Inventory Imbalance

June 17, 2025

[1]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

1 Inventory Imbalance as a Predictor of Short-Term Returns

Forecasting short-term asset returns is notoriously difficult due to the high noise-to-signal ratio, low predictability, and the dominance of microstructure effects at high frequencies. Unlike medium- or long-term horizons, where macroeconomic trends and valuations can play a role, short-term price movements are primarily driven by the flow of orders, liquidity provision, and the behavior of market participants reacting to new information. As a result, alpha research in this space often focuses on extracting signals from limit order book dynamics, trade imbalances, and quote adjustments.

One such signal is inventory imbalance, which compares the volume resting on the bid side of the order book to that on the ask side. The intuition is that when there's significantly more volume on the bid than on the ask, it reflects an asymmetry in market pressure; either due to latent demand or strategic behavior by liquidity providers. This imbalance can indicate that buyers are more aggressive or that sellers are pulling back, both of which may precede a short-term upward price adjustment. Conversely, a heavy ask-side imbalance may signal selling pressure and potential short-term declines. Thus, inventory imbalance can serve as a proxy for latent directional intent in the market, making it a valuable candidate for short-horizon alpha modeling.

1.0.1 The Target Variable

Before we think about a predictive feature, we need to decide what we will try to predict with the feature. A standard choice in financial modeling is log-returns, defined as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

We use returns instead of absolute prices because we're not interested in predicting the price level itself, but rather the direction and magnitude of price changes**. Log-returns are approximately normally distributed for small time intervals, additive over time (unlike simple returns), and centered around zero which makes them more suitable as targets for statistical or machine learning models.

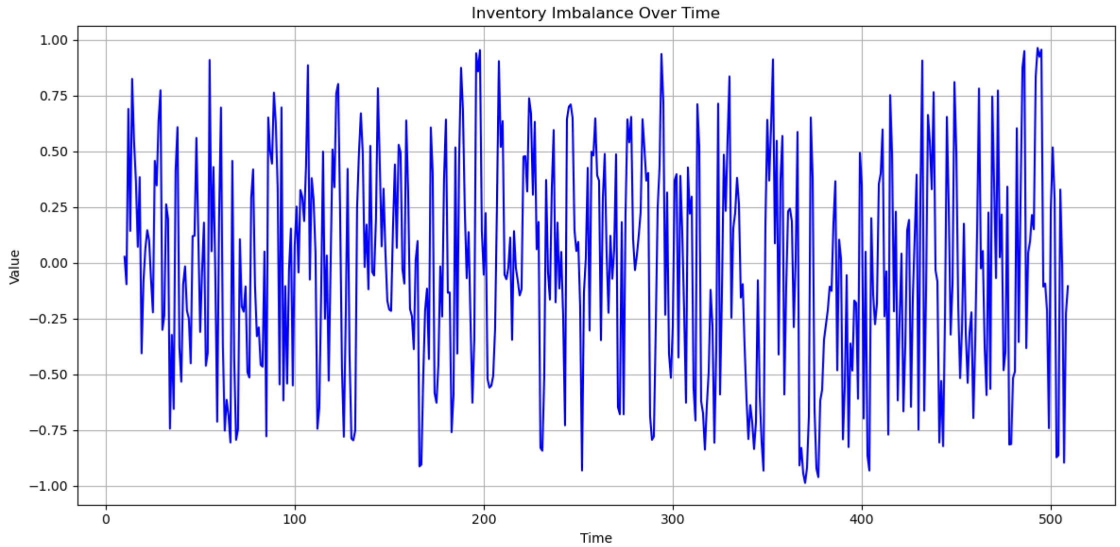
1.0.2 Inventory Imbalance Signal

A key feature we examine in short-term price prediction is **inventory imbalance**, which captures the relative strength of demand versus supply at the top of the order book. The intuition is straightforward: when there is significantly more size on the ask side than the bid side, it suggests downward pressure on price, and vice versa.

We define the inventory imbalance signal as:

$$\text{inv\_imbalance}_t = -1 \times \frac{\text{BS}_t - \text{AS}_t}{\text{BS}_t + \text{AS}_t + \epsilon}$$

where BS and AS are the average quoted sizes at the top of the bid and ask, respectively, and ϵ is a small constant to prevent division by zero. The signal is positive when ask-side depth dominates, indicating potential selling pressure, and negative when bid-side depth dominates, indicating potential buying pressure. We multiply by -1 to align the signal so that positive values suggest expected upward price movement. This normalization ensures the signal ranges roughly between -1 and 1, making it easy to compare across time and conditions.



1.0.3 Information Coefficient (IC)

We often want to evaluate alphas or predictive signals by looking at how well they correlate with future returns. One standard way to do this is by computing the **Information Coefficient (IC)**, which measures the strength and direction of the relationship between the signal and realized future returns.

