The Next Tick on Nasdaq Stockholm: Predicting Price Direction in Limit Order Books Using Order Imbalance

Jamil Najjar (41021) and Julian Kramer (41032)

A thesis presented for the degree of MSc in Finance
Stockholm School of Economics

December, 2018

Abstract

We explore complete Level II limit order books for eight stocks listed on Nasdaq Stockholm during 2016 and investigate the use of the imbalance between bid and ask volumes in predicting the direction of price change in an ultra-high-frequency environment. Specifically, we test whether a top-of-the-book (Level I) measure of order imbalance and a deeper-in-the-book (Level II) measure can predict the direction of a change in the mid-price of a security for up to three events before the change occurs. Logistic models are used to fit the data and prediction power is judged based on the percent of correctly predicted observations out-of-sample. We find that order imbalance has statistically significant explanatory power in predicting price-change direction and that the logistic predictor considerably outperforms a naive one. In addition, we find that Level I order imbalance is more informative than Level II imbalance and that prediction power quickly decays in the order book. Finally, we compare our findings to those in Gould and Bonart (2016), on which the methodology in this paper is based. We find that unlike the case for US stocks, relative tick size does not play a role in explaining the shape of the order book for Swedish stocks. Rather, we argue that liquidity is the main factor driving order book behavior and price-direction predictability.

Acknowledgments: We would like to thank our supervisor, Dr. Michael Halling, for his support and the Swedish House of Finance for providing access to the Nasdaq HFT database.

1 Introduction

Increasingly, low-latency networking and advancing computing power are changing the reality of trading on electronic exchanges; First, high-frequency trading (HFT) has become prevalent. Hagströmer and Nordén (2013) estimate that Nasdaq Stockholm member firms that were identified as being primarily HFT traders submitted an average of 30% of all orders on OMXS30 stocks during February 2012 and generated 25% of trading volume for the same stocks over the same period. Second, order flow data is key to survival. Brogaard et al. (2017) explain that modern financial markets, which allow anyone to submit limit orders, have changed our view of price discovery, as traditionally trades were seen to represent private information while maker makers' quotes represented public information. They also show that since limit orders far outnumber market orders, price discovery happens mainly through limit orders. Easley et al. (2012) argue that HFTs react strategically to information revealed by low-frequency traders (LFT). In turn, LFT strategies use smart algorithms and limit orders to optimize their trades and avoid predatory HFT behavior. Therefore, market participants constantly analyze market transactions and the arrival and cancellation of limit orders on an exchange for informational content and predictive power. But most importantly, the availability of high-frequency financial data has invited the interest of researchers studying topics such as the price discovery process and modeling dynamics in limit order books. See Khashanah et al. (2014) for a survey of academic research on HFT and Cont (2011) for a survey of high-frequency financial data models. The HFT literature is vast, yet the majority of studies focus on US markets.

In this paper, we focus on price discovery in Swedish high frequency data. We explore complete (event-by-event) limit order books from Nasdaq Stockholm. A limit order book (LOB, order book, or book) is an aggregation of all outstanding limit buy (bid) and sell (ask or offer) orders on an exchange. Figure 1 illustrates a snapshot of a basic LOB. The best bid and ask prices are said to be at the top of the book, or Level I (L1). The next best bid and ask prices are Level II (L2), and so on. The total volume of all orders outstanding at each price, or the length of the queue at each price, is the bid/ask *size* (queue size). In Figure 1, The mid-price can only change if an order is placed inside the spread or if either the best bid queue or the best ask queue is depleted. Queues are depleted when orders are cancelled or are matched (transaction events).

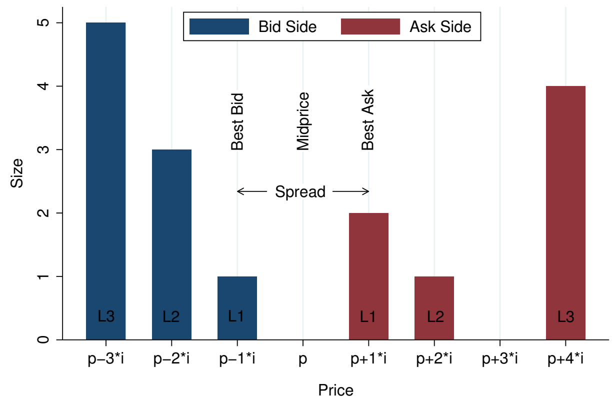


Figure 1: Limit Order Book example. Blue bars represent aggregate buy limit orders at each price. Red bars represent aggregate sell limit orders at each price. The midprice is p, the smallest tick size is i, and the spread is the difference between the highest buy order price (the bid price) and the lowest sell order price (the ask price). In this example, the spread is 2*i.

The *mid-price* (*price* or p) and the *bid-ask spread* (*spread* or s) are defined in Eq.(1.1) and Eq.(1.2), respectively.[1]

$$p_t = \frac{askpx_t^{L1} + bidpx_t^{L1}}{2} \tag{1.1}$$

$$s_t = askpx_t^{L1} - bidpx_t^{L1} \tag{1.2}$$

where p_t is the *mid-price* at time t, $bidpx_t^{L1}$ and $askpx_t^{L1}$ are the the best bid and ask prices at time t, respectively, and s_t is the *bid-ask spread* at time t.

A well-known measure that has been shown to have predictive power over short time intervals is the imbalance between supply and demand in a LOB. Order imbalance between bid and ask orders summarizes the shape, or the state, of the order book reflecting investors' information and incentives to trade. Gould and Bonart (2016) find a strong statistically significant relationship between order imbalance and the direction of the subsequent mid-price movement on Nasdaq. Cao et al. (2004) study the Australian Stock Exchange and find that order imbalance explains future short-term returns. Goldstein et al. (2017) also study the Australian market and find that order imbalance is a strong predictor of mid-price

[1] It is important to keep in mind that ultra-high frequency data is irregularly spaced because events arrive in unequally spaced time intervals. Therefore, we treat time, t, throughout this paper as event-based time. For a discussion, see Easley et al. (2012).

movement. Cont et al. (2014) show that over short time intervals, price changes of US stocks are mainly driven by the order flow imbalance. Zheng et al. (2012) analyze French stocks and find that the liquidity on the best bid and ask, in addition to other variables, is informative for predicting incoming market orders.

We define order imbalance (book imbalance, imbalance, or OI) as the normalized difference between bid and ask volumes and measure it at the first two levels in the book (Equations (1.3) and (1.4)). The normalized difference scales the imbalance into the range $[-1, 1]$. While most studies we surveyed focused on L1 imbalance, the fact that L1 volumes are smaller than L2 volumes makes L2 (and deeper-in-the-book) imbalances of interest. Cao et al. (2004) find that L2 to L10 imbalances provide addition power in explaining future short-term returns. Cont et al. (2014) find that L2 imbalance adds a small increase to the explanatory power in their regression while L3 to L5 imbalances' contributions can be neglected.

$$OI_t^{L1} = \frac{bidsz_t^{L1} - asksz_t^{L1}}{bidsz_t^{L1} + asksz_t^{L1}} \tag{1.3}$$

$$OI_t^{L2} = \frac{(bidsz_t^{L1} + bidsz_t^{L2}) - (asksz_t^{L1} + asksz_t^{L2})}{(bidsz_t^{L1} + bidsz_t^{L2}) + (asksz_t^{L1} + asksz_t^{L2})} \tag{1.4}$$

where OI_t^{L1} and OI_t^{L2} are the Level I imbalance and Level II imbalance at time t, respectively, $bidsz_t^{L1}$ is the best bid size, $bidsz_t^{L2}$ is the second-best bid size, $asksz_t^{L1}$ is the best ask size, and $asksz_t^{L2}$ is the second-best ask size,

Equation (1.3) matches the definitions found in most studies we surveyed (see Gould and Bonart (2016), Goldstein et al. (2017), Brogaard et al. (2014), and Lipton et al. (2013)) while Eq.(1.4) is an extension we make of Eq.(1.3) to include all orders at the first two levels in the book. Some researchers define order imbalance differently. For example, Zheng et al. (2012) define imbalance as the log ratio of bid size to ask size while Cao et al. (2004) use values ($size * price$) rather than size to calculate book imbalance. It is also worth keeping in mind that our measure of order imbalance does not take into account any hidden orders. Hidden orders may be used when orders are large and/or when the market is illiquid. Therefore, not including them may distort our estimate of market supply and demand. See Avellaneda et al. (2011) for a statistical model of order imbalance with hidden liquidity. Since our study is focused on the most liquid stocks, we estimate that hidden orders should have a minimal impact on our results.

In this paper we investigate using L1 and L2 order imbalances to predict the direction of a change in the price of a security in the short period of time before the price change occurs.

