidity patterns.

Limitations

- **Non-smooth transition**: The kink at x_0 may be unrealistic.
- **Estimation challenge**: Requires accurate detection of x_0.
- **No concavity**: Does not capture diminishing marginal impact.

Possible Approaches

- **Empirical Learning**: Fit $g_{t,s}(x)$ on historical data and learn $x_0, \eta_{t,s}$ from features (spread, depth, volatility).
- **Execution Optimization**:

$$\min_{\{x_t\}} \sum_t g_t(x_t), \quad \text{subject to } \sum_t x_t = X_{\text{total}}$$

- **Robustness Evaluation**: Assess parameter stability across days/tickers.
- **Simulation Testing**: Compare naive and optimized execution on historical traces.

Summary

By analyzing normalized impact curves across time and tickers, we observe consistent patterns: a flat-slippage region up to a liquidity threshold, followed by linear or concave cost growth. We propose a ReLU-style model that captures this behavior with interpretable parameters.

This foundation supports both empirical prediction and optimization-based execution scheduling, enabling robust and adaptive strategies.