The **Pearson IC** captures the **linear correlation** between a signal and future returns:

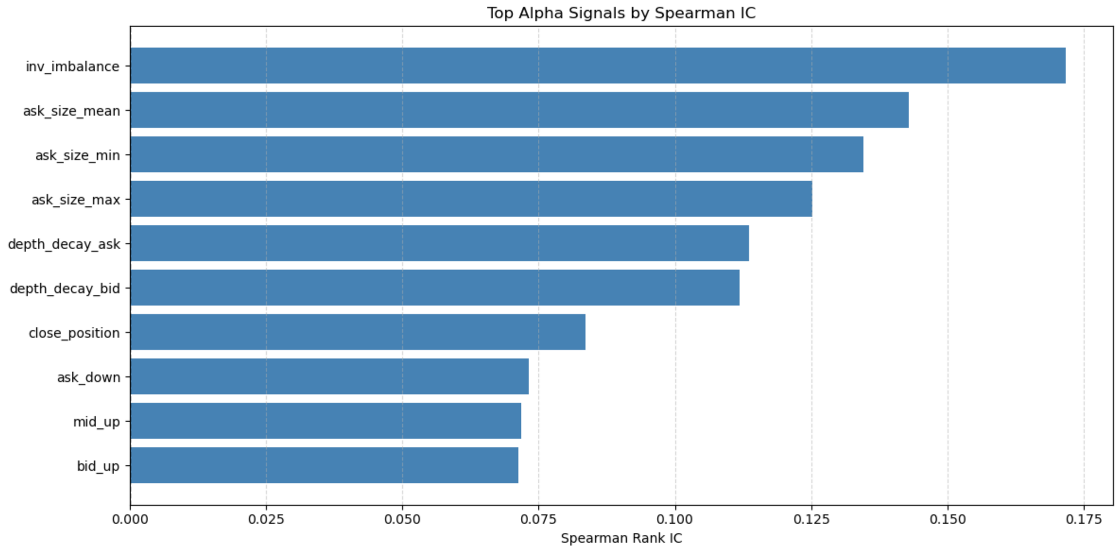
$$\mathrm{IC}_{\mathrm{Pearson}} = \frac{\mathrm{Cov}(\alpha_t, r_{t+1})}{\sigma_\alpha \sigma_r}$$

The **Spearman IC** instead measures the **rank correlation**, capturing whether higher-ranked signals tend to correspond to higher-ranked returns, regardless of their actual magnitude:

$$\mathrm{IC}_{\mathrm{Spearman}} = \mathrm{Corr}(\mathrm{rank}(\alpha_t), \mathrm{rank}(r_{t+1}))$$

Here, α_t is the signal at time t, and r_{t+1} is the realized return at a future time (often a short horizon like 1-minute ahead). A high positive IC indicates that the signal successfully predicts the direction of future returns, while a negative IC suggests an inverse relationship. IC is a widely used metric in quantitative finance to **evaluate the quality of alphas**, where values above 0.05 are generally considered meaningful in high-frequency contexts.

Below is a graph of the Spearman IC of inventory-imbalance signal along with some other computed features.



1.0.4 From Alpha to Strategy

While strong Information Coefficients are a promising first sign, alphas need to be tested much more rigorously before they can be used in live trading. A predictive signal must ultimately translate into an actionable strategy, and that strategy should generate consistent, risk-adjusted returns after accounting for transaction costs, latency, and slippage. Furthermore, when using linear-combinations of alphas, the signals should ideally be linearly-independent / orthogonal, so signals describe unique dimensions of the data. So preparing an alpha for production-trading requires far more through testing and work than we've done here. However, to demonstrate the process of how an alpha might be translated into a strategy, we design and implement a simple trading strategy based on the inventory imbalance signal.

The strategy operates by observing the inventory imbalance It at three evenly spaced intervals. We define:

- It = inv_imbalance$_t$
- It-1 = inv_imbalance$_{t-\Delta}$
- It-2 = inv_imbalance$_{t-2\Delta}$

We generate a **buy signal** when It shows upward momentum across all three snapshots and its absolute value is above a certain threshold:

$$\text{Buy if: It-2} < \text{It-1} < \text{It} \quad \text{and} \quad |\text{It-2}|, |\text{It-1}|, |\text{It}| > \theta$$

Similarly, a sell signal is generated when the signal shows consistent downward momentum and exceeds the same threshold. Once a signal is raised, we simulate entering a trade and holding it for a fixed horizon (etc., 1 second, 5 seconds, 100 seconds), calculating log returns as:

$$r_{t+h} = \ln\left(\frac{P_{t+h}}{P_t}\right)$$

Our crude backtest applies this logic across the dataset and evaluates the strategy using metrics such as cumulative PnL, Sharpe ratio, win rate, and number of trades. This simplified approach helps us assess whether the alpha is capable of generating directional profits under idealized conditions. While it ignores execution frictions, it serves as a useful first filter before considering more complex signal combinations or real-world deployment.