The relationship between these two variables is modeled with a simple logistic model. We benchmark goodness-of-fit for the logistic model against that of a naive model based on the percent of correctly predicted observations in-sample. Prediction power is assessed based on the percent of correctly predicted observations out-of-sample. We follow the general methodology and outline in Gould and Bonart (2016), in which US stocks were investigated. An important finding in their paper is that the behavior and predictability across stocks differ considerably depending on the relative tick size of a stock, which is the ratio of price to the minimum possible tick size. O'Hara et al. (2018) find that relative tick size affects liquidity provision and investor behavior. We investigate whether the same observations and effects can be found in Swedish data.

The rest of the paper is organized as follows. Section 2 describes the stock selection process, the selected stock sample, grouping stocks into big-tick and small-tick groups, preparing the HFT dataset, and statistical properties of the HFT dataset. Section 3 details the methodology, the models, the prediction process, and the assessment criteria. Section 4 discusses statistical analysis of order imbalance, regression results, and prediction performance. Section 5 discusses the results and compares them to the case for US data. Section 6 summarizes our findings and discusses caveats and areas of improvements.

2 Data

2.1 Stock Sample Selection

We choose to study eight equities traded on Nasdaq Stockholm during 2016, which had 253 trading days, including four days of half-day trading.[2] We reason that data for a full year would not only contain a large number of observations, but also cover seasonal effects, such as earnings season and macro events, which may be reflected in market volatility and the price-formation process. We use Thomson Reuters Eikon to screen through a list of the most liquid stocks, by average daily trading value, that were traded on Nasdaq Stockholm during 2016. We exclude stocks that are dual-listed or have very active ADRs on other exchanges.

Next, we consider relative tick sizes. Gould and Bonart (2016) use relative tick size to group stocks into large-tick and small-tick stocks and find strong evidence of different LOB shape and price predictability between the two groups. The relative tick size is defined as the ratio between the stock price and the tick size (smallest change in price allowed by the

[2] Source: https://www.nasdaqomxnordic.com/digitalAssets/102/102819_tsn-fixed-trading-calendar-2016-2018.pdf

exchange). The idea is that although the tick size can be the same across all stocks, the lowest-priced stocks will have the largest relative tick sizes. For example, if the tick size is \$0.01, then it represents 1% of the value of a stock trading at \$1, and 0.01% of the value of a stock trading at \$100. In EU markets, however, tick sizes increase in steps depending on the price range.[3] Yet, these 'dynamic' tick sizes are not continuous which means that there will be large-tick and small-tick stocks within each price range. Therefore, to classify stocks into large-ticks and small-ticks we build a relative tick size score, which we find by using the average daily relative tick size for each stock and then normalizing that value using the average and standard deviation for the whole group (40 of the most liquid stocks on the exchange). This score would give a clear indication of how far away each stock's average relative tick size is from the group's average, and it is an objective and intuitive way of comparing the relative size of a one-tick change in the price of an instrument as compared to other stocks in the group. The higher the score, the larger the price impact is. For a discussion on relative tick size and tick size regimes see Verousis et al. (2018) and O'Hara et al. (2018).

Finally, eight stocks are chosen to represent variation across sectors, market capitalizations, relative tick size scores, trading volumes, and overall price performance. Table 1 summarizes trading statistics for each stock. Half of the group were selected to be large-tick stocks (SHB, ALFA, SECU, and SOBI) while the other half consists of small-tick stocks (HM, INVE, SWMA, and HEXA). Note that except for HEXA and SOBI, all stocks were members of the OMXS30 index during 2016. Figure 2 compares their daily price performance over the year.

2.2 HFT Data

For each stock, and for each trading day in 2016, reconstructed Level II order books are obtained from the Nasdaq HFT database at the Swedish House of Finance Date Center. The order books are reconstructed from historic Nasdaq ITCH feeds and contain nanosecond-stamped views of the top two levels of the order book at each event. Figure 12 in the appendix shows an example of the raw data and contains variable definitions. An event may represent a transaction or the arrival, modification, replenishing, or cancellation of a limit order at the best bid/ask (Level I) or the second-best bid/ask (Level II).[4] Therefore, an event may

[3] For details about the tick size schedules on Nasdaq Stockhholm, refer to https://business.nasdaq.com/Docs/INET-Nordic-Market-Model.pdf

[4] For details about order types and the ITCH feed, see http://www.nasdaqomx.com/digitalAssets/102/102135_nordic-equity-totalview-itch-3.01.pdf

Stock	Sector	P_{high} (SEK)	P_{low} (SEK)	M.Cap (BSEK)	Volume (mil)	Value (MSEK)	Relative tick size	Return (%)
HM	Consumer	305.1	234.5	383	2.9	768	-1.12	85
SHB	Financials	134.7	90.25	216	3.6	392	1.72	114
INVE	Financials	345.6	256.8	229	1.1	319	-1.35	111
SWMA	Consumer	318.9	249.7	56	0.6	186	-1.31	98
HEXA	Industrials	380.9	260.7	111	0.6	185	-1.49	105
ALFA	Industrials	154.4	121.3	56.7	1.4	182	1	99
SECU	Consumer	152.9	110	47.2	1.15	152	1.04	112
SOBI	Healthcare	133.3	89.5	28.6	1.1	119	1.56	80

Table 1: Summary trading statistics for eight stocks trading on Nasdaq Stockholm during 2016. The stocks are HM (H&M), SHB (Handelsbanken), INVE (Investor), SWMA (Swedsih Match), HEXA (Hexagon), ALFA (Alfa Laval), SECU (Securitas), and SOBI (Swedish Orphan Biovitrum). For each stock, the table shows the sector, highest and lowest price (SEK), average market capitalization (SEK bil), average daily volume traded (mil), average daily value traded (SEK mil), relative tick size score, and price performance (return) over the full year. The stocks are ordered according to average value traded.

not necessarily correspond to a change in the mid-price. Figure 3 demonstrates this visually with an example from the dataset.

Nasdaq Stockholm equities' continuous trading hours are 09:00-17:25 on regular trading days and 09:00-12:55 on half-day trading days. To avoid the more volatile conditions and abnormal order flow around the opening and closing of the market which may introduce noise into our data, we remove the first and last 30 minutes of trading (See Gould and Bonart (2016) and Cao et al. (2004) for similar treatment of intra-day data). Figure 4 compares price patterns for the eight stocks on Jan 4, 2016. The first 30 minutes see strong price adjustment as investors price in new information that became available before market open. The data is also cleaned of any non-continuous trading events.

Another reason for possible noisy behavior closer to market open and close is the fact that many day-trading strategies start building portfolios at market open and close their positions by market close (see Aldridge (2013)). Indeed, this is true for market makers who aim to keep low inventories and may have accumulated unwanted positions by end of day (see Biais and Foucault (2014)). Figure 5 shows larger trading volumes for HM shares towards the beginning and end of trading on Jan 4, 2016.

We summarize selected information from the order books of each stock in Table 2. Although the number of instruments is small, some interesting observations can be made by looking at Tables 1 and 2 together. First, the average daily number of events is correlated

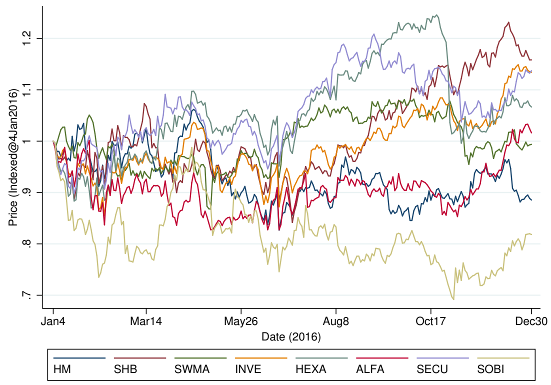


Figure 2: Daily closing price patterns for eight stocks trading on Nasdaq Stockholm during 2016. Each stock's prices are indexed at 100% on the first trading day (Jan 4, 2016). (253 days/points.)

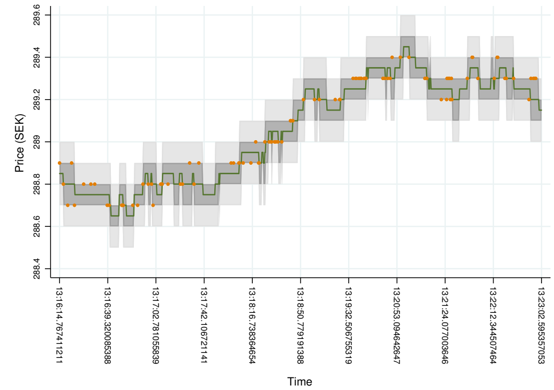


Figure 3: Price quotes for HM shares during a window of 1000 events (about 8 minutes) on Jan 4, 2016. The dark shaded area depicts the best bid and ask. The light shaded area depicts Level II bids and asks. The green line is the mid-price. Transactions are marked with orange circles. (1000 points.)

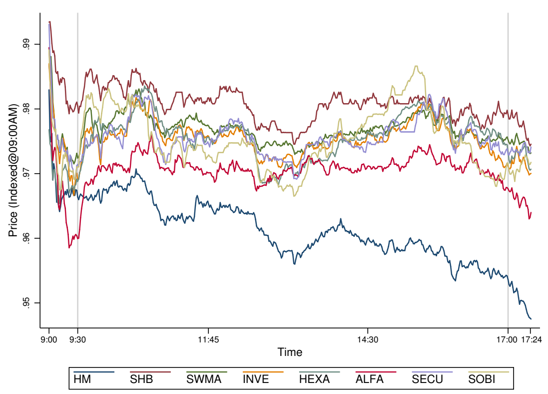


Figure 4: Intra-day price patterns for eight stocks trading on Nasdaq Stockholm on Jan 4, 2016 during regular trading hours. (505 minutes/points.)

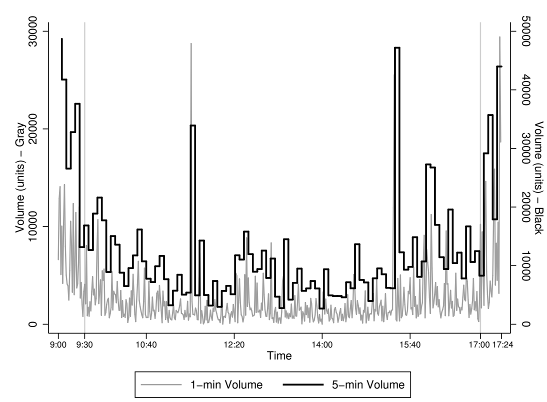


Figure 5: Intra-day transaction volumes smile for HM shares on Jan 4, 2016 during regular trading hours. 1-minute volumes are plotted in gray (left axis, 505 points). 5-minute volumes are plotted in black (right axis, 100 points).

with our measure of liquidity (average daily value traded). Second, the average spread is inversely correlated with the average daily volume traded. Third, and most interestingly, the number of upticks (as a percentage of the number of price changes) is shown to be unrelated to the overall price performance over the year, which suggests that price discovery is ultimately determined by price jumps that are greater than the minimum tick size. In addition, this may be a confirmation of the fact that the majority of HFT activity represents market making, which takes both sides of the spread. Menkveld (2013) finds that HFTs predominantly earn the spread, as most of their trades are passive, and suffer losses on their net positions. Hagströmer and Nordén (2013) find that market makers constitute about 86% of HFT limit order traffic on Nasdaq Stockholm.

Stock	Events (1000's)	Price-changes (1000's)	Upticks (% of Price-changes)	Spread (SEK)	Bid size (units)	Ask size (units)
HM	57	4.3	50.23	0.15	2953	3439
SHB	36	2.5	50.17	0.14	10766	10307
INVE	46	3.3	50.31	0.16	1789	1701
SWMA	29	2.7	50.26	0.17	966	929
HEXA	17	4.1	50.57	0.22	435	451
ALFA	31	2.0	50.12	0.15	4164	3960
SECU	22	1.4	50.26	0.15	4008	3829
SOBI	10	1.6	50.04	0.15	2104	2026

Table 2: Summary LOB statistics for eight stocks trading on Nasdaq Stockholm during 2016. The stocks are HM (H&M), SHB (Handelsbanken), INVE (Investor), SWMA (Swedsih Match), HEXA (Hexagon), ALFA (Alfa Laval), SECU (Securitas), and SOBI (Swedish Orphan Biovitrum). For each stock, the table shows average number of daily events (1000's), average number of daily mid-price changes (1000's), percent of mid-price changes that were upticks, average spread (SEK), average size of the best bid (SEK), and average size of the best ask (SEK).

In addition, we consider a representative day and stock in order to make additional exploratory observations about the ultra-high frequency environment. For this we choose to look at HM shares on Jan 4, 2016 between 09:30 and 17:00. Within this time frame, there were 83989 events, 6503 price changes (ticks), and 5602 transactions. We find that 88% of these transactions did not coincide with a change in price. Indeed, transactions do not necessarily have to deplete a bid/ask queue, orders can be matched with hidden orders inside the spread, and matched iceberg orders may replenish at the best bid/ask, leaving the mid-price unchanged. Interesting as well is how fast the first two levels of the order book evolve. We calculate the median time between events was 766 microseconds and the median time between price changes was 370 microseconds, which indicates clustering in the data. We also calculate that the median time between a price-change event and the event preceding it is 190 microseconds. These numbers are compared to a median time of 17 seconds between transactions and explain the higher levels of investment in computing and networking resources that HFT trading strategies require.

We analyze the price changes, or ticks, within this short window further. Table 3 is a tabulation of the sizes of the 6503 price changes. Note that almost all (96.71%) price changes are equal to the smallest possible tick size of 0.05. While these are almost balanced, the few price jumps greater than the smallest possible tick size are imbalanced and explain the cumulative price change of -4.00 SEK over the day. Furthermore, we investigate whether there is serial correlation in price changes. For this, we create a binary indicator of the direction of the same 6503 price changes. For an up-tick, we give it a value of 1, and a value of 0 for a downtick. Figure 10 in the appendix depicts the partial autocorrelation of the tick direction indicator. The figure shows significant negative serial correlation in the direction of price changes.

Tick size (SEK)	Frequency (units)	Percent (%)	Cumulative % (%)
-0.15	1	0.02	0.02
-0.10	138	2.12	2.14
-0.05	3,122	48.01	50.15
0.05	3,167	48.70	98.85
0.10	71	1.09	99.94
0.15	4	0.06	100.00

Table 3: Tabulation of tick sizes for HM shares on Jan 4, 2016. (6503 price changes (ticks).)

Finally, to visualize the relationship between order imbalance and price changes, Figure 6 plots Level 1 and Level 2 order imbalances (Equations (1.3) and (1.4)) against the mid-price of HM stock in a short window of 100 events on Jan 4, 2016. Although 100 events is too small a sample to base any conclusions on, the patterns in the plot are interesting. Both L1 and L2 order imbalances appear to lead the price signal in the few events before a tick arrives.

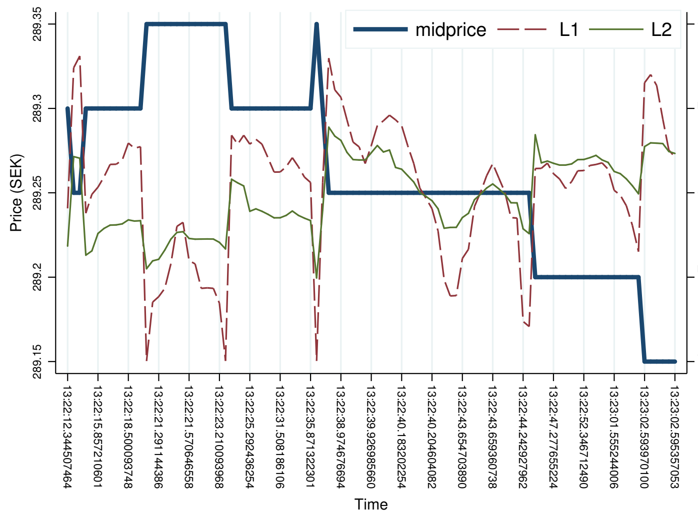


Figure 6: Mid-price quotes of HM shares (blue line, left axis) against LI order imbalance (red dashed line, right axis) and L2 order imbalance (green line, right axis) for a short period of 100 events during continuous trading hours on Jan 4, 2016. (100 points.)

3 Methodology

From the discussions in Section 1 and 2, we hypothesize that order imbalance has power in predicting the sign of the next tick within the very small window of time before the tick arrives. In this section, we develop the methodology for testing the relationship between the direction of price-changes (tick sign) and the order imbalance. We model a simple relationship between both variables and test model performance both in-sample and out-of-sample.

3.1 Dataset Construction

To construct the dataset, we start with calculating the *mid-price* (p_t), *spread*, (s_t) and order imbalance variables (OI_t^{L1} and OI_t^{L2}) using Equations (1.1), (1.2), (1.3) and (1.4), respectively, for each stock. We also create a price-change indicator by only considering observations where the mid-price changes (i.e., $p_t \neq p_{t-1}$). From this, we construct a price-change direction indicator, y_t, such that

$$y_t = \begin{cases} 1 & \text{if } p_t > p_{t-1} \\ 0 & \text{if } p_t < p_{t-1} \end{cases} \tag{3.1}$$

y_t will become our binary dependent variable and it indicates whether the price-change was upward or downward (positive and negative ticks). The decision to restrict our focus to the direction (sign) of price changes is due to the limited variation in the magnitudes of price changes, which is expected when working in an event-by-event environment. (See Table 3 for an example.)

Next, for each price change, we calculate order imbalances at each of the three preceding events. While Gould and Bonart (2016) chose to measure order imbalance at a random point in the window between each price change, we believe that focusing on the few events preceding a price change is more intuitive and more relevant for practical applications. Further, it allows us to compare the informativeness of these 'lags' of OI before a price change.

Importantly, we want to have an equal number of datapoints for each stock and give each trading day an equal weighting in the data. To achieve this, we draw 100 uniform random samples from each trading day, which yields a final dataset of 253,000 observations for each stock. In addition to requiring less computing power, the small number of daily samples and the random draws help avoid the strong serial correlation in high-frequency data.

Finally, we randomize and split the data into a training set (80%) and a testing set (20%).[5] The training set will be used to fit the models and the testing set to assess out-of-sample performance.

3.2 Logistic Models

We model the relationship between the direction of price change, y, and order imbalance, OI, using a logistic regression.[6] The logistic model is also used by Zheng et al. (2012) to model the relationship between order imbalance, liquidity, and the arrival of market orders. The logistic regression is a dummy (binary) dependent variable model (DDM) for a binary variable y and a continuous explanatory variable OI. The model introduces a continuous latent dependent variable, y^*, such that

$$y = \begin{cases} 1 & \text{if } y^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

[5]The 80-20 split is arbitrary. Any split is acceptable as long as the number of datapoints is large enough in each set and the sets are mutually exclusive.

[6]For more about the logistic regression, see Hosmer Jr et al. (2013) and Wooldridge (2015).

then defines the following relationship between y^* and OI

$$y^* = \beta_0 + \beta_1 * OI + e$$

where e is an error term that is assumed to be independent of OI. The logistic model assumes that the error term follows a logistic distribution and the model is interpreted probabilistically such that the probability that $y = 1$ is determined by the Cumulative Distribution Function (CDF) of the error term. The model is estimated by maximum likelihood and it forecasts probabilities (of observing $y = 1$):

$$P(y = 1 \,|\, OI) = \Lambda(b_0 + b_1 * OI) = \frac{exp(b_0 + b_1 * OI)}{1 + exp(b_0 + b_1 * OI)}$$

where $\Lambda()$ is the CDF of the logistic function.

Using the logistic model, we formalize our hypothesis of the existence of a direct relationship between y and OI and estimate the following models using the training set:

$$P(y_t = 1 \,|\, OI^{L1}_{t-1}) = \Lambda(b_0 + b_1 * OI^{L1}_{t-1}) \tag{A1}$$

$$P(y_t = 1 \,|\, OI^{L1}_{t-2}) = \Lambda(b_0 + b_1 * OI^{L1}_{t-2}) \tag{A2}$$

$$P(y_t = 1 \,|\, OI^{L1}_{t-3}) = \Lambda(b_0 + b_1 * OI^{L1}_{t-3}) \tag{A3}$$

$$P(y_t = 1 \,|\, OI^{L2}_{t-1}) = \Lambda(b_0 + b_1 * OI^{L2}_{t-1}) \tag{B1}$$

$$P(y_t = 1 \,|\, OI^{L2}_{t-2}) = \Lambda(b_0 + b_1 * OI^{L2}_{t-2}) \tag{B2}$$

$$P(y_t = 1 \,|\, OI^{L2}_{t-3}) = \Lambda(b_0 + b_1 * OI^{L2}_{t-3}) \tag{B3}$$

where y is the price-direction indicator, OI is order imbalance, and $\Lambda()$ is the logistic function CDF. Models (A1), (A2), and (A3) estimate the probability that a price change is an uptick given L1 imbalance values at the 1, 2, and 3 events ('lags') that precede a price-change event, respectively. Models (B1), (B2), and (B3) estimate the same probability given L2 imbalance values at the 1, 2, and 3 events ('lags') that precede a price-change event, respectively.

3.3 Assessing Goodness-of-Fit and Prediction Performance

Goodness-of-fit and prediction performance are assessed by measuring the percent of observations correctly predicted for both the training and testing subsets, respectively. We compare these to the performance of a naive estimator. In addition, the area under the Receiver Operating Characteristic (ROC) curve is calculated for all the models for comparison.

Once the logistic models (Equations (A1) to (B3) in Section 3.2) for each stock are fitted using the training subset, we use them to forecast probabilities (of observing $y = 1$) both in-sample (using the training subset) and out-of-sample (using the testing subset). The naive estimator is generated by creating uniform random probabilities (of observing $y = 1$). Next, the forecasted probabilities of each stock's models and the random probabilities of the naive model are classified into binary predictions using the rule in Eq. (3.2). Finally, we compare these predictions to the actual values to find the percent of observations that our models and the naive estimator predict correctly. For convenience, we call this measure $\tilde{R}^2$.

$$\hat{y} = \begin{cases} 1 & \text{if } P(y = 1 \,|\, OI) > 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

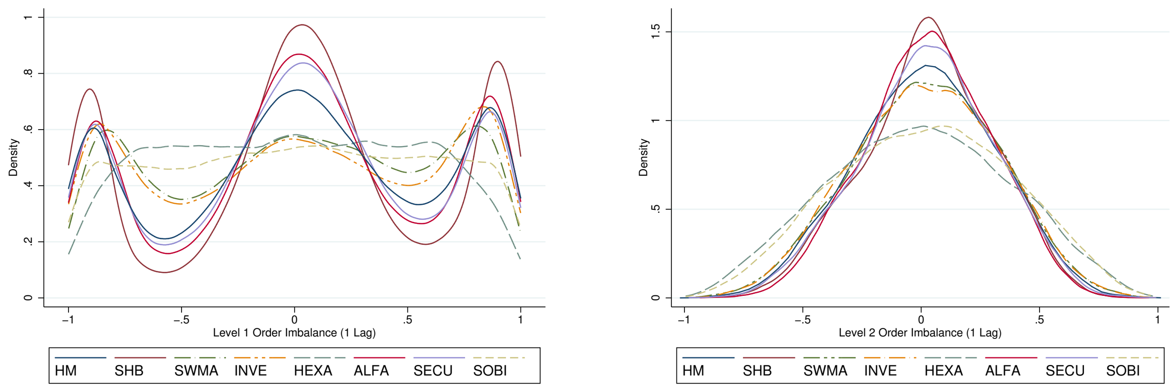
While the measure of Percent Correctly Predicted (PCP) is convenient, it may hide weakness in correctly predicting the least likely outcome.[7] A more accurate way to assess performance is to measure both the sensitivity (percent of correctly predicted upticks) and specificity (percent of correctly predicted downticks). However, these measures will obviously depend on the cut-point used to classify probabilities in Eq. (3.2). In calculating PCP, we use 0.5 as a cut-point since upticks and downticks have an equally likely probability in our data (see Table 2). However, this may not always be the case. For example, consider a trading algorithm that trains its model in a rolling window. In this case the ratio of upticks to downticks will keep changing. Therefore, researchers like to plot ROC curves that show sensitivity against specificity for cut-points in the range $[0, 1]$ and summarize a model's discrimination ability by measuring the area under this curve (called AUC). For each stock, we plot out-of-sample ROC curves and measure AUC for all our models in order to better assess and compare discrimination ability and choose the best performing ones.

[7] For an example, visit http://thestatsgeek.com/2014/05/05/area-under-the-roc-curve-assessing-discrimination-in-logistic-regression

4 Results

4.1 Order Imbalance

We start by studying the covariate, OI. The distributions of the first lags of Level I and Level II order imbalances (OI^{L1}_{t-1} and OI^{L2}_{t-1}) for each stock are presented in Figure 7.



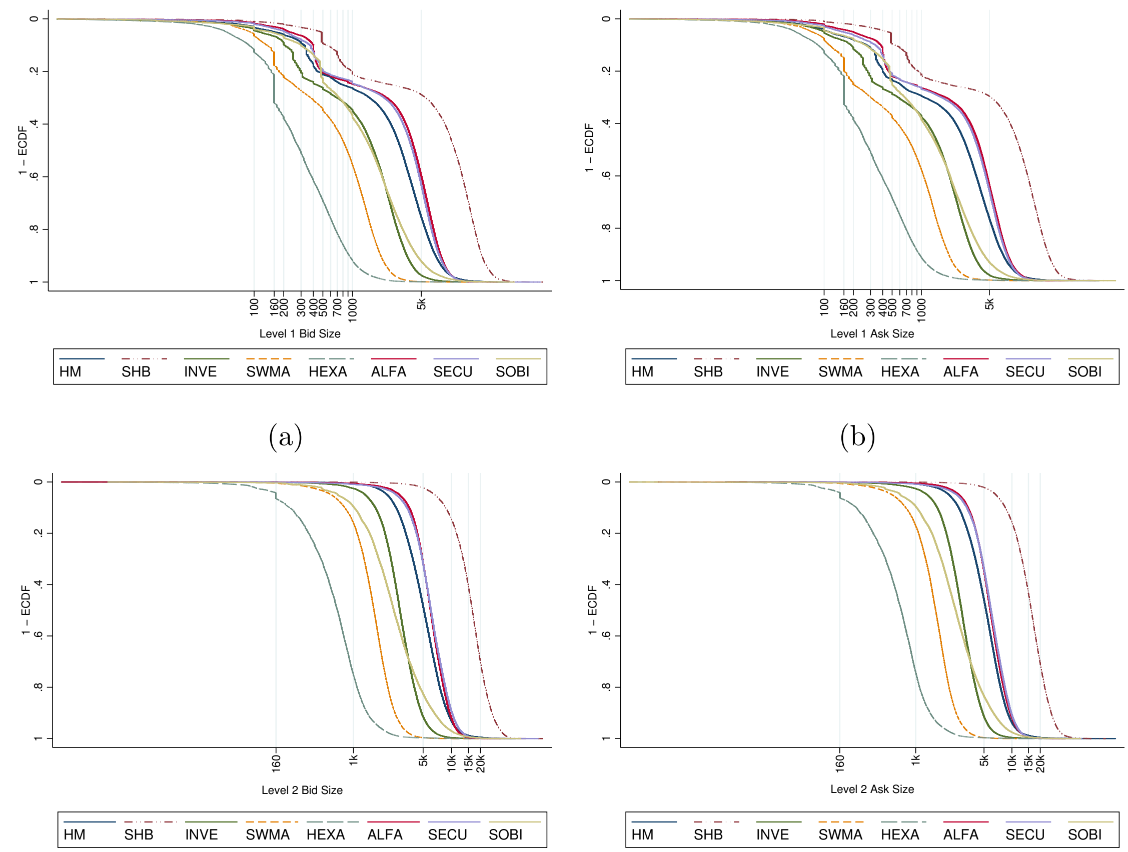
(a) (b)

Figure 7: Frequency distributions of the 1-lag L1 order imbalance, OI^{L1}_{t-1} (panel (a)) and L2 order imbalance, OI^{L2}_{t-1} (panel (b)), for each stock. (100 points/day, total 25300 points for 2016.)

From the distributions of OI^{L1}_{t-1} and OI^{L2}_{t-1}, we observe that in general, both levels of the imbalance show continuous values in their full domain $[-1, 1]$ and peak densities that are centered around 0, which is when bid and ask sizes (volumes) are equal. While L2 imbalance is closer to a normal distribution with almost zero densities at the edges (-1 and 1), L1 imbalance is trimodal with two more peaks, or shoulders, around the values -0.85 and 0.85. L1 distributions show that imbalance values around -0.5 and 0.5 are the least common. A second observation we make is that stocks behave differently; we divide them into three groups. The first group members (HM, SHB, ALFA, and SECU) have the most variation in their L1 distributions, and their L2 distributions have high kurtosis. The second group members (INVE and SWMA) behave more moderately. The most salient are members of the third group (HEXA and SOBI), which have almost flat L1 distributions in the range $[-0.7, 0.7]$ for HEXA and $[-0.9, 0.9]$ for SOBI and no shoulders. Their L2 distributions have fat tails and much less kurtosis than the other six stocks. Our third observation is about the asymmetry of L1 distributions. While L2 distributions are almost symmetric, L1 distributions are not: positive imbalances have higher densities (higher troughs and higher peaks in Figure 7(a)) than negative ones, for all stocks. In addition, we compare the distributions of the three lags of L1 and L2 imbalances for HM in Figure 11 in the appendix. The plot indicates that as we move closer to a price change, imbalance distributions become less flat and have more variation. The variation between successive lags is stronger for L1 imbalances than it is for L2 ones. The same is true for all eight stocks.

Interestingly, unlike the order imbalance distributions found by Gould and Bonart (2016) for US stocks, we do not observe a difference between the distributions of small-tick stocks and those of large-tick ones. This may be due to several factors. First, our methodologies differ. The relative tick size score we introduce to classify stocks into big-tick and small-tick can be inaccurate or irrelevant. Also, we measure order imbalance in the few events before price-changes while Gould and Bonart measure OI at a random point in the interval between two price-changes. Second, there exist differences in the liquidity and structure of US and Swedish markets. For example, the maker-taker model in the US, whereby an order that makes a market (adds liquidity) receives a rebate instead of paying a commission, is not adopted by Nasdaq Stockholm. Therefore, market players make their decisions based on different incentives and, consequently, will not be using the same trading strategies in both markets. This is expected to be reflected in the shape of the order book. See Chan (2017) for more on the maker-taker model.

Moreover, we find it useful to study the distributions of bid and ask sizes (volumes). The Empirical Cumulative Density Functions (ECDFs) of L1 and L2 bid and ask sizes are plotted in Figure 8. We notice that for L1 queues, round numbers are common to all stocks, but to different degrees. For example, a best ask size of 160 makes about 11% of all best ask sizes in the data for HEXA, while the most common ask size for SHB is 500. Also, small-tick stocks see smaller round sizes, as expected, on both the bid and the ask sides. For sizes greater than 1000, all curves are smooth. L2 queues have smooth curves, except for HEXA which shows a small jump at 160. Here also we can see that small-tick stocks' curves start decreasing at sizes smaller than those where large-tick stocks do. In other words, small-tick stocks witness higher frequencies of small queue sizes, which are also more likely to be round numbers. This is expected and is simply due to their higher stock prices. Gould and Bonart (2016) find an even stronger difference in behavior between small-tick and large-tick US stocks in their paper. Their ECDFs for large-tick stocks are smoother than the ones we plot for Swedish ones and their ECDFs for small-tick stocks have much sharper declines at round numbers. This explains why their OI distribution plots contain spikes while our plots are smooth.



(a) (b)

(c) (d)

Figure 8: Empirical Cumulative Density Functions (ECDF) of queue sizes for each stock. Panel (a): L1 bid size distributions. Panel (b): L1 ask size distributions. Panel (c): L2 bid size distributions. Panel (d): L2 ask size distributions. For convenience, y-axis shows (1-ECDF) and x-axis is in logs. (100 points/day, total 25300 points for 2016.)

4.2 Regression Results

The models in Section 3.2 were fitted using the training subsample. The regression results are presented in Table 4. First, we note that the intercepts, which can be interpreted as baseline probabilities (i.e., the probability that the next tick is positive when the imbalance is zero), have small magnitudes and are statistically insignificant for most cases. For example, the largest magnitude for an intercept is that for SHB Model (A1) and translates to a 44.77% probability of the next tick being up, given a balanced book.[8] In general, the small magnitudes are expected because positive and negative ticks should be equally likely when the order book is balanced. Second, all slope estimates are statistically significant at the 99%

[8]To convert a 'log odds' (x) to a probability (p): p=exp(x)/(1+exp(x))

L1 models	1-lag (A1)		2-lag (A2)		3-lag (A3)	
	$b_0\,(se)$	$b_1\,(se)$	$b_0\,(se)$	$b_1\,(se)$	$b_0\,(se)$	$b_1\,(se)$
HM	-.16* (.019)	3.3* (.05)	-.002 (.017)	2.3* (.03)	-.004 (.016)	2.0* (.03)
SHB	-.21* (.018)	3.3* (.05)	-.02 (.016)	1.9* (.03)	-.01 (.016)	1.8* (.03)
INVE	-.06* (.019)	3.3* (.04)	.03 (.017)	2.5* (.04)	.03 (.016)	2.1* (.03)
SWMA	-.07* (.018)	3.1* (.04)	-.003 (.016)	2.3* (.04)	.004 (.016)	1.9* (.03)
HEXA	.05* (.016)	2.1* (.03)	.04* (.015)	1.1* (.03)	.04* (.014)	0.8* (.03)
ALFA	-.18* (.019)	3.6* (.05)	-.01 (.017)	2.7* (.04)	-.01 (.016)	2.3* (.04)
SECU	-.14* (.018)	3.1* (.04)	.001 (.016)	2.3* (.04)	.006 (.016)	1.9* (.03)
SOBI	-.01 (.016)	2.2* (.03)	.014 (.015)	1.3* (.03)	.01 (.015)	1.1* (.03)
L2 models	1-lag (B1)		2-lag (B2)		3-lag (B3)	
	$b_0\,(se)$	$b_1\,(se)$	$b_0\,(se)$	$b_1\,(se)$	$b_0\,(se)$	$b_1\,(se)$
HM	-.07* (.016)	4.1* (.07)	.02 (.015)	3.2* (.06)	.02 (.015)	2.7* (.06)
SHB	-.14* (.017)	4.8* (.08)	-.03 (.015)	3.1* (.06)	-.02 (.015)	2.9* (.06)
INVE	-.04* (.016)	4.2* (.07)	.02 (.016)	3.4* (.06)	.03 (.015)	2.9* (.06)
SWMA	-.07* (.016)	3.9* (.06)	-.01 (.015)	3.0* (.06)	-.004 (.015)	2.5* (.06)
HEXA	.05* (.015)	1.5* (.04)	.05* (.014)	.87* (.04)	.04* (.014)	0.7* (.04)
ALFA	-.13* (.017)	5.1* (.08)	-.03 (.016)	4.2* (.07)	-.02 (.015)	3.6* (.07)
SECU	-.12* (.016)	4.2* (.07)	-.02 (.015)	3.4* (.06)	-.02 (.015)	2.8* (.06)
SOBI	-.03 (.015)	2.0* (.04)	-.003 (.015)	1.4* (.04)	-.003 (.014)	1.2* (.04)

Table 4: Coefficient estimates and standard errors for fitted L1 models with 1, 2, and 3 lags (upper pane) and L2 models with 1, 2, and 3 lags (lower pane), for each stock. An asterisk (*) denotes significance at the 99% level. (in-sample obs.: 20160/stock.)

level and have positive signs. Thus, the relationship between the direction of price-change and preceding order imbalance measures is statistically significant. The positive signs mean the relationship is an increasing function in OI.[9] Third, for all stocks and for L1 and L2 models, the magnitudes of slope estimates increase the closer we are (in event-time) to the price-change event. In other words, the same order imbalance value will predict a higher probability of the next tick being up the closer it is to the price-change event. Fourth, we note that except for HEXA and SOBI, slope estimates have higher magnitudes in all L2 models than in their respective L1 models, meaning that L2 models are, in general, more likely than L1 models to predict the next tick as an uptick given the same order imbalance value. In addition, HEXA and SOBI have considerably smaller slope magnitudes when compared to those of the other six stocks, across levels and lags.

[9] In interpreting the estimates, it is important to remember that for a logistic model, marginal effects, or the impact of a change in an explanatory variable on the probability of a positive outcome is not fixed, but rather depends on the value of the explanatory variable (where you measure the marginal effect). (Hosmer Jr et al. (2013) and Wooldridge (2015).)

4.3 Goodness-of-Fit and Predictive Power

To assess goodness-of-fit, we find the percent of correctly predicted training-set observations (in-sample performance). For predictive performance, we find the percent of correctly predicted testing-set observations (out-of-sample performance). These percentages are then benchmarked against the performance of a naive estimator, which predicts equally likely positive and negative ticks. Table 5 details the results. We note that for all cases, in-sample performance is higher than that of the naive model. The models are, thus, doing a good job fitting the data. Also, for all cases, out-of-sample and in-sample performance are very close. The maximum difference is 1.23% for INVE Model (A1). This confirms the hypothesis that order book imbalance can reliably predict tick signs. In addition, L1 models are superior to L2 models which may indicate that the information at the top of the book is best for predicting the direction of price change. Furthermore, we note that for L1 and L2 models and for all stocks, performance increases the closer we get (in event-time) to the price-change event. Another way of describing this is by saying that the predictive power decays with each lag, or event. In addition, the decrease in performance is much stronger for L1 models than for L2 ones. This is true for both in-sample and out-of-sample performance. Most interesting, however, is that all models for HEXA and SOBI underperform their respective counterparts for other stocks.

As discussed in Section 3.3, we are also interested in calculating the area under the out-of-sample ROC curve, or AUC, for all models and stocks. We list the AUCs in Table 6 and plot the ROC curves in Figure 9. The ROC curves make it easier to visually compare the performance of different models for each stock. The AUC table confirms the results in Table 5. Note that the AUC for a naive model is 0.5 and, therefore, all our logistic models outperform the naive model. The table also shows that, for all stocks, Model (A1) is by far the best discriminator, out-of-sample. Finally, HEXA and SOBI clearly stand out as the least predictable stocks.

L1 models	1-lag (A1)		2-lag (A2)		3-lag (A3)		Naive	
	in $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$
HM	76.17%	75.52%	72.94%	71.82%	70.79%	69.34%	49.62%	51.03%
SHB	72.31%	72.40%	69.54%	68.46%	68.37%	67.55%	50.05%	49.96%
INVE	78.79%	80.02%	75.39%	76.28%	72.83%	73.44%	50.37%	50.40%
SWMA	77.18%	77.98%	72.75%	72.70%	69.92%	68.60%	49.98%	49.25%
HEXA	70.80%	70.88%	61.26%	60.52%	58.24%	57.23%	49.74%	49.26%
ALFA	75.56%	75.17%	73.70%	73.47%	71.64%	71.49%	49.91%	49.14%
SECU	72.97%	73.90%	71.75%	71.09%	69.06%	68.68%	49.55%	50.97%
SOBI	71.58%	70.52%	65.34%	64.77%	62.54%	62.04%	50.68%	49.54%
L2 models	1-lag (B1)		2-lag (B2)		3-lag (B3)		Naive	
	in $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$	*in* $\tilde{R}^2$	*out* $\tilde{R}^2$
HM	69.99%	69.66%	67.90%	67.88%	65.70%	65.73%	49.62%	51.03%
SHB	69.17%	69.66%	66.48%	66.16%	65.22%	65.15%	50.05%	49.96%
INVE	71.74%	72.89%	69.50%	70.12%	67.54%	68.16%	50.37%	50.40%
SWMA	70.24%	70.88%	67.54%	68.00%	65.24%	65.04%	49.98%	49.25%
HEXA	61.27%	62.14%	56.78%	56.16%	55.21%	56.13%	49.74%	49.26%
ALFA	71.19%	70.46%	69.68%	69.33%	67.63%	67.13%	49.91%	49.14%
SECU	68.66%	68.02%	67.88%	67.24%	65.28%	65.29%	49.55%	50.97%
SOBI	64.29%	63.36%	60.85%	60.50%	58.92%	58.30%	50.68%	49.54%

Table 5: Percent Correctly Predicted, in-sample (*in* $\tilde{R}^2$) and out-of-sample (*out* $\tilde{R}^2$) for L1 models with 1, 2, and 3 lags (upper pane), L2 models with 1, 2, and 3 lags (lower pane), and a naive model, for each stock. (in-sample obs.: 20160/stock, out-of-sample obs.: 5140/stock.)

Stock	Model (A1) L1 1-lag	Model (A2) L1 2-lag	Model (A3) L1 3-lag	Model (B1) L2 1-lag	Model (B2) L2 2-lag	Model (B3) L2 3-lag
HM	.8602	.8029	.7615	.7742	.7348	.7042
SHB	.8426	.7608	.7482	.7912	.7238	.7082
INVE	.8911	.8284	.7956	.8059	.7647	.7383
SWMA	.8702	.7938	.7448	.7772	.7328	.6973
HEXA	.7720	.6391	.6055	.6650	.5899	.5826
ALFA	.8638	.8169	.7867	.7902	.7548	.7276
SECU	.8413	.7868	.7524	.7620	.7232	.6929
SOBI	.7878	.6915	.6562	.6860	.6440	.6158

Table 6: Out-of-sample AUC values (area under ROC curve) for L1 and L2 models with 1, 2, and 3 lags, for each stock. (5140 obs./stock). Note: AUC for the naive model is 0.5.

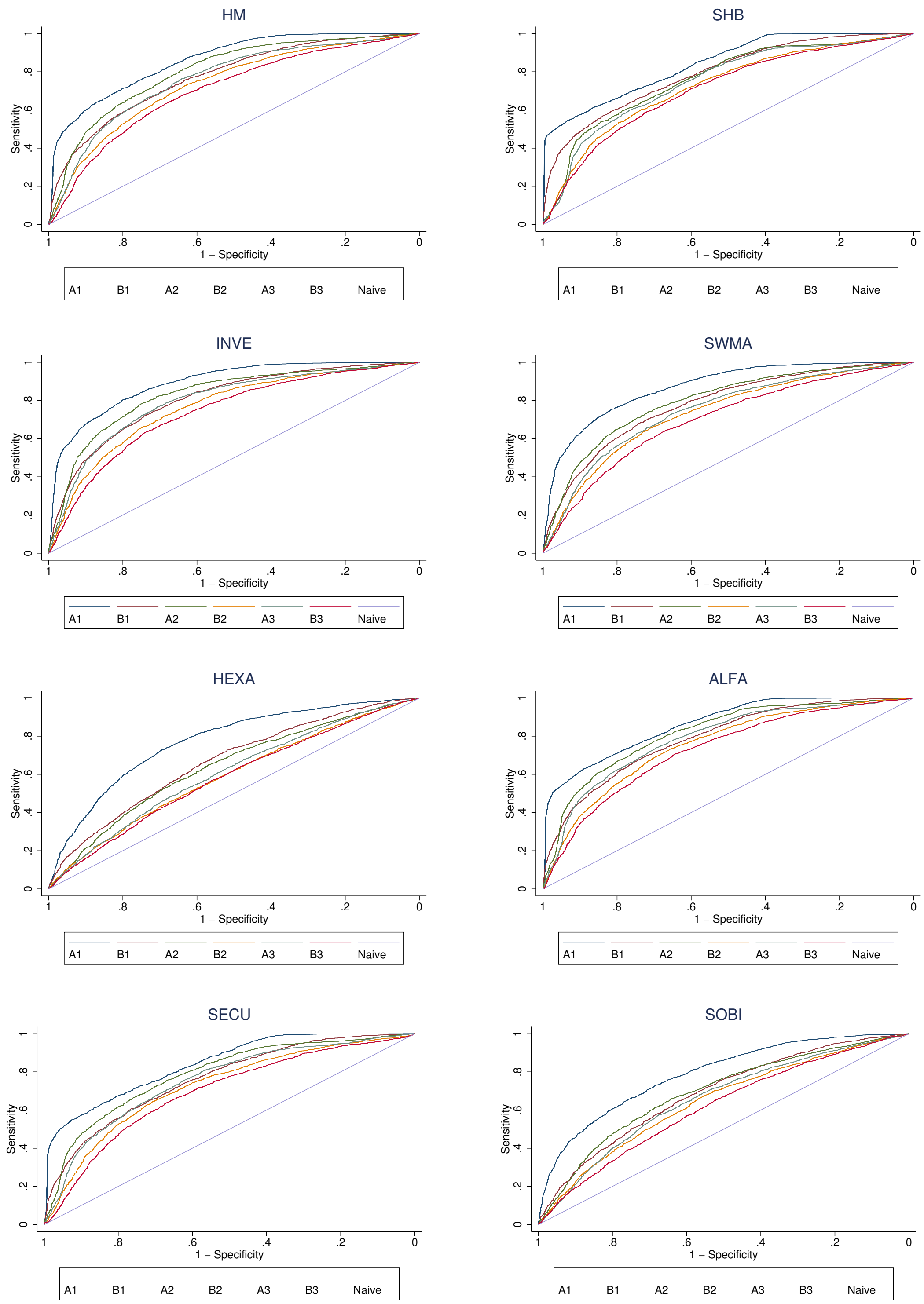


Figure 9: Out-of-sample ROC curves for L1 and L2 models with 1, 2, and 3 lags (Models (A1) to (B3) in Section 3.2) and a naive model, for each stock. (5140 obs./stock).

5 Discussion

We have shown that a simple measure of order imbalance can be used to reliably predict the direction of the next tick. In all the cases in Table 5, the gain in prediction power over a naive predictor is substantial. Even in the worst performing out-of-sample case, HEXA Model (B3), 56% of price changes were predicted correctly. Compared to the 50% benchmark, this gain represents 12% of the maximum possible improvement, which is not negligible in the HFT world. Gould and Bonart (2016) suggest that prediction can be further improved by considering more extreme values of order imbalance. Indeed, when considering only values of OI in the range $[-1, -0.5] \cup [0.5, 1]$, HEXA's Model (B3) out-of-sample performance improves to 62%. While thresholds closer to -1 and 1 produce even better results, we chose -0.5 and 0.5 since imbalance distributions (Figure 7) show local minima at these points for L1 imbalances and a concentration in the range $(-0.5, 0.5)$ for L2 imbalances.

More importantly, we find that L2 order imbalance is not as informative as L1 order imbalance (Section 4.3). The reason may be that since L2 order queues are much larger than L1 queues (Figure 8), L2 information will have a higher probability of containing noise. For example, while information traders are more likely to place orders at the best bid/ask or even step inside the spread, L2 information may be crowded by orders from a broader range of strategies and market participants. The larger size of L2 queues also increases the proportion of 'stale' information at L2 prices. In Section 4.3, we showed that the predictive power of OI quickly decays in successive events.[10]

One of the main findings in Gould and Bonart (2016) is that small-tick and large-tick stocks have very different order imbalance distributions and that small-tick stocks are harder to forecast. Interestingly, we find weak evidence of a strong difference in behavior between the two groups in our results. First, and as discussed in Section 4.1, order imbalance distributions do not show the same patterns found in Gould and Bonart. In the case for Swedish stocks, the main driver of the difference in order imbalance seems to be liquidity, as defined by the average daily trading volume (Table 1). Second, we could not identify a similar difference in prediction performance between the two groups (Section 4.3). In fact, if we only consider Model (A1) and exclude data for SOBI, the evidence points in the opposite direction, i.e., Models (A1) for large-tick stocks outperform the same models for small-tick stocks, although not by a considerable margin.

[10]As a robustness exercise, we repeated our analysis using different sample sizes and definitions of OI, as discussed in Section 1. We also tested different shapes of the relationship between y and OI. The results remained consistent.

Our results also show that HEXA, a small-tick stock, and SOBI, a large-tick stock, can be placed in one group since they are the least predictable stocks (Section 4.3) and their imbalance distributions are strikingly different from the other six stocks (Section 4.1). Comparing stocks in Table 1, we find that HEXA and SOBI share one characteristic: they are not members of the OMXS30 index, while the other six stocks are. It can be argued that index membership is a very strong liquidity indicator, since it invites a larger pool of investors and strategies to trade a stock. The findings in McDermott and Hegde (2000) and Xu (2012) support this argument. In the context of our analysis, we believe this fact supports the idea that liquidity could be the main driver behind the differing order imbalance behavior and price predictability across our sample of stocks. Lower liquidity is reflected in wider spreads and invites hidden or iceberg orders into the order book, which can also make our measure of the order imbalance less informative for non-index-member stocks, in turn obscuring the price discovery process. For this reason, Zheng et al. (2012) control for liquidity in their models while Avellaneda et al. (2011) introduce a model for order imbalance with hidden liquidity.

We stress that comparing our findings to Gould and Bonart's is not a simple apples-to-apples comparison. As discussed in Section 4.1, our methodologies differ slightly and there are differences between US and Swedish markets. With that in mind, we reflect upon the two points provided by Gould and Bonart in their paper as an explanation for the difference in their results between large-tick and small-tick stocks. The first explanation they present is that the large-tick sample has much tighter spreads than the small-tick one. In fact, the mean spread in their large-tick sample is almost equal to the minimum tick size. With no new orders arriving inside the spread, the only way the mid-price would change is by depleting the best bid/ask, which means that the relationship between the size of the best bid/ask and the probability of price change is more direct, and so, the order imbalance should have higher explanatory power. We confirm this idea with a test. We run our analysis for HM when the spread is equal to the minimum tick size (0.1 SEK) and find that the out-of-sample explanatory power of Model (A1) increases from 76% to 92%. Given the fact that our large-tick stocks are not as liquid as their US counterparts, they are less likely to have spreads that are equal to the minimum tick size (Table 2). This could explain why our results do not display the same divergence in the predictability of large-tick and small-tick stocks that we see in Gould and Bonart.

The second point Gould and Bonart make is that even when the spread is wider than the minimum tick size, the 'cost' of placing an order inside the spread is relatively higher for large-tick stocks. Therefore, large-tick investors prefer to wait at the best bid/ask. On the other hand, small-tick investors find it relatively cheaper to place orders inside the spread. Consequently, small-tick stocks will have smaller bid/ask sizes. O'Hara et al. (2018) also find that relative tick size plays a large role in affecting transaction costs and trader behavior. Gould and Bonart argue further that bid/ask sizes that are very small are not informative and, therefore, excluding situations where both the bid and ask sizes are small will improve performance. In other words, small-tick stocks, which are more likely than large-tick ones to have small bid/ask sizes, will have less informative imbalances and, thus, be less predictable. While we do find evidence in our data of small-tick stocks having smaller bid/ask sizes than large-tick stocks do (Figure 8), we do not find evidence of better performance when excluding observations with small bid/ask sizes. For example, we run our analysis for HM using different thresholds of minimum bid/ask sizes and find that out-of-sample performance fluctuates in a narrow range around the numbers for the full subsample. In contrast, when we excluded the largest bid/ask sizes, the out-of-sample performance improves. Regardless of these observations, we believe that the 'relative cost' of stepping inside the spread is less relevant for stocks trading on Nasdaq Sweden because the dynamic tick size regime reduces liquidity-related issues that result from having the same tick size for all stocks. For further discussion on the relationship between the minimum tick size, liquidity, and efficiency, see Brown and Yang (2016).

6 Conclusions

We examined order imbalance, a simple measure that summarizes supply and demand status in a limit order book, and presented empirical evidence of its ability to predict the direction of price changes in an ultra-high-frequency environment. We also showed that order imbalance has more predictive power at the top of the book than at the second level in the book. In addition, we presented theoretical and empirical evidence of the role of liquidity in driving price-direction predictability in the context of our analysis. Finally, we compared our findings to those in Gould and Bonart (2016). Unlike the case for US stocks, we found no difference in order imbalance characteristics or price predictability between large-tick and small-tick stocks. We argued that this is due to the dynamic tick size regime adopted in European markets and the wider bid-ask spreads of Swedish stocks.

Our results have practical implications for HFT investors. For example, an algorithm using order imbalance as an indicator can optimize computing resources by considering Level

I information only. However, we point to some caveats and areas of improvement for this study. First, we investigated the predictability of price direction given a price-change event. A future study can investigate using LOB information to predict the occurrence of a price-change event, which would be useful for an order imbalance-based trading strategy. For example, Mizrach (2006) finds that the number of bids (offers) is more informative of the probability of an uptick (downtick) than the quoted depth. Second, our models can be further improved by controlling for liquidity, as in Zheng et al. (2012). Third, we only study eight stocks and, therefore, we are cautious about generalizing our results to the larger pool of instruments. A future study may look at a larger number of stocks and at other instruments, perhaps more liquid, such as index futures. Fourth, this study considered the first two levels in the order book and only three events before a price change occurs, it would be interesting to find out how predictability changes if we consider more events and deeper levels in the book. Finally, we proposed index membership as a proxy for liquidity and an important factor in determining price predictability. HEXA's addition to OMXS30 in July 2018 is an opportunity to examine this claim. A follow up study can compare HEXA's price predictability before and after its inclusion in the index.

References

Aldridge, Irene, 2013, High-frequency trading: a practical guide to algorithmic strategies and trading systems, volume 604. (John Wiley & Sons).

Avellaneda, Marco, Reed, Josh, and Stoikov, Sasha, 2011, Forecasting prices from level-i quotes in the presence of hidden liquidity, Algorithmic Finance, 1(1):35–43.

Biais, Bruno and Foucault, Thierry, 2014, Hft and market quality, Bankers, Markets & Investors, 128(1):5–19.

Brogaard, Jonathan, Hendershott, Terrence, and Riordan, Ryan, 2014, High-frequency trading and price discovery, The Review of Financial Studies, 27(8):2267–2306.

Brogaard, Jonathan, Hendershott, Terrence, and Riordan, Ryan, 2017, Price discovery without trading: Evidence from limit orders, Finance Down Under 2017: Building on the Best from the Cellars of Finance, Melbourne.

Brown, A and Yang, F, 2016, Market liquidity and price efficiency: Evidence from two field experiments, working paper, (University of East Anglia).

Cao, Charles, Hansch, Oliver, and Wang Beardsley, Xiaoxin, 2004, The informational content of an open limit order book, EFA 2004 Maastricht Meetings Paper No. 4311.

Chan, Ernest P, 2017, Machine Trading: Deploying Computer Algorithms to Conquer the Markets. (John Wiley & Sons).

Cont, Rama, 2011, Statistical modeling of high-frequency financial data, IEEE Signal Processing Magazine, 28(5):16–25.

Cont, Rama, Kukanov, Arseniy, and Stoikov, Sasha, 2014, The price impact of order book events, Journal of financial econometrics, 12(1):47–88.

Easley, David, López de Prado, Marcos M, and O'Hara, Maureen, 2012, Flow toxicity and liquidity in a high-frequency world, The Review of Financial Studies, 25(5):1457–1493.

Goldstein, Michael A, Kwan, Amy, and Philip, Richard, 2017, High-frequency trading strategies, working paper.

Gould, Martin D and Bonart, Julius, 2016, Queue imbalance as a one-tick-ahead price predictor in a limit order book, Market Microstructure and Liquidity, 2(02):1650006.

Hagströmer, Björn and Nordén, Lars, 2013, The diversity of high-frequency traders, Journal of Financial Markets, 16(4):741–770.

Hosmer Jr, David W, Lemeshow, Stanley, and Sturdivant, Rodney X, 2013, Applied logistic regression, volume 398. (John Wiley & Sons).

Khashanah, Khaldoun, Florescu, Ionut, and Yang, Steve, 2014, High-frequency trading: a white paper, IRRC Institute, September.

Lipton, Alexander, Pesavento, Umberto, and Sotiropoulos, Michael G, 2013, Trade arrival dynamics and quote imbalance in a limit order book, arXiv preprint arXiv:1312.0514.

McDermott, John B. and Hegde, Shantaram P., 2000, The liquidity effects of additions to the s&p500 index, Working paper, (University of Connecticut).

Menkveld, Albert J, 2013, High frequency trading and the new market makers, Journal of Financial Markets, 16(4):712–740.

Mizrach, Bruce, 2006, The next tick on nasdaq: Does level ii information matter?, Working paper, (Rutgers University).

O'Hara, Maureen, Saar, Gideon, and Zhong, Zhuo, 2018, Relative tick size and the trading environment, Review of Asset Pricing Studies, Forthcoming.

Verousis, Thanos, Perotti, Pietro, and Sermpinis, Georgios, 2018, One size fits all? high frequency trading, tick size changes and the implications for exchanges: market quality and market structure considerations, Review of Quantitative Finance and Accounting, 50 (2):353–392.

Wooldridge, Jeffrey M, 2015, Introductory econometrics: A modern approach. (Nelson Education).

Xu, Yuanbin, 2012, Impact of changes in the nasdaq 100 index membership, working paper, (Brock University).

Zheng, Ban, Moulines, Eric, and Abergel, Frédéric, 2012, Price jump prediction in limit order book, arXiv preprint arXiv:1204.1381.

Appendix

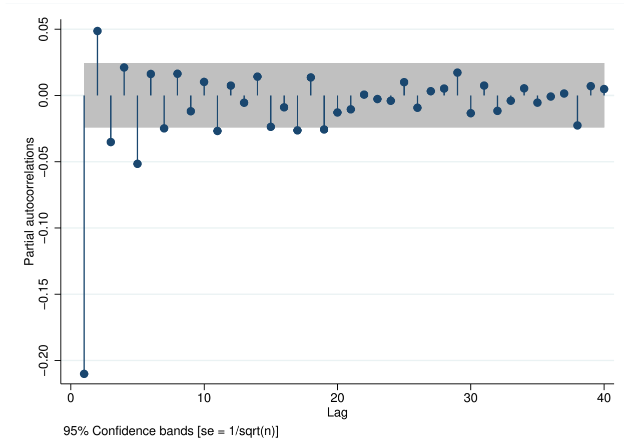


Figure 10: Partial autocorrelations of tick signs for HM shares on Jan 4, 2016. (6503 ticks.)

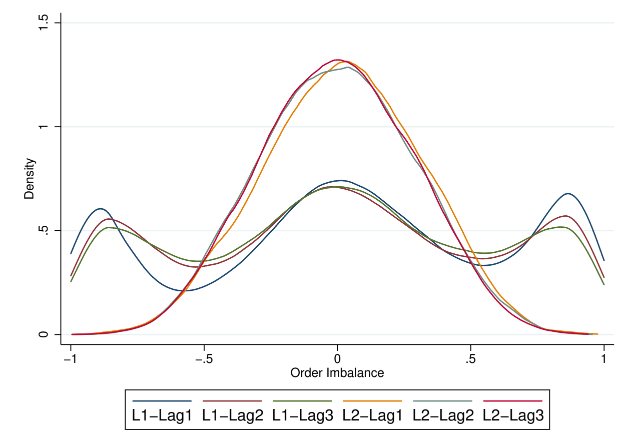


Figure 11: Frequency distributions of L1 and L2 order imbalances for HM in the three events prior to a price change. (100 points/day, total 25300 points for 2016.)

symbol	isin	date	time	askpx1	asksz1	bidpx1	bidsz1	askpx2	asksz2	bidpx2	bidsz2	txpx	txsz	status	market	currency	instrumentid
HM B	SE0000106270	2016-01-04	13:16:49.275785398	288.8000	2555	288.7000	3044	288.9000	4418	288.6000	872	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.275832470	288.8000	1960	288.7000	3044	288.9000	4418	288.6000	1272	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.275855949	288.8000	2360	288.7000	3044	288.9000	3730	288.6000	1867	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.276497711	288.8000	2039	288.7000	3044	288.9000	3730	288.6000	1867	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.276843052	288.8000	2039	288.7000	2764	288.9000	3730	288.6000	2188	288.7000	280	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.281800537	288.8000	2319	288.7000	2764	288.9000	3730	288.6000	2188	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927392705	288.8000	2319	288.7000	2216	288.9000	3730	288.6000	2188	288.7000	548	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927447412	288.8000	2319	288.6000	2188	288.9000	3730	288.5000	5397	288.7000	2216	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927458915	288.8000	2555	288.6000	2188	288.9000	3730	288.5000	5397	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927501133	288.8000	2555	288.6000	2528	288.9000	3730	288.5000	5397	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927514876	288.8000	2555	288.6000	2864	288.9000	3730	288.5000	5397	N/A	N/A	T	XSTO	SEK	126
HM B	SE0000106270	2016-01-04	13:16:49.927531413	288.8000	2555	288.6000	2543	288.9000	3730	288.5000	5397	N/A	N/A	T	XSTO	SEK	126

Figure 12: A sample of 12 consecutive events from the HM limit order book on Jan 4, 2016. The reconstructed Level II order books are obtained from the Nasdaq HFT database at the Swedish House of Finance Date Center. The order books are reconstructed from historic Nasdaq ITCH feeds and contain nanosecond-stamped views of the top two levels of the order book at each event. For each event, the raw data contains the instrument's symbol on the exchange, the instruments ISIN, date, nanosecond time stamp, L1 ask price (askpx1) and size (asksz1), L1 bid price (bidpx1) and size (bidsz1), L2 ask price (askpx2) and size (asksz2), L2 bid price (bidpx2) and size (bidsz2), transaction price (txpx) and size (txsz), if any, status of the instrument (T=Continuous Trading), the Market Identifier Code (Market), ISO code of the currency, and a unique identifier of the instrument at the Swedish House of Finance (SHoF).
Source: Swedish House of Finance Research Data Center. For more information, refer to the latest Nasdaq TotalView-ITCH protocol interface specification.