An Agent-Based Model of Competition Between Financial Exchanges: Can Frequent Call Mechanisms Drive Trade Away from CDAs?

Zhuoshu Li
Washington University in St. Louis
zhuoshuli@wustl.edu

Sanmay Das
Washington University in St. Louis
sanmay@wustl.edu

ABSTRACT

In the debate over high frequency trading, the frequent call (Call) mechanism has recently received considerable attention as a proposal for replacing the continuous double auction (CDA) mechanisms that currently run most financial markets. One natural question, which has begun to spur the development of new models, is the effect of competition between platforms that use these two different mechanisms when agents can strategize over platform choice. In this paper we contribute to this nascent literature by developing an agent-based model of competition between a Call market and a CDA market. Our model incorporates patient informed traders (both high-frequency and not) who are willing to wait for order execution at their preferred price and impatient background traders who demand immediate execution. We show that there is a strong tendency for the Call market to absorb a significant fraction of trade under most equilibrium and approximate-equilibrium conditions. These equilibria typically lead to significantly higher welfare for the background traders, an important measure of social value, than the operation of an isolated CDA market.

Keywords

Competing platforms; agent-based modeling; market microstructure

1. INTRODUCTION

Most modern financial exchanges operate using the continuous double auction (CDA) mechanism, which in principle allows for trading in continuous time, at least to within our measurement and implementation capabilities. With companies investing in faster infrastructure for trading and events like the "flash crash" of May 2010, high frequency trading (HFT) has become an increasingly debated topic in both the media and policy spheres [15]. Proponents claim that high-frequency trading improves liquidity and price discovery. Improved liquidity means lower transaction costs for average investors, while better price discovery serves the social information aggregation and dissemination role of market prices [2, 16]. However, there is increasing evidence that at least one form of high frequency trading, namely latency arbitrage, has reached a point of socially diminishing returns. Budish et al demonstrate this both empirically and through a simple model: empirically, they show that correlations between virtually identical assets being traded in different markets break down at very small timescales, while they are essentially perfect at larger timescales [3]. This can almost be thought of as a law of physics – there is no natural force tying the assets or markets together, so there is no way to make them actually move simultaneously. What is problematic is the "arms race" this creates to extract the maximum profit from squeezing this reaction time down as much as possible. Budish et al show that this is not only socially inefficient, it can actually create thinner markets. Along similar lines, Wah and Wellman build a model where an asset is traded on two markets, and there is an infinitely fast latency arbitrageur present. They show how the presence of the arbitrageur can hurt social welfare [22]. Both sets of authors recommend frequent batch auctions as a market structure that could replace CDAs, since the minimum time period between trades is specified, and there is no benefit to being faster than that.

An important question for the possible use of frequent batch (or call) auctions is how they would work in the presence of existing CDA markets. Competition between exchanges or platforms that try to attract trade is a vast topic, and there is evidence in many domains that platforms with better welfare properties *assuming that there is only one platform available* may not be able to capture enough of the market for these properties to become evident when they face competition from other platforms. For example, in living-donor paired kidney exchange, even though exchanges that wait to build thickness may be socially preferable, exchanges that match greedily can make them nonviable [6]. Therefore, even though they may have desirable welfare properties, could call auction based markets actually take volume away from CDA markets if both existed simultaneously? Wah et al have engaged this question using empirical game theoretic analysis [21]. They develop a model where the environment is populated by fast (HFT) and slow (non-HFT) traders. They argue that a frequent call market in the wild could attract sufficient volume for viability from two perspectives: first, in equilibrium, welfare of slow traders is generally higher in the call market, where they are relatively protected from sniping and adverse selection, and second, fast traders are willing to follow the slow traders to either market, including to the call market, so it could serve as a basin of attraction. Wah et al's model does not consider traders who have a preference for immediacy, and

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

they also restrict traders to choose a single market and then do not allow traders to move. While their results are quite promising, we seek to build a richer model that combines aspects of classic financial market microstructure models and agent-based models that are known to replicate important properties of order books.

Another line of literature relates to the TAC Market Design Competition (CAT) [17]. In this competition, participants aim to design better mechanisms to maximize a score (a combination of profit, market share and transaction success rate) when traders are drawn from a known population of different types. CAT gives a general view of competition among different markets, but our paper focuses on a comparison of two more specific market mechanisms and how they influence the social welfare of traders.

1.1 Overview and Contributions

In this paper, we build an agent-based model to analyze platform competition between financial exchanges. Agent-based modeling seeks to fill the hole in simple stylized models which may not represent agent behavior in sufficiently complex manners to really capture the essence of the important phenomena. The last two decades have seen substantial work on agent-based modeling of financial markets, using both sophisticated [4, 7, 9, 18] and simple [8, 12, 14] trader models in the population. Our model is as parsimonious as possible while attempting to capture the essential relevant behaviors that are important to understanding the behavior of these markets. As such, it follows the basic structure of classic models of market microstructure such as those of Glosten and Milgrom [11] and of Kyle [13], in which there are informed traders, who possess superior information and trade in search of profit, liquidity (or background) traders, who trade for exogenous reasons (e.g. retirement funds that receive cash and need to track indices, or investors liquidating portfolios in retirement or in order to buy a house, say) and demand immediacy, and market makers, who may be employed in order to facilitate price discovery and trade execution.

Our key measure of welfare is the price of immediacy – the expected loss suffered by background traders. This measures the cost that the "average trader" pays in order to execute transactions. This is a different measure than that of Wah and Wellman [23] or Wah et al [21], who use surplus. These are both reasonable measures, but surplus is most meaningful in private value models, where some meaning can be attributed to different agents having different valuations for an asset. Our model follows in a tradition of common value models, where the asset has a true underlying value, and different traders may have different estimates of that true value. The existence of background traders in our model provides a useful proxy for estimating the cost of trading. It is worth noting that this doesn't mean that background traders are necessarily losing money – typically such traders would stay in the market for much longer, and under reasonable models of price appreciation, these "losses" can be thought of as transaction costs for buy-and-hold type investors.

First, we look at simple models of individual markets and confirm that our model satisfies the basic intuitions one would expect. Namely, informed traders (in particular, low latency traders) make more profit (and background traders are consequently made worse off) in CDA markets than in frequent call markets. A zero-profit market maker (with no specialized information) can greatly improve the position of background traders, taking away most of the profit opportunities from informed traders in CDA markets. Next, we model competition between a CDA market and a frequent call market when informed traders pick which market to place their orders in based simply on which market is more mispriced with respect to their current belief. We show that the informed traders do better overall when they choose to place orders in the market that is more mispriced from their perspective. We show that, when informed traders are all using this strategy, a majority of orders flows to the call market, and background traders are better off than in a single CDA market.

Note that all of the above analysis is not in an equilibrium setting – we assume that all informed traders use the same strategy. We can use the insights developed in these models to begin analyzing strategic market choice. We do so by introducing a learning framework, where informed traders learn a parametric form for the expected profit of choosing to place an order in a market (and a non-parametric probability of order execution) given the distance of that market's "current price" from the trader's estimate of the true value of a stock. We show that, when all agents use this learning approach, they converge to an approximate equilibrium where a majority of trades again flow to the frequent call market.

2. MARKET MODEL

2.1 The CDA and Call Markets

Our model of competing markets consists of two markets, one employing a *continuous double auction* (CDA) mechanism and the other one employing a *frequent call* (CALL) mechanism. We begin by describing the details of each individual market, which will serve as the foundation for our model of competing markets.

Each market is running in the continuous-time interval [0,T]. A single security is traded in the market. There is an underlying "true value" process. The initial true value of the security v_0 is drawn from a Gaussian distribution with mean v_{initial} and standard deviation σ_{initial}. Then, the true value jumps according to a Poisson process with rate parameter λ_{jump}. If the true value jumps, the new true value v_t is generated from $v_t \sim \mathcal{N}(v_{t-dt}, \sigma_j)$, where v_{t-dt} is the price instantaneously before the jump (we restrict $v_t \geq 0$, so all values are truncated at 0).

In the CDA market, outstanding orders are maintained in two priority queues: one for bids (the buy orderbook) and one for asks (the sell orderbook). Bids and asks are prioritized by price first and time second. When a new order comes in, it is added to the corresponding order book. A trade is executed if the highest bid exceeds or is equal to the lowest ask. The execution involves the orders at the top of the bid and ask queues, at the price of the older of the two orders involved.

The CALL market is similar to that described by Budish et al [3]. It clears in fixed intervals of time τ (the *call interval*). At each clearing time, the market collates all of the orders and computes the aggregate demand and supply functions of all bids and asks, respectively. The market clears where supply equals demand, with all executions occurring at the same price, called the market-clearing price. None of the orders are visible to any traders during the call interval. The market announces the market-clearing price after each clearing (*market announcement*). When no order was executed at the last clearing time, if both the buy and sell orderbooks are not empty, the market announcement will be the mid-point of the highest bid and lowest ask, otherwise it will be the most recent available market-clearing price. All untraded orders roll into the next call.

2.2 Valuation model

Each trader has a private valuation for the security (or equivalently for our purposes, a private signal of the true value). We have two types of traders, informed (IF) traders and background (BG) traders. Each informed trader IF_i receives a private signal of the security value, $w_{i,t} \sim \mathcal{N}(v_{\hat{t}}, \sigma_{\text{trader}})$, where $v_{\hat{t}}$ is the underlying true value of the security at some time $\hat{t}$, where $\hat{t} \leq t$, and σ_{trader} is a noise parameter. We define two types of informed traders, namely, low latency (LL) traders and high latency (HL) traders. High latency traders have staler information, i.e., $w_{i,t} \sim \mathcal{N}(v_{t-\delta}, \sigma_{\text{trader}})$, and low latency traders observe information with no delay, thus $w_{i,t} \sim \mathcal{N}(v_t, \sigma_{\text{trader}})$. Background traders do not have any private information; each arriving background trader wishes to either buy or sell one unit (with equal probability). They demand immediacy, that is, they want to get their orders executed as soon as possible, so they are willing to take any market price.

2.3 Agent arrival process

There are a fixed number of traders of each type. Informed traders and background traders both arrive at the market according to separate Poisson processes, with informed traders arriving with rate λ_{IF} and background traders' arriving with rate λ_{BG}. In the event that an informed trader arrival occurs, a specific IF trader is selected uniformly at random from all the IF traders to place/replace an order; similarly, if a background trader arrival occurs, a specific BG trader is selected uniformly at random from all the BG traders to place/replace an order.

2.4 Agent strategies in individual markets

Each of the informed traders and background traders is only allowed to maintain a single unit order in the market. When informed and background traders reenter the market, they can replace existing orders that have not yet been executed. We model informed traders as using limit orders and background traders as using market orders exclusively.

Informed traders' strategy: $S_{\text{individual}}$.

We consider trading strategies in the *Zero Intelligence* (ZI) family for informed traders. There is a large literature involving ZI strategies, including some controversy, which we will not rehash here [4, 19]. While ZI strategies are clearly not the "best" trading strategies in isolation, it is also generally believed that they model order arrival processes well, and they are a standard method for choosing prices in complex agent-based market simulations [9, 18]. We first define, in the CDA market,

$$p^*_{t,\text{CDA}} = \begin{cases} \dfrac{(BID_t.p) + (ASK_t.p)}{2}, \text{if } BID_t \text{ and } ASK_t \text{ exist,} \\ \text{the most recent execution price, if any order} \\ \hfill \text{book is empty,} \end{cases}$$

where $(BID_t.p)$ and $(ASK_t.p)$ refer to the price of BID_t and ASK_t respectively. And in the CALL market,

$$p^*_{t,\text{CALL}} = \text{the most recent } \textit{market announcement}.$$

When an informed trader IF_i places an order, a limit price is generated from

$$p_{i,t} \sim \mathcal{N}(p^*_{t,\text{market}}, \sigma_{\text{price}}),$$

where market $\in \{\text{CDA}, \text{CALL}\}$. Based on $p_{i,t}$ and $w_{i,t}$, the informed trader IF_i's strategy at time t is as follows,

$$p_{i,t} \begin{cases} > w_{i,t}, \text{places a unit sell order,} \\ < w_{i,t}, \text{places a unit buy order,} \\ = w_{i,t}, \text{uniformly at random places a unit buy or sell order.} \end{cases}$$

Note that $p_{i,t} = w_{i,t}$ is a zero probability event. If $p_{i,t} > (ASK_t.p)$ and the order is a buy order, then it executes immediately and therefore effectively functions as a market order. Similarly if $p_{i,t} < (BID_t.p)$ and the order is a sell order. We call the strategy above $S_{\text{individual}}$.

Background traders' strategy.

The background traders choose whether they want to buy or sell a unit uniformly at random. Once the direction is decided, the order is routed to the market and handled in a special manner as a market order through a "waiting" mechanism. The market is aware of the direction of a market order and the fact that this indicates the trader would like to execute the order at any available market price. However, market orders are not visible to any other traders in both the CDA and CALL markets, since they may need to wait for execution if there is no corresponding limit order on the other side in the CDA market, and at least until the next call in the CALL market.

Market maker's strategy.

In the CDA market, we also incorporate a market maker in some of our experiments. To increase the liquidity of the market, the market maker maintains a unit buy order and a unit sell order at all times. This market maker is implemented using the Bayesian market making algorithm (BMM) of Brahma et al [1], with parameters tuned to maintain near zero-profit. BMM is a learning algorithm that learns from the current bid and ask prices and the direction of incoming trades, augmented with jump prediction and a technique to widen its spread in times of uncertainty. BMM updates its own belief whenever there is an execution, and it immediately replaces its orders. Our implementation closely follows that of Brahma et al, except that we only need to use it for unit orders in our model.

2.5 CDA and CALL market operation

CDA market operation.

In the CDA market, BID_t and ASK_t are based only on orders from informed traders (and also possibly the market maker). If the order book only has market orders that are waiting from background traders, the market shows the order book as empty. The scenario that we want to simulate is that background traders are waiting in the market to buy or sell; as soon as an unfilled corresponding order becomes available, they will immediately take the other side of that order. We need to specify the execution priority in the situation where one side of the market has both market orders and limit orders from informed traders. In this case it must be that the other side of the market is empty (note that this never happens with a market maker present), otherwise the market orders on the first side would have executed. In this situation, we prioritize by time. Procedure 1 illustrates the operation of a CDA market when a new buy order arrives (a sell order arrival is similar).

Procedure 1 CDA market operation when a new buy order arrives

Input: buy orderbook, sell orderbook

```
A new buy order OD_1 arrives, (OD_1.p) is the price of OD_1
if not empty(sell orderbook) then
  if OD_1 is a limit order then
    if (OD_1.p) ≥ (ASK_t.p) and ASK_t comes earlier than any market order then
      Execution(ASK_t, OD_1) at (ASK_t.p)
      (in which case, OD_1 is the highest bid and sell orderbook has limit orders from IF traders or BMM)
    else
      if sell orderbook contains market orders then
        Execution(the oldest market ask, OD_1) at (BID_t.p)
        (in which case, OD_1 is the highest bid)
      end if
    end if
  else {OD_1 is a market order}
    if ASK_t is available and comes earlier than any market order then
      Execution(ASK_t, OD_1) at (ASK_t.p)
    else
      Execution(the oldest market ask, OD_1) at the most recent execution price
    end if
  end if
end if
```

CALL market operation.

The main difference from a standard aggregation mechanism in our implementation involves the background traders. Background traders would like to buy or sell at any price, so all the market orders are always at the top of both the sell orderbook and the buy orderbook in the CALL market. All the market orders in each orderbook are prioritized by submission time, with earlier submissions having higher priority. At each clearing time, the market collates all of the orders and computes the aggregate demand and supply functions of all bids and asks, respectively. The market clears where supply equals demand, with all executions occurring at the same price, the market-clearing price. If the market only clears market orders, the market-clearing price is the most recent market announcement. If only market orders clear on one side of the market, while some limit orders clear on the other side, the market-clearing price is determined by the side that has limit orders being cleared. More specifically, if cleared buy orders consist of only market orders and cleared sell orders include limit orders, the clearing price will be the highest ask of all the cleared limit orders; if cleared sell orders consist of only market orders and cleared buy orders include limit orders, the clearing price will be the lowest bid of all the cleared limit orders. When some limit orders clear on both sides of the market, the clearing price is the midpoint of the highest ask and lowest bid of all cleared limit orders.

3. COMPETING MARKETS

In the competing markets model, we assume that one CDA market and one CALL market run simultaneously. A single security is traded in both markets. Thus, there is only one underlying "true value" process, but the CDA and CALL markets can price the security differently. The traders choose to place orders in only one market at a time, although they can switch markets each time they re-enter. Each market is running in the same manner as when there is an individual market and each trader can maintain only one unit order in the whole system.

3.1 Agent strategies in competing markets

Informed traders' strategy.

On what basis should a trader choose which market to place an order in? One important factor is the distance between the trader's belief and $p^*_{t,\text{market}}$,

$$d^{\text{market}}_{i,t} = |w_{i,t} - p^*_{t,\text{market}}|, \tag{1}$$

where market $\in \{\text{CDA}, \text{CALL}\}$. We call this the *belief distance*. Comparing $d^{\text{CDA}}_{i,t}$ with $d^{\text{CALL}}_{i,t}$, the informed trader IF_i can choose one of two strategies. One is to place the order in the market that has larger $d^{\text{market}}_{i,t}$, S_{LARGE}, and the other is to place the order in the market that has smaller $d^{\text{market}}_{i,t}$, S_{SMALL}. The tradeoff here is that IF_i gets lower probability of execution but higher profit if she places the order in the market that has larger $d^{\text{market}}_{i,t}$. We will discuss the effects of these two different strategies in Section 4. Strategy 1 shows a summary of IF_i's strategy in the competing markets. After IF_i decides in which market to place the limit order, she follows $S_{\text{individual}}$ to decide the direction and price of the order in the selected market.

Strategy 1 IF_i's strategy in the competing markets at time t

1: **if** following S_{LARGE} **then**

2: $$d^{\text{CDA}}_{i,t} \begin{cases} > d^{\text{CALL}}_{i,t}, & \text{places an order at CDA market following } S_{\text{individual}} \\ < d^{\text{CALL}}_{i,t}, & \text{places an order at CALL market following } S_{\text{individual}} \end{cases}$$

3: **else** {following S_{SMALL}}

4: $$d^{\text{CDA}}_{i,t} \begin{cases} < d^{\text{CALL}}_{i,t}, & \text{places an order at CDA market following } S_{\text{individual}} \\ > d^{\text{CALL}}_{i,t}, & \text{places an order at CALL market following } S_{\text{individual}} \end{cases}$$

5: **end if**

Background traders' strategy.

When a background trader enters the market to place or replace a new order, she first compares the price in both markets. For instance, if she wants to buy, she will compare

ASK_t and $p^*_{t,\text{CALL}}$ (if they are available), and select the market which has the lower price to place a market order there. If ASK_t is not available, that is the sell orderbook in the CDA market shows as empty, she will place the order in the CALL market. She follows a similar process for sell orders.

After every market clearing in the CALL market, the background traders check whether their orders have been executed. If not, and the corresponding order books in the CDA market are not empty, that is the sell orderbook is not empty if a BG trader wants to buy and the buy orderbook is not empty if a BG trader wants to sell (here empty means the order books do not have orders from informed traders – market orders from background traders are not visible to any trader), the background traders move their existing orders from the CALL market to the CDA market. In the implementation, the background traders with orders that did not execute are randomly permuted. Each of them moves their order to the CDA market in this random order, until there are no corresponding limit orders on the other side of the market in the CDA. This process is atomic in time.

Strategy 2 shows the overall framework for implementing buy orders for background traders in the competing markets model. The sell strategies are similar.

Strategy 2 The overall framework for implementing the buy orders for background (BG) traders in the competing markets model

$t = 0$
while $t \leq T$ **do**
  An event happens after $\Delta\ t$
  $t = t + \Delta t$
  if A BG trader arrives in the market **then**
    BG_i is selected from all BG traders uniformly at random
    if ASK_t is available **then**
      if $ASK_t <= p^*_{t,\text{CALL}}$ **then**
        BG_i places a market order in CDA market
      else
        BG_i places a market order in CALL market
      end if
    else {sell orderbook in CDA market shows as empty}
      BG_i places a market order in CALL market
    end if
  end if
  if CALL market clear **then**
    After each market clearing
    A = {All the BG traders who have buy orders in CALL market}
    while A is not empty and ASK_t is available **do**
      Select BG_i uniformly at random from A
      $A = A - \{\text{BG}_i\}$
      BG_i moves her order to the buy orderbook in CDA market, the new order age is re-generated from the age counter.
      CDA market clears
    end while
  end if
end while

4. SIMULATION RESULTS

In this section, we simulate four different environments, namely CDA vs CALL competing markets (competing markets), an individual CDA market (i-CDA market), an individual CALL market (i-CALL market) and an individual CDA market with BMM (i-CDA-BMM market). The parameters are set as follows. Each simulation run lasts $T = 100,000$ units of time. The initial true value of the security v_0 is drawn from $\mathcal{N}(v_{\text{initial}} = 50, \sigma_{\text{initial}} = 4)$. The true value jump parameter $\sigma_j = 4.0$ and the rate parameter for the jump is $\lambda_{\text{jump}} = 0.0001$, which means there is a jump every 10000 units of time on average. We have 20 informed traders, 10 high latency traders and 10 low latency traders, and 20 background traders. Reentry rates are fixed across the environments, with informed traders arriving in the market at rate $\lambda_{\text{IF}} = 2$, and background traders entering at rate $\lambda_{\text{BG}} = 1$. In all settings, CALL markets clear every 1 unit of time, $\tau = 1$. The standard deviation of informed traders' belief is $\sigma_{\text{trader}} = 2.0$, and the high latency traders' information delay is $\delta = 1000$ units of time. The time between jumps is 10000 units of time on average, so high latency traders have information with no delay a significant fraction of the time, as $v_{t-\delta} = v_t$. Following a jump, high latency traders receive staler information for the next 1000 units of time. The standard deviation of the distribution from which informed traders draw ZI prices is $\sigma_{\text{price}} = 4.0$. We simulate both S_{LARGE} and S_{SMALL} strategies for informed traders.

Across our experiments, we are interested in the total profit, expected per-order profit and order execution percentage for each trader type. At time T, all shares held by traders are liquidated at price v_T. Unfilled orders are abandoned. The expected per-order profit of each trader type is total profit divided by the total number of executed and replaced untraded orders. In the environment with competing markets, we also calculate the total and expected per-order profit in CDA market and CALL market separately.

Figure 1 shows that the informed traders make higher profit in both per-order and in total when using S_{LARGE}. The difference are small but statistically significant. Therefore, we would expect the informed traders to choose S_{LARGE} if given these two options (if they had to choose one as a group), confirming our intuition that traders gravitate to markets in which they perceive more mispricing. Because of this, for the rest of our analysis, we use S_{LARGE} as the strategy for informed traders in the competing markets.

As mentioned in the introduction, one measure of social welfare is the "price of immediacy" which is the loss suffered by background traders. Figure 2 shows that, for the non-competing settings, background traders perform better in the i-CALL market than the i-CDA market in terms of both the expected per-order (left figure) and total (right figure) profit (consequently, informed traders have lower profit in the i-CALL market than the i-CDA market). The i-CDA-BMM market has much higher social welfare, as measured by background trader losses, than both i-CALL and i-CDA markets. This confirms some of the results of Wah and Wellman [23] in a completely different model and setting. One possible solution to the problems resulting from HFT may then be to have market-making agents who are regulated and deployed to perform this specific role in CDAs. They could be compensated separately for this role. However, (1) there are additional risks associated with this role [5] and (2) markets have been moving away from having designated

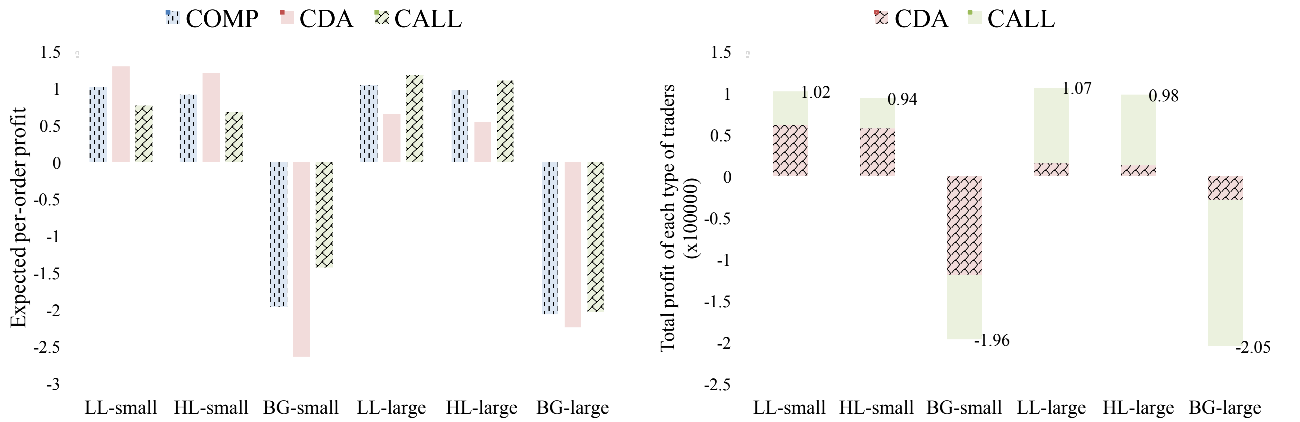

Figure 1: Comparison of expected per-order (left) and total (right) profit in the competing markets under the small distance (S_{SMALL}) and large distance strategies (S_{LARGE}). LL-*, HL-* and BG-* represent low latency, high latency, and background traders respectively. On the left, the first bar in each group shows the expected per-order profit in the whole competing system, while the second and third show the contributions of the CDA and CALL markets to that total. On the right, the stacked bars show total profit, with contributions from each of CDA and CALL shown within in different shades.

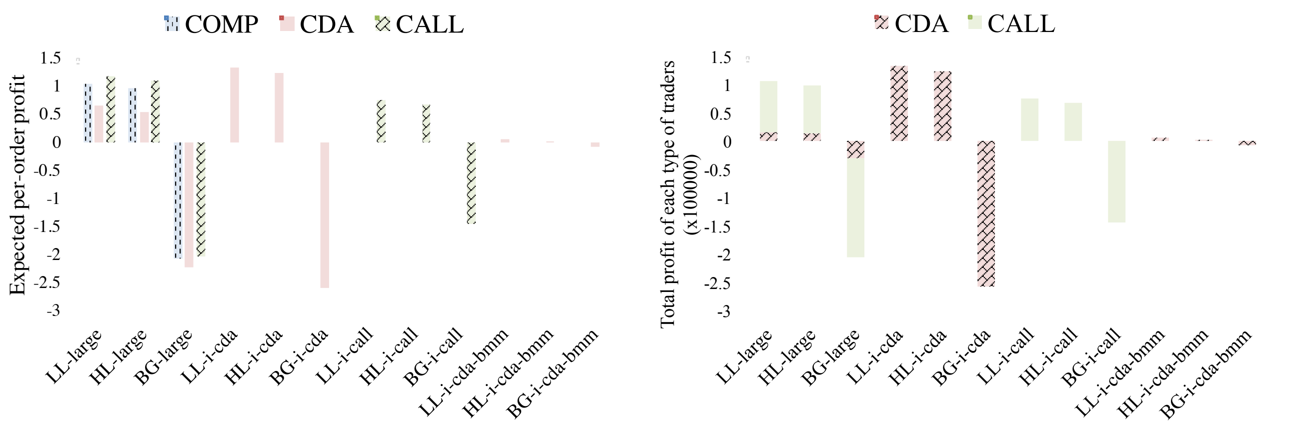

Figure 2: The expected per-order (left) and total profit (right) of different types of traders in the competing markets (*-large), the i-CDA market (*-i-cda), the i-CALL market (*-i-call) and the i-CDA-BMM market (*-i-cda-bmm).

specialists and allowing HFTs and others to fulfil the role of market-makers. Given these practical realities, it is important to understand how markets can function without them, so we focus on comparing situations with no market maker.

In the competing markets (first 3 panels of Figure 2), we analyze the expected per-order and total profit in the whole system, and also in the CDA market and CALL market separately. Similar to running an individual market, the order execution percentage is close to 100 for background traders in the competing markets (shown in Figure 3), and so background traders are not losing out in terms of order execution. Considering expected per-order profit, background traders do better in the CALL market than the CDA market in the competing markets (see left figure of Figure 2, *-large). Overall, they are doing worse in the CALL market than the CDA market in the competing system (shown on the right of Figure 2, *-large), but this is because a vast majority of orders are going to the CALL market (see Figure 4), and they lose more money there. In sum, they are doing slightly better in the competing markets than they do in the i-CDA market in terms of both the expected per-order and total profit. The better news is that the CALL is absorbing a large fraction of the orders, driving trade away from the CDA (see Figure 4). This is promising, because if the CALL could absorb all the trades, the BG traders would be better off, as the system would reduce to the i-CALL market. We note that these results are robust for a wide range of strategy parameters, information delay and arrival rates.

5. LEARNING TRADERS

The analysis above shows that frequent call markets absorb a large fraction of trade when we assume that all informed traders use the same strategy. We now use the insights developed in these models to analyze strategic market choice. In this section, we introduce a learning framework where informed traders learn a parametric form for the expected profit of choosing to place an order in a market, and also a non-parametric probability of order execution given

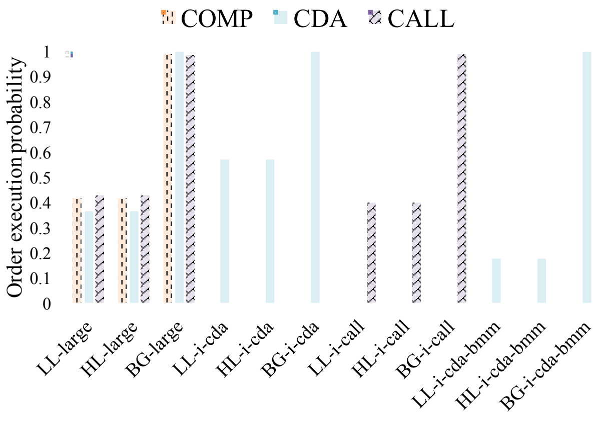

Figure 3: Order execution percentage in the competing markets (*-large), the i-CDA market (*-i-cda), the i-CALL market (*-i-call) and the i-CDA-BMM (*-i-wmm).

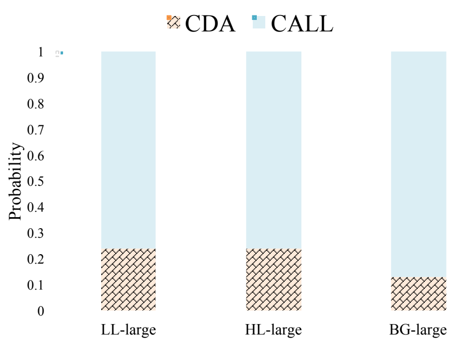

Figure 4: Illustration of proportion of orders entering CDA market vs. CALL market in the competing markets under S_{LARGE}.

the distance between that market's price and the trader's belief (the *belief distance* defined previously). We then analyze the behavior of the system with these learning traders.

5.1 Learning algorithm

The expected profit π_t of an order placed at time t, contingent on its execution, according to a trader's belief, is

$$\pi_t = \begin{cases} w_t - \text{execution price, if buy} \\ \text{execution price } - w_t, \text{if sell} \end{cases} \quad (2)$$

where w_t is the trader's belief about the true value at time t. The trader computes expected profit assuming her belief is correct. The main idea here is to predict the expected profit if the order is placed in a particular market. Traders must learn an estimate of this as a function of the belief distance. In this paper, we allow all the traders to learn an expected profit function that is quadratic in the belief distance. Each learning trader LIF_i uses an online regression algorithm for reinforcement learning based on one developed by Walsh et al [24]. The form of the learning model is:

$$y_{i,t}^{\text{market}} = \mathbf{D}_{i,t}^{\text{market}} \mathbf{Q}_i^{\text{market}}, \quad (3)$$

where $y_{i,t}^{\text{market}}$ predicts the expected per-order profit contingent on execution, market $= \{\text{CDA}, \text{CALL}\}, \mathbf{D}_{i,t}^{\text{market}} = [(d_{i,t}^{\text{market}})^2, d_{i,t}^{\text{market}}, 1]$ and $\mathbf{Q}_i^{\text{market}}$ contains the weight parameters of the model for market $= \{\text{CDA}, \text{CALL}\}$. Thus, the predicted expected per-order profit not contingent on execution is given by

$$\mathbb{E}(y_{i,t}^{\text{market}}) = \Pr(\text{exe}|d_{i,t}^{\text{market}}) \mathbf{D}_{i,t}^{\text{market}} \mathbf{Q}_i^{\text{market}} \quad (4)$$

The probability of execution $\Pr(\text{exe}|d_{i,t}^{\text{market}})$ is learned non-parametrically by counting successful and unsuccessful executions in bins of the belief distance.

The trader uses an ϵ-greedy algorithm to select a market to trade in along the learning path. Whenever the trader makes a decision, with probability $1-\epsilon$, she places an order in the market with higher predicted expected profit, and with probability ϵ, she randomly picks one market to place the order ($\epsilon = 0.1$ in our case). After market selection, the trader chooses a price based on the ZI strategy $S_{\text{individual}}$.

5.2 Results

Our goal is to use this model of learning traders to investigate two questions: (1) If all the informed traders use the same learning algorithm, do they converge to (approximate) equilibrium strategies? (2) Can we characterize any equilibria of the competing markets system?[1] We use an experimental framework similar to Section 4. A CALL market and a CDA market run simultaneously from $[0, T]$, $T = 100,000$. The CALL interval $\tau = 1$. There are 20 learning informed traders (LIF). These are all low latency traders who observe information with no delay; therefore $w_{i,t} \sim \mathcal{N}(v_t, \sigma_{\text{trader}})$, where v_t is the underlying true value of the security at time t, and σ_{trader} is the noise parameter. The reentry rate of learning informed traders is $\lambda_{\text{LIF}} = 2$. There are 20 background traders (with reentry rate $\lambda_{\text{BG}} = 1$) following Strategy 2. In addition, we also simulate the existence of a pool of *fixed* informed traders (FIF) who are committed to a particular market, either CDA or CALL. This is to ensure that there is some flow of trade in each market – otherwise there are degenerate equilibrium paths where all traders start off by going to one of the markets, and there is never incentive to deviate to the other. There are 5 fixed (low-latency) informed traders in each market who place orders following $S_{\text{individual}}$. The reentry rate of fixed informed traders is defined as λ_{FIF}. We vary the reentry rate of fixed informed traders $\lambda_{\text{FIF}} = \{1, 0.1, 0.0005\}$.

Outcomes of the learning process.

The first question is whether the learning process followed by the informed traders converges, and, if so, whether the learned representations are a good approximation to the true profit function. Empirically, we find that the estimates of $\mathbf{Q}_i^{\text{CDA}}, \mathbf{Q}_i^{\text{CALL}}$ under different λ_{FIF} settings all do converge. Further, each trader's parameters converge to very similar ranges (see Figure 5). We check whether these parameters are a good approximation by fixing the parameters of all

[1] Note that traders only choose which market to place an order in (albeit as a function of the belief distances to both markets, so this can be a complex decision space). Once the market is determined, the choice of price is according to the ZI strategy. The problem becomes exponentially more complex if traders can strategize over both market choice and price. Over time, traders should learn the expected profit of the ZI strategy in a market, a useful proxy for the profit potential of that market. It could be interesting to interact this learning problem with different pricing strategies, but some restriction will always be necessary to gain any traction.

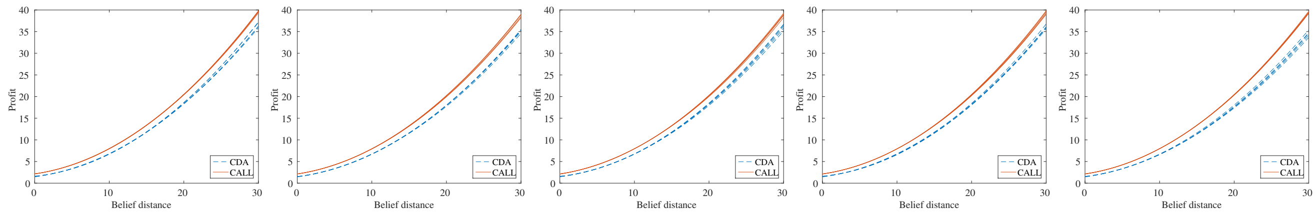

Figure 5: Five example plots of the curves traders learn for profits in the CDA and CALL markets. Each graph shows the learned curves of 5 traders in one instantiated learning simulation.

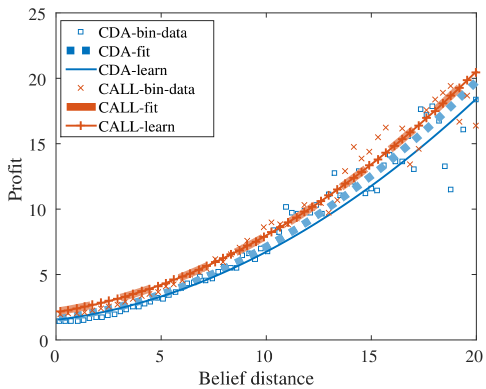

Figure 6: The learned curve vs. the best fit to realized after-the-fact data for profit as a function of belief distance in CDA and CALL markets.

the traders except one and having them play strategies using those parameters. For the remaining trader, we flip a coin to determine the choice of market, and test whether the profit achieved is well-fit by the curve given by the learned parameters. Figure 6 shows that the learning curves of the trader for each market are very close to the polynomial curves that are best-fit to the profits achieved using the randomized strategy, confirming both that the quadratic space is a good fit and that the learned parameters are correct for the environment.

Equilibrium.

As mentioned above, there are two main questions we would like to engage. First, since all the traders are converging to a particular set of learned parameters, do these parameters constitute an equilibrium or an approximate equilibrium (under the specified space of strategies – i.e., where the strategy is a mapping from d^{CALL} and d^{CDA} to one of the two markets, which can be specified by the quadratic form of the expected profit function and the nonparametric probability of execution model)? Is there a profitable deviation (some other set of parameters that one of the traders could use and increase her profits)?

To find deviations, we search the parameter space (holding the execution probability model constant) for this trader using Bayesian Optimization (BO), a powerful framework for optimization of a black-box function or expensive objective function that uses very few function evaluations [20]. Here, the objective function is the expected profit of a trading strategy that uses the parameters $\mathbf{Q}_i^{\text{CDA}}, \mathbf{Q}_i^{\text{CALL}}$ when the other traders are using their learned strategies. Market selection is determined by $y_{i,t}^{\text{CDA}}$ and $y_{i,t}^{\text{CALL}}$ based on Equation (4), so the actual values of $\mathbf{Q}_i^{\text{CDA}}$ and $\mathbf{Q}_i^{\text{CALL}}$ are not important in themselves. The important thing is how they decide the relation between $y_{i,t}^{\text{CDA}}$ and $y_{i,t}^{\text{CALL}}$ at each prediction. Therefore, based on the value of parameters of the learned curves from Figure 5, we constrain our search space from $[\mathbf{-10}, \mathbf{10}]$, as this is enough to represent the relation between predicted profit in these two markets. We utilize an existing code base for BO [10] to search the space.

Our results show that the learned parameters yield an approximate equilibrium, achieving between 90-95% (0.91, 0.95, 0.95 for $\lambda_{\text{FIF}} = .0005, 0.1, 1$ respectively) of the profit of the best response strategy found by BO. In the learned approximate equilibrium, typically above 90% of orders are placed in the CALL market. Interestingly, the best response strategy found by BO always resulted in the deviating trader placing *every single order* in the call market. So we then asked whether all informed traders placing all their orders in the call market is an equilibrium, and found that, except under exceptional conditions, it is (that is, BO returned a set of parameters for the remaining trader that resulted in that trader placing all its orders in the CALL market as well). The only condition which we found under which it is not an equilibrium is when there is very little liquidity from FIF traders in the CDA market, but still some background traders – in this case the deviating informed trader can essentially become a price setter and trade with the background traders at whatever prices it chooses.

6. CONCLUSION

We have developed an agent-based model in the tradition of classic microstructure models to engage the question of whether frequent call markets can drive liquidity away from CDA markets. If they could do so, this would have the potential to increase welfare both by reducing transaction costs for average market participants and by reducing the incentive for firms to engage in the latency "arms race." Our results are promising. Even in the presence of impatient background traders who primarily demand immediacy and are willing to pay for it, we show in both a simple zero-intelligence model, and more sophisticated learning and equilibrium settings, that call markets have the potential to attract a large fraction of the order flow. In addition to the policy implications, we believe the modeling approach taken in this paper constitutes a useful bridge between classic financial microstructure models and more complex agent-based models, preserving intuition from the former, while allowing us to examine richer environments and questions.

Acknowledgments

We thank Elaine Wah and Michael Wellman for useful discussions. This research is supported by US National Science Foundation awards 1527037 and 1414452.

REFERENCES

[1] A. Brahma, M. Chakraborty, S. Das, A. Lavoie, and M. Magdon-Ismail. A Bayesian market maker. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 215–232, New York, NY, USA, 2012. ACM.

[2] J. Brogaard, T. Hendershott, and R. Riordan. High-frequency trading and price discovery. *Review of Financial Studies*, 27(8):2267–2306, 2014.

[3] E. B. Budish, P. Cramton, and J. J. Shim. The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, pages 1547–1621, 2015.

[4] D. Cliff, J. Bruten, et al. Zero not enough: On the lower limit of agent intelligence for continuous double auction markets. *HP Laboratories Technical Report HPL*, 1997.

[5] S. Das. The effects of market-making on price dynamics. In *Proceedings of the International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 887–894, Estoril, Portugal, May 2008.

[6] S. Das, J. P. Dickerson, Z. Li, and T. Sandholm. Competing dynamic matching markets. In *Proceedings of the Conference on Auctions, Market Mechanisms and Their Applications (AMMA)*, 2015. To appear.

[7] L. Epstein, A. Levin, J. Mestre, and D. Segev. Improved approximation guarantees for weighted matching in the semi-streaming model. *SIAM Journal on Discrete Mathematics*, 25(3):1251–1265, 2011.

[8] J. D. Farmer and D. Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.

[9] J. D. Farmer, P. Patelli, and I. I. Zovko. The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6):2254–2259, 2005.

[10] J. Gardner, M. Kusner, K. Q. Weinberger, J. Cunningham, and Z. Xu. Bayesian optimization with inequality constraints. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 937–945, 2014.

[11] L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.

[12] A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *working paper*, 2014.

[13] A. S. Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

[14] B. LeBaron, W. B. Arthur, and R. Palmer. Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, 23(9):1487–1516, 1999.

[15] M. Lewis. *Flash Boys: A Wall Street Revolt*. WW Norton & Company, 2014.

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

[17] J. Niu, K. Cai, S. Parsons, P. McBurney, and E. H. Gerding. What the 2007 TAC market design game tells us about effective auction mechanisms. *Autonomous Agents and Multi-Agent Systems*, 21(2):172–203, 2010.

[18] A. Othman. Zero-intelligence agents in prediction markets. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 879–886, 2008.

[19] I. Palit, S. Phelps, and W. L. Ng. Can a zero-intelligence plus model explain the stylized facts of financial time series data? In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 653–660. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[20] J. Snoek, H. Larochelle, and R. P. Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, pages 2951–2959, 2012.

[21] E. Wah, D. R. Hurd, and M. P. Wellman. Strategic market choice: Frequent call markets vs. continuous double auctions for fast and slow traders. In *Proceedings of the Conference on Auctions, Market Mechanisms and Their Applications (AMMA)*, 2015. To appear.

[22] E. Wah and M. P. Wellman. Latency arbitrage, market fragmentation, and efficiency: A two-market model. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, pages 855–872. ACM, 2013.

[23] E. Wah and M. P. Wellman. Welfare effects of market making in continuous double auctions. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 57–66. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[24] T. J. Walsh, I. Szita, C. Diuk, and M. L. Littman. Exploring compact reinforcement-learning representations with linear regression. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 591–598. AUAI Press, 2009.