Experiments in High-Frequency Trading: Comparing Two Market Institutions[*]

Eric M. Aldrich[†]
Department of Economics
University of California, Santa Cruz

Kristian López Vargas[‡]
Department of Economics
University of California, Santa Cruz

February 28, 2019

Abstract

We implement a laboratory financial market where traders can access costly technology that reduces communication latency with a remote exchange. In this environment, we conduct a market design study on high-frequency trading: we contrast the performance of the newly proposed Frequent Batch Auction (FBA) against the Continuous Double Auction (CDA), which organizes trades in most exchanges worldwide. Our evidence suggests that, relative to the CDA, the FBA exhibits (1) less predatory trading behavior, (2) lower investments in low-latency communication technology, (3) lower transaction costs, and (4) lower volatility in market spreads and liquidity. We also find that transitory shocks in the environment have substantially greater impact on market dynamics in the CDA than in the FBA.

Keywords: Market design, Auctions, High-Frequency Trading, Continuous Double Auction, Frequent Batch Auction.

JEL Classification: C91, D44, D47, D53, G12, G14.

1   Introduction

Telecommunications technology has transformed financial markets in the last decade. Order submission and execution times at major exchanges have declined from seconds to microseconds (1 millionth of a second). As a consequence, traders are rewarded for reacting quickly to information, resulting in a new market participant: high-frequency trading (HFT) firms. HFT firms use computerized strategies to transact large volumes in fractions of a second and now account for more than half of transactions at major exchanges worldwide. Proponents claim that HFT has improved market liquidity (the ease with which trades can occur) and reduced transaction costs (Narang, 2010). Opponents argue that the multi-billion-dollar cost of HFT infrastructure is ultimately borne by investors, its liquidity is illusory, and it is a destabilizing force in financial markets (Lewis, 2015).

Most academic papers studying HFT use proprietary data with trader identification, and are able to classify accounts as either aggressive (primarily liquidity consuming) or passive (primarily liquidity providing). Passive accounts are almost uniformly associated with improved market performance. Hagströmer and Nordén (2013); Malinova et al. (2014); Brogaard et al. (2015); Jovanovic and Menkveld (2015) and Menkveld and Zoican (2017) are examples of papers that find such positive effects. The effects of aggressive HFT, however, