Learning-Based Trading Strategies in the Face of Market Manipulation

Xintong Wang [1] **Chris Hoang** [1] **Michael P. Wellman** [1]

Abstract

We study learning-based trading strategies in markets where prices can be manipulated through *spoofing*: the practice of submitting spurious orders to mislead traders who use market information. We explore two variations based on the *heuristic belief learning* (HBL) trading strategy (Gjerstad, 2007), which learns transaction probabilities from observed order activity. The first variation selectively ignores orders at certain price levels, particularly where spoof orders are likely to be placed. The second considers the full order book, but adjusts its offer price to correct for bias in decisions based on the learned heuristic beliefs. To evaluate the two proposed variations, we employ agent-based simulation for several market settings where background traders can adopt (non-learning) zero intelligence strategies or HBL, in basic form or the two variations. We demonstrate through empirical game-theoretic analysis that both variations can reduce the learning agents' vulnerability to spoof orders in equilibrium, and thereby increase the overall background-trader surplus in the market.

1. Introduction

The increasing automation of trading and the unprecedented interconnectedness of trading venues have transformed the financial market from a human decision ecosystem to an algorithmic one. With trades happening in an extremely short time scale, often beyond the limit of human decision making, traders develop and use autonomous agents controlled by complex algorithms to extract, process, and react to new information. While this may reduce transaction costs and improve efficiency in some respects, the evolution of market operations and trading technology has also made new forms of disruptive and manipulative practices possible.

In this paper, we focus on a common form of order-based market manipulation, called *spoofing*. Spoofing refers to the practice of submitting large spurious orders to buy or sell some security. The orders are not intended for execution, but rather to mislead other traders by feigning a strong buy or sell interest in the market. Spoof orders may persuade other traders—those who learn from market information—to believe that prices may soon rise or fall, thus altering their own behavior in a way that moves the price. To profit on its feint, the spoofer follows by submitting a real order on the opposite side of the market and as soon as the real order transacts, cancels all the spoof orders.

In normal markets, there is real information to be gleaned from the order book, and thus strategies that *learn* from observable market activity have an advantage over those that neglect such information. The less sophisticated strategies, however, have the advantage of being oblivious to spoofers, and thus not manipulable. The question we investigate is whether learning-based strategies can be made similarly robust to spoofing, or more broadly how to design strategies that reasonably trade off effectiveness in non-manipulated markets for robustness against manipulation. Specifically, we are interested in exploring strategies by which individual traders can adopt to learn from market information but in less vulnerable ways.

We start from the *heuristic belief learning* (HBL) trading strategy (Gjerstad, 2007), previously adopted in the agent-based model of spoofing by Wang & Wellman (2017). We explore two variations of the original HBL strategy. In the first, we selectively ignore orders at certain price levels, particularly where spoof orders are likely to be placed. This approach is inspired by the *cloaking* mechanism of Wang et al. (2018), in which the the market itself hides certain price levels. Here, the traders themselves decide what prices to ignore. In a second variation, HBL agents consider the full order book, but adjust their offer price by a constant offset. The adjustment serves to correct biases in learned price beliefs either caused by manipulation or the intrinsic limitation built in the belief function.

We employ agent-based simulation to evaluate the two proposed variations on HBL under equilibrium settings across a variety of parametrically distinct market environments with and without manipulation. The market is populated with multiple background trading agents and one exploiter,

[1]University of Michigan, Ann Arbor. Correspondence to: Xintong Wang <xintongw@umich.edu>.

ICML-19 Workshop on AI in Finance, Long Beach, CA, 2019.

trading a single security. The exploiter profits by first buying the underlying security at low prices and later selling at higher ones. To increase profit, it can try to manipulate the market through spoofing after its original purchase. Background traders with private values are further divided into non-spoofable agents employing the *zero intelligence* (ZI) strategy, and learning agents who employ HBL, either in its basic form or the two variations introduced here. We demonstrate that both variations can reduce the learning agents' vulnerability to spoof orders, and thus increase the overall background-trader surplus in the market. Empirical game-theoretic analysis implied that trading agents can strategically adapt to the presence of manipulation, even without any intervention to regulate the market.

The paper is structured as follows. In the next section we present additional background on market manipulation. In Section 3, we describe the market model and the two variations of HBL trading strategy in detail. Section 4 presents experiments and our main findings. Section 5 concludes with discussions.

2. Background on Market Manipulation

The US Securities and Exchange Commission (SEC) formally defines manipulation as "intentional conduct designed to deceive investors by artificially affecting the market." With the automation of trading, manipulators nowadays employ automated programs to spread deceitful information through rapid submission and cancellation of voluminous orders, as other investors learn from market information (including the misleading orders) to make trading decisions. Manipulators in turn profit from investors' misled beliefs about supply and demand.

Since 2016, the SEC has brought legal action on over a hundred cases of manipulation (U.S. Securities and Exchange Commission, 2017; 2018b). In one recent case, a group of individuals and ten associated entities was charged with participation in fraudulent schemes that generated over $27 million from unlawful stock sales (U.S. Securities and Exchange Commission, 2018a). The manipulation was conducted through a series of illegal promotional activities in a short period of time to artificially boost each issuer's stock price, giving an appearance of active trading volume at certain prices.

In 2010, the Dodd-Frank Wall Street Reform and Consumer Protection Act was signed into federal law, outlawing spoofing as a deceptive practice. Given the challenges of directly detecting disruptive practices, regulators advocate simplifying investment strategies, improving intermediary integrity, and enhancing financial cybersecurity (Lin, 2015).

Some researchers in computational finance recommend a systematic imposition of cancellation fees, as this may render manipulative strategies that involve massive cancellations uneconomical (Biais & Woolley, 2012; Prewit, 2012). Opponents argue that cancellation fees could instead discourage the activity of liquidity providers to minimize risk from adverse selection and adapt to new information (Copeland & Galai, 1983; Foucault et al., 2003). Wang & Wellman (2017) built a computational model of spoofing, illustrating the strategic interactions between a manipulator and two groups of background traders: those who use and do not use market information to trade. In their model, the designed spoofing strategy is profitable and can effectively mislead other rational traders. A comparison of equilibrium outcomes shows that manipulation hurts market welfare and decreases the number of learning traders. In a follow-up work, Wang et al. (2018) proposed to deter spoofing through strategically cloaking certain market information, introducing risks and difficulties for the manipulator to post misleading bids. Simulation results derived from equilibrium analysis demonstrate the designed cloaking mechanism can significantly diminish the efficacy of spoofing, but at the cost of degrading the general usefulness of market information.

3. Market Model

3.1. Market Mechanism

We build on an existing computational model of spoofing (Wang & Wellman, 2017) to conduct experiments. Market prices take on discrete integer values with a tick size of one. Time is also discrete over a finite trading horizon T. The fundamental value r_t of the traded security changes throughout the trading period according to a mean-reverting stochastic process (Chakraborty & Kearns, 2011; Wah et al., 2017): for $t \in [0, T]$,

$$r_t = \max\{0, \kappa\bar{r} + (1-\kappa)r_{t-1} + u_t\} \text{ and } r_0 = \bar{r}, \quad (1)$$

where $\kappa \in [0, 1]$ specifies the degree to which the time series reverts back to the fundamental mean $\bar{r}$. $u_t \sim N(0, \sigma_s^2)$ represents a systematic random shock upon the fundamental at time t, where σ_s^2 is the fundamental shock variance.

Agents trade a single security by submitting limit orders that specify the maximum (minimum) price at which they would be willing to buy (sell) some number of units. The market maintains a *limit order book* of outstanding orders, from which traders may learn at their own discretion.

3.2. Agents in the Market

The market is populated with multiple background traders and one exploiter. Background traders represent investors with private values which specify preferences on longing or shorting the underlying security. The exploiter without any private value seeks to profit by buying at lower prices and later selling at higher ones. In order to profit more, it can manipulate the market with spoof orders to push prices up.

The position preference of background trader i is captured by a private value vector Θ_i of length $2q_{\max}$, where $q_{\max}$ is the maximum number of units one can be long or short at any time. Element θ_i^{q+1} represents the marginal gain from buying an additional unit, given the current net position q. We generate Θ_i from a set of $2q_{\max}$ values independently drawn from $N(0, \sigma_{PV}^2)$, where σ_{PV}^2 denotes the private value variance. We then sort elements to reflect diminishing marginal utility and assign θ_i^q accordingly. The trader's overall *valuation* for a unit of the security is the sum of its private value and the fundamental.

Arrivals of a background trader to the market follow a Poisson process with an arrival rate of λ_{a}. On each entry, the trader observes an agent and time specific noisy fundamental $o_t = r_t + n_t$ with the observation noise following $n_t \sim N(0, \sigma_n^2)$, where σ_n^2 represents the observation variance. The noisy observation captures different perceptions of the intrinsic value of the underlying security. As this noisy observation gives imperfect information about the fundamental, traders can benefit from considering market information, which is influenced by the aggregate observations of all the other traders. To react to a new observation, the background trader withdraws its previous order (if not transacted) upon arrival and submits a new single-unit order to either buy or sell as instructed with equal probability. The order price is jointly decided by the background trader's valuation and *trading strategy* (see Section 3.3).

3.3. Background Trading Strategies

3.3.1. Estimation of the Final Fundamental

Because holdings of the security are evaluated at the end of the trading period, background traders estimate the final fundamental value based on their noisy observations. Given a new noisy observation o_t, a trader estimates the current fundamental by updating its posterior mean $\tilde{r}_t$ and variance $\tilde{\sigma}_t^2$ in a Bayesian manner. Let t' denote the trader's preceding arrival time. We first update the previous posteriors ($\tilde{r}_{t'}$ and $\tilde{\sigma}_{t'}^2$) by taking account of mean reversion for the interval since preceding arrival ($\delta = t - t'$):

$$\begin{aligned} \tilde{r}_{t'} &\leftarrow (1-(1-\kappa)^{\delta})\bar{r} + (1-\kappa)^{\delta}\tilde{r}_{t'} \ ; \\ \tilde{\sigma}_{t'}^2 &\leftarrow (1-\kappa)^{2\delta}\tilde{\sigma}_{t'}^2 + \frac{1-(1-\kappa)^{2\delta}}{1-(1-\kappa)^2}\sigma_s^2. \end{aligned}$$

The new posterior estimates at time t are then given by:

$$\tilde{r}_t = \frac{\sigma_n^2}{\sigma_n^2 + \tilde{\sigma}_{t'}^2}\tilde{r}_{t'} + \frac{\tilde{\sigma}_{t'}^2}{\sigma_n^2 + \tilde{\sigma}_{t'}^2}o_t \ ; \ \tilde{\sigma}_t^2 = \frac{\sigma_n^2\tilde{\sigma}_{t'}^2}{\sigma_n^2 + \tilde{\sigma}_{t'}^2}.$$

Based on the posterior estimate of $\tilde{r}_t$, the trader computes $\hat{r}_t$, its estimate at time t of the terminal fundamental r_T, by adjusting for mean reversion:

$$\hat{r}_t = \left(1-(1-\kappa)^{T-t}\right)\bar{r} + (1-\kappa)^{T-t}\tilde{r}_t. \qquad (2)$$

3.3.2. Zero Intelligence

We employ an extended and parameterized version of *zero intelligence* (ZI) as a representative strategy that is non-spoofable, as it decides order prices without utilizing order book information. The strategy has been widely adopted in agent-based finance due to its simplicity and effectiveness for market modeling (Gode & Sunder, 1993; Farmer et al., 2005).

The ZI trader computes a limit-order price by shading its valuation with a random offset uniformly drawn from $[R_1, R_2]$. Specifically, a ZI trader i arriving at time t with position q generates a limit price by:

$$p_i(t) \sim \begin{cases} U[\hat{r}_t + \theta_i^{q+1} - R_2, \hat{r}_t + \theta_i^{q+1} - R_1] & \text{if buying,} \\ U[\hat{r}_t - \theta_i^{q} + R_1, \hat{r}_t - \theta_i^{q} + R_2] & \text{if selling.} \end{cases}$$

The ZI further takes into account the current quoted price, controlled by a strategic surplus threshold parameter $\eta \in [0, 1]$. Before submitting a new limit order, if the ZI could achieve a fraction η of its requested surplus by accepting the most competitive order, it would take that quote by submitting an order at the same price.

3.3.3. Variations of Heuristic Belief Learning

We describe in detail the two proposed variations of *heuristic belief learning* (HBL). The two variations, together with its basic form, are our representative trading strategies that learn from the market's aggregated order book information. We are interested in investigating their competitiveness to other trading strategies and robustness to manipulation across different market settings.

We first provide a brief description on how the basic HBL works. The HBL trading strategy is centered on belief functions that traders form on the basis of observed market data $\mathcal{D}$ in memory. Upon arrivals, agents estimate the probability that orders at various prices would be accepted in the market according to the heuristic [1]:

$$f_t(P|\mathcal{D}) = \begin{cases} \frac{\text{TBL}_t(\text{P})+\text{AL}_t(\text{P})}{\text{TBL}_t(\text{P})+\text{AL}_t(\text{P})+\text{RBG}_t(\text{P})} & \text{if buying,} \\ \\ \frac{\text{TAG}_t(\text{P})+\text{BG}_t(\text{P})}{\text{TAG}_t(\text{P})+\text{BG}_t(\text{P})+\text{RAL}_t(\text{P})} & \text{if selling.} \end{cases} \qquad (3)$$

Based on the interpolated probabilities, an agent chooses a limit price that maximizes its own expected surplus at the current valuation estimate. That is,

$$\text{P}_i^*(t) = \begin{cases} \arg\max_p (\hat{r}_t + \theta_i^{q+1} - p) f_t(p) & \text{if buying,} \\ \arg\max_p (p - \theta_i^{q} - \hat{r}_t) f_t(p) & \text{if selling.} \end{cases}$$

[1] It uses the observed frequencies of *transacted* and *rejected* orders (T and R), *bids* and *asks* (B and A), and orders with prices *less* than or equal to and *greater* than or equal to P (L and G) within the HBL's memory. For example, $\text{TBL}_t(\text{P})$ is the number of transacted bids found in memory with price less than or equal to P up to time t.

HBL with Price Level Blocking. Our first variation of HBL allows agents to strategically construct $\mathcal{D}$ by neglecting limit orders at a certain price level. Specifically, the strategy extends basic HBL with a blocking parameter L, which specifies the index of a price level to ignore symmetrically from inside of the limit order book. For example, when $L = 1$, the trading agent constructs the dataset $\mathcal{D}$ by considering any order but the best bid and asks, whereas when $L = 0$, the agent acts as a basic HBL. Trading agents can use this flexibility to incorporate information that is more likely to give true insight of the fundamental value. Therefore, in markets with manipulation, agents may strategically exclude spoof orders appearing at a certain price level from its belief function.

HBL with Price Offsets. The second variation of HBL considers all market information, and simply extends the optimal price $\mathrm{P}_i^*(t)$ derived from surplus maximization with a random offset uniformly drawn from $[R_1, R_2]$. Specifically, a trader i arriving at time t with the calculated price $\mathrm{P}_i^*(t)$ submits a limit order for a single unit of the security at price

$$p_i(t) \sim \begin{cases} U[\mathrm{P}_i^*(t) - R_1, \mathrm{P}_i^*(t) - R_2] & \text{if buying,} \\ U[\mathrm{P}_i^*(t) + R_1, \mathrm{P}_i^*(t) + R_2] & \text{if selling.} \end{cases}$$

We explore whether such price offsets can help to overcome limitations brought by either the designed heuristic belief function or any market manipulation.

3.4. Exploitation and Spoofing Strategies

The strategy that an exploiter adopts has three stages. At the beginning of a trading period $[0, T_{\text{spoof}}]$, the exploiter buys as many units as possible by accepting any sell order at prices lower than the fundamental mean $\bar{r}$.

During the second stage $[T_{\text{spoof}}, T_{\text{sell}}]$, if the exploiter manipulates, it submits and maintains spoof buy orders at one tick behind the best bid with volume $Q_{\text{sp}} \gg 1$. This spoofing strategy aims to artificially boost prices, so that units purchased earlier in the first stage can be later sold at higher prices. If the exploiter does not spoof, it simply waits until the selling stage.

During the last stage $[T_{\text{sell}}, T]$, the exploiter begins to sell by accepting any buy orders at a price higher than $\bar{r}$. The exploiter, if also manipulates, continues to spoof until the end of the trading period T or when all units previous purchased are sold.

3.5. Surplus

A background trader's surplus is its net profits from trading plus the *final valuation* of holdings at T, whereas a exploiter's payoff is its gain or loss from trading. The market's final valuation of background trader i with final holdings H is $r_T H + \sum_{k=1}^{k=H} \theta_i^k$ for long position $H > 0$, or alternatively, $r_T H - \sum_{k=H+1}^{k=0} \theta_i^k$ for short position $H < 0$.

4. Empirical Game-Theoretic Analysis

We conduct agent-based simulation of the market model described in Section 3 to evaluate the two HBL variations over a range of market environments with and without spoofing. A *game* is defined by a specific market environment and a strategy set from which each background trader can choose. For each game, we run thousands of simulations to evaluate any *profile* of strategy assignment, calculate the average payoff of agents adopting the same strategy, and derive an approximate Nash Equilibrium. We are interested in measuring a trading strategy and the market performance in equilibrium, where agents have no incentive to deviate to other strategies, given an environment and other players' choices.

Based on fixed strategy profiles, we also perform controlled experiments to understand the impact of manipulation and the performance of different learning strategies. Specifically, in paired instances, we control all other stochastic factors (e.g., agent arrivals, fundamental evolution, and private values), such that any change in agent behavior is caused by the experimental factor of interest (e.g., the presence of manipulation and the blocking of a certain price level by background traders).

The section is structured as follows. Section 4.1 specifies a set of parametrically defined market settings. Section 4.2 summarizes the empirical game-theoretic analysis (EGTA) methodology we adopted to identify equilibrium solutions. In Sections 4.3 and 4.4, we present the EGTA studies and results on effectiveness and robustness of the two HBL variations.

4.1. Market Environment

The global fundamental time series is generated according to Eq.(1) with a fundamental mean $\bar{r} = 10^5$ and a mean reversion constant $\kappa = 0.05$. Each trading period lasts $T = 10,000$ discrete time steps. We consider three representative environments listed in Table 1 that vary in market shock σ_s^2 and observation noise σ_n^2 to cover different market conditions. Specifically, **LSHN** represents a market with low shock and high observation noise, **MSMN** a market with medium shock and medium observation noise, and **HSLN** a market with high shock and low observation noise. Intuitively, market shock controls the intensity of fluctuations in the fundamental time series and consequently, influences the predictability of future price outcomes. Observation noise limits the extent to which agents can rely on their own information, and thus encourages the trading agents to learn from the market's aggregated order book information.

Environment	Shock σ_s^2	Observation Noise σ_n^2
LSHN	10^5	10^9
MSMN	5×10^5	10^6
HSLN	10^6	10^3

Table 1. Market environments considered in the experiments.

Our market is populated with 64 background traders and one exploiter. Background traders arrive at the market according to a Poisson distribution with rate $\lambda_a = 0.005$ and observe a noisy fundamental o_t. Private values are drawn from a Gaussian distribution with zero mean and a variance of $\sigma_{PV}^2 = 5 \times 10^6$. The maximum number of units that they can hold at any time is $q_{\text{max}} = 10$. The background traders can choose from a restricted strategy set defined in Table 2 to maximize their payoffs.

The exploiter follows the strategy described in Section 3.4. If it manipulates, the exploiter submits spoof orders with volume $Q_{\text{sp}} = 200$ at time $T_{\text{spoof}} = 1000$. Since then, the exploiter maintains its spoof orders at a tick behind the market's best bid to push prices up, until all the earlier bought units are sold. At time $T_{\text{sell}} = 2000$, it starts to liquidate its position by selling units at prices above $\bar{r}$.

4.2. Empirical Game-Theoretic Analysis

We provide a brief overview of empirical game-theoretic analysis (EGTA), a methodology for performing strategy selection to find equilibria in games defined by discretized strategy space and simulated payoff data. We refer to Wellman (2016) for detailed descriptions.

EGTA takes an iterative process to identify candidate equilibria in subgames (games over strategy subsets), and searches for potential deviations until a candidate is confirmed. Explorations start with subgames where all agents play a single strategy, and incrementally spread to other strategies. Equilibria from a subgame are considered as candidate solutions of the full game, and will be refuted if there exists a beneficial deviation to a strategy outside the subgame set. If we examine all deviations without refuting, the candidate is confirmed. We continue to refine the empirical subgame with additional strategies and corresponding simulations until at least one equilibrium is confirmed and all non-confirmed candidates are refuted.

We model the market as a role-symmetric game, which is defined by an environment and agents representing two roles: background traders and a single exploiter. Since game size can grow exponentially in the number of players and strategies, we apply the *deviation-preserving reduction* (DPR) (Wiedenbeck & Wellman, 2012) technique to approximate large games with many agents as games with fewer players. We obtain this approximation through aggregation, which

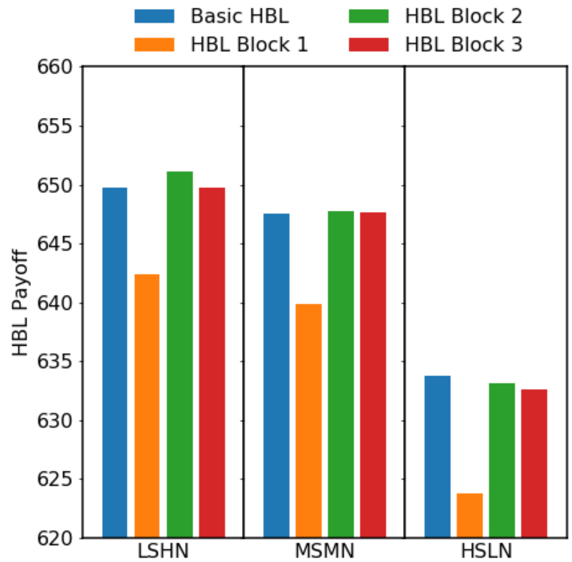


Figure 1. Comparisons of payoffs achieved by basic HBLs and its variations that block orders from the first, second, and third price level respectively across market settings.

preserves payoffs from single-player deviations. To facilitate DPR, we choose values to ensure that the required aggregations come out as integers. For example, in this study, we choose 64 background traders and one exploiter, so that a market can be aggregated to a smaller one with four background traders and one exploiter; as one background trader deviates to a new strategy, the remaining 63 can be further reduced to three traders.

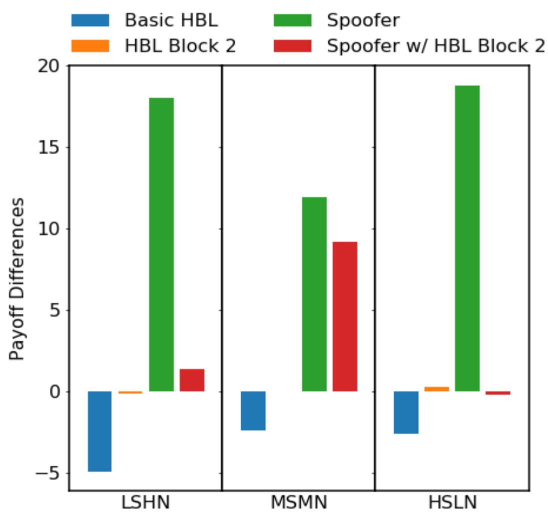
4.3. HBL with Price Level Blocking

A learning trader who chooses to ignore certain orders faces a natural trade-off between losing useful information and correctly blocking spoof orders, thus being robust to manipulation. We first examine, under *non-spoofing* environments, how important are orders at each price level for background agents to make informed trading decisions. We start with one equilibrium profile of each market environment where background traders are restricted to choose from the *basic* HBL strategy and five parametrically different ZI strategies in Table 2. Based on the equilibrium profile, we perform controlled experiments by letting the learning traders ignore orders from a selected price level throughout the trading period. Figure 1 demonstrates the payoffs obtained by basic HBL and HBLs that respectively block orders at the first, second and third price level in the order book. We find that consistently across market settings, HBL agents benefit the most by learning from market best bids and asks, and can achieve fairly similar performance even when orders from the second or third level are ignored.

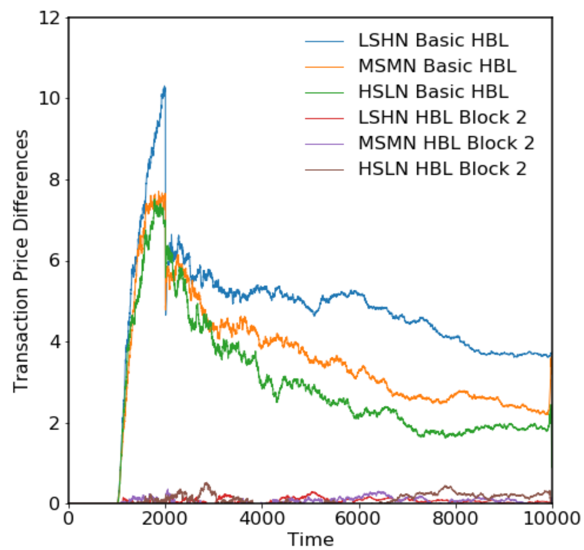
To understand the learning-and-blocking strategy in the face

Strategy	ZI_1	ZI_2	ZI_3	ZI_4	ZI_5	HBL_1	HBL_2	HBL_3	HBL_4	HBL_5	HBL_6
L	-	-	-	-	-	0	1	2	0	0	0
R_1	0	0	0	0	0	0	0	0	-10	-20	-40
R_2	1000	1000	1000	500	250	0	0	0	0	0	0
η	0.4	0.8	1	0.8	0.8	1	1	1	1	1	1

Table 2. Background trading strategies included in empirical game-theoretic analysis.



(a) Differences in HBL payoffs and the exploiter profit in markets with and without spoofing.



(b) Transaction price differences in markets with and without spoofing.

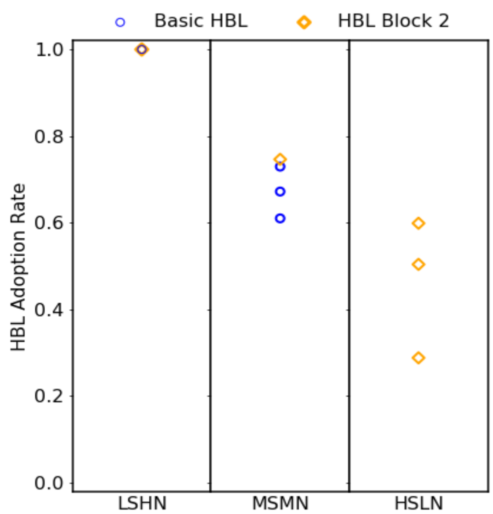
Figure 2. Payoff differences and transaction price rises caused by spoofing across market settings. The group of learning traders chooses either the basic HBL or HBL that blocks the second price level, where spoof orders are placed.

of manipulation, we further let the exploiter inject spoof orders at a tick behind the best bid (i.e., the second price level), and compare the differences in payoffs and transaction prices of the paired two markets with and without spoofing. As what we expect, Figure 2a shows that learning traders who correctly block spoof orders can achieve similar payoffs as in markets without manipulation, due to the robustness to manipulation. However, in markets where traders adopt basic HBL, the exploiter is able to get much higher profits, while the background agents suffer from the spoofed beliefs and obtain lower payoffs. Figure 2b further demonstrates that if spoof orders can be correctly blocked, prices in market will be less affected with price rises caused by manipulation close to zero. However, in markets where spoof orders are not blocked, transaction prices increase subsequently to the arrival of spoof orders at $T = 1000$.

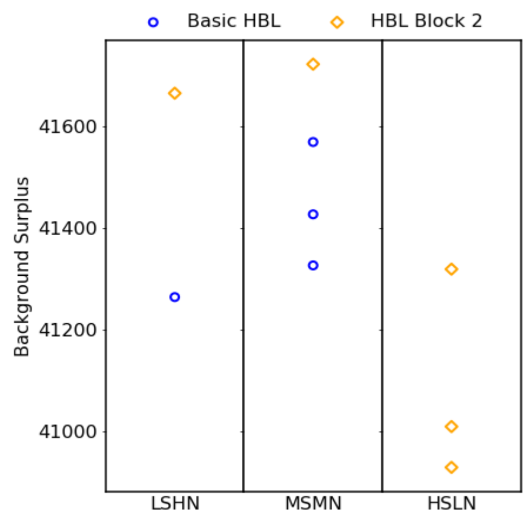
In the final set of experiments, we conduct EGTA to find Nash equilibria in games where the exploiter is consistently manipulating and background traders choose any trading strategies from the ZI family and HBLs that block a selected price level. We find though both HBL with and without order blocking can appear at the equilibrium, markets that populate with learning agents who correctly block spoof orders at the second price level can achieve much larger HBL proportions and consequently, higher total surplus in equilibrium (See Figure 3).

4.4. HBL with Price Offsets

We study the second variation of HBL which works by adding a price offset to the optimal value learned by the heuristic belief function. Different from the first variation which strategically constructs a dataset $\mathcal{D}$ by excluding potential spoof order samples, this second HBL variation considers all orders in memory and relies on a price adjustment to adapt to different market environments. We start the exploration with a set of offset intervals, ranging from positive $[R_1, R_2]$ values that induce more conservative bids to negative values that bid more aggressively compared to a basic HBL. Table 2 includes only strategies that are competitive enough to appear in equilibrium.



(a) Total HBL (and its level blocking variations) adoption rate in equilibrium across markets where the exploiter manipulates.



(b) Total background surplus achieved in equilibrium across markets where the exploiter manipulates.

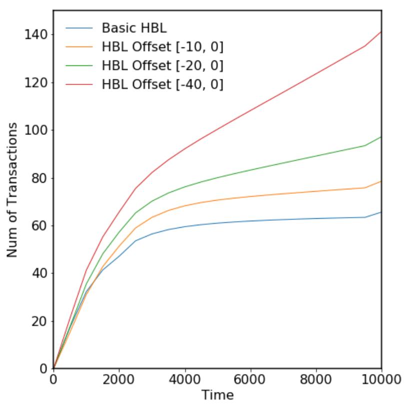
Figure 3. Learning trader proportion and total background surplus achieved in equilibria in games with manipulation. Each yellow (blue) marker specifies an outcome at one equilibrium where the group of learning traders does (does not) block orders from the second price level.

As in Section 4.3, we first conduct EGTA in games without spoofing and let background traders choose from ZI and a set of HBL with parametrically different price offsets. Figure 5 compares equilibrium outcomes in two *non-spoofing* markets, one where background traders are provided the flexibility to adjust prices with offsets and the other where agents are restricted to the standard HBL. We find that HBL traders, by bidding a little bit more aggressively (i.e., adopting negative price adjustments), can achieve consistently higher proportions in equilibrium (see Figure 5a). Specifically, this price adjustment enables learning traders to adapt better to high shock environments where prices are less predictable from past observations. Thanks to the overall improved HBL ratio in equilibrium, total background surplus also increases across all market environments as shown in Figure 5b.

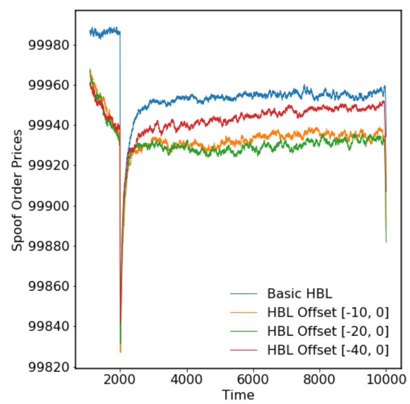
We, again, perform controlled experiments to investigate reasons behind the efficacy of including such price offsets. We focus on environment "MSMN" without spoofing and start from one found equilibrium where all background traders adopt HBL_6. We are interested in how transaction volume, spoof order prices, and differences in transaction prices vary as background traders deviate to HBL adopting different offsets and the exploiter also spoofs. The results are shown in Figure 4.

First, we measure the total number of transactions happened throughout the trading period in markets where HBL traders adopt different price offsets. As expected, Figure 4a shows that as traders submit orders at more aggressive prices, the market tends to have more transactions. We then let the exploiter manipulate, holding the remaining environment parameters and background strategy profiles unchanged. We are interested in comparing the price of spoof orders in those different market settings. We find, perhaps surprisingly, that the spoof-order prices are generally lower in markets where HBL traders bid more aggressively by adopting negative offsets (see Figure 4b). This may be because as more bids and asks result in transactions, the outstanding best bids are actually lower than those of markets where agents adopt basic HBL and submit orders at more conservative prices. Consequently, spoof orders are submitted at lower prices to avoid being transacted in the markets. This is further confirmed in Figure 4c, which demonstrates transaction price differences caused by manipulation. It shows though market populated with HBL of price offsets has a transaction price rise after the exploiter starts to spoofing at $T = 1000$, differences in transaction prices quickly decay compared to those of markets populated with basic HBLs.

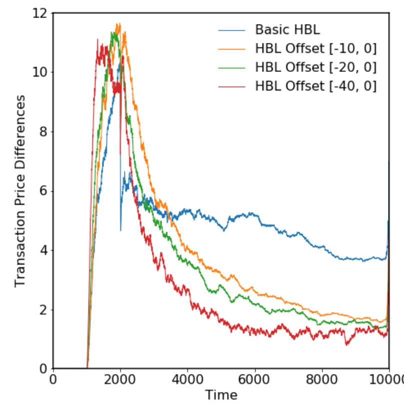
Finally, we investigate the performance of the second HBL variation in games with spoofing. We find in the face of manipulation, markets where trading agents are provided options of HBL with price offsets remain to achieve larger HBL adoption rates and higher surplus than those in markets where agents are restricted to standard HBL. However, manipulation does decrease social welfare, due to the consideration of spoof orders in the belief function.



(a) Number of transactions.

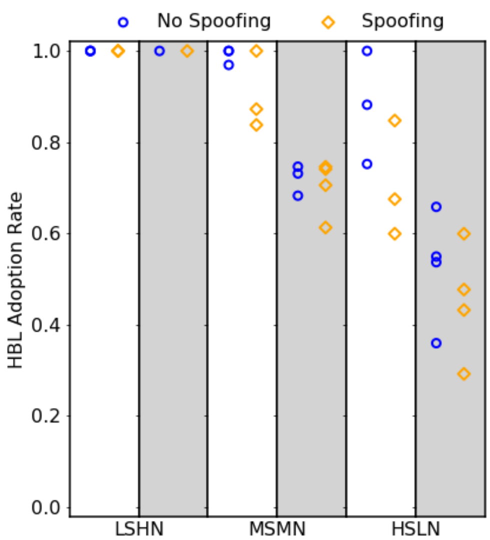


(b) Average prices of spoof orders.

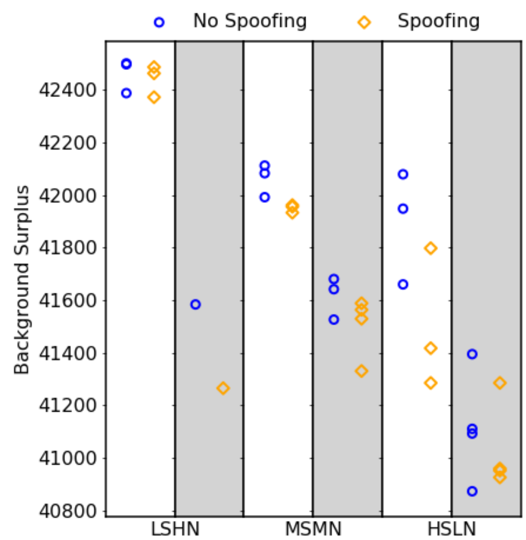


(c) Differences in transaction prices.

Figure 4. The accumulated number of transactions, the average price of spoof orders, and the transaction price rises caused by spoofing throughout the trading period across markets where background agents adopt basic HBL and its variations with different price offsets.



(a) Total HBL adoption rate at equilibrium in markets without and with spoofing.



(b) Total background surplus achieved at equilibrium in markets without and with spoofing.

Figure 5. The total adoption rate of HBL strategies and background surplus achieved in equilibria in games with and without spoofing. Markets represented in shaded background indicate that agents are restricted to the basic HBL strategy. Each yellow (blue) marker specifies an outcome at one equilibrium found in a specific game environment with (without) spoofing.

5. Conclusion

We study learning-based trading strategies by which individual traders can adopt to exploit market information but in less vulnerable ways in the face of market manipulation. We explored two variations based on the original HBL strategy that learns from the full order book. The first variation selectively blocks orders at certain price levels, particularly where spoof orders are likely to be placed. The second considers the full order book, but adjusts the offer price by an offset, aiming to correct any biases in the learning process. We employ agent-based simulation to evaluate the two proposed variations on HBL under equilibrium settings in markets where background traders can adopt the non-learning ZI and the HBL strategies, both in its basic form and the two variations. We demonstrate that both variations can reduce the learning agents' vulnerability to spoof orders, and thus increase the overall background-trader surplus in the market. Empirical game-theoretic analysis implied that trading agents can strategically adapt to the presence of manipulation, even without any intervention to regulate the market.

References

Biais, B. and Woolley, P. High frequency trading. Technical report, Toulouse University, 2012.

Chakraborty, T. and Kearns, M. Market making and mean reversion. *11th ACM Conference on Electronic Commerce*, pp. 307–314, 2011.

Copeland, T. E. and Galai, D. Information effects on the bid-ask spread. *Journal of Finance*, 38(5):1457–1469, 1983.

Farmer, J. D., Patelli, P., and Zovko, I. I. The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Sciences*, 102:2254–2259, 2005.

Foucault, T., Röell, A., and Sandås, P. Market making with costly monitoring: An analysis of the SOES controversy. *Review of Financial Studies*, 16(2):345–384, 2003.

Gjerstad, S. The competitive market paradox. *Journal of Economic Dynamics and Control*, 31:1753–1780, 2007.

Gode, D. K. and Sunder, S. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, pp. 119–137, 1993.

Lin, T. C. W. The new market manipulation. *Emory Law Journal*, 66:1253–1314, 2015.

Prewit, M. High-frequency trading: Should regulators do more. *Michigan Telecommunications and Technology Law Review*, 19:131–161, 2012.

U.S. Securities and Exchange Commission. Division of enforcement annual report, 2017.

U.S. Securities and Exchange Commission. SEC charges microcap fraudsters for roles in lucrative market manipulation schemes, 2018a.

U.S. Securities and Exchange Commission. Division of enforcement annual report, 2018b.

Wah, E., Wright, M., and Wellman, M. P. Welfare effects of market making in continuous double auctions. *Journal of Artificial Intelligence Research*, 59:613–650, 2017.

Wang, X. and Wellman, M. P. Spoofing the limit order book: An agent-based model. In *16th International Conference on Autonomous Agents and Multiagent Systems*, pp. 651–659, 2017.

Wang, X., Vorobeychik, Y., and Wellman, M. P. A cloaking mechanism to mitigate market manipulation. In *27th International Joint Conference on Artificial Intelligence*, pp. 541–547, 2018.

Wellman, M. P. Putting the agent in agent-based modeling. *Autonomous Agents and Multi-Agent Systems*, 30:1175–1189, 2016.

Wiedenbeck, B. and Wellman, M. P. Scaling simulation-based game analysis through deviation-preserving reduction. In *11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 931–938, 2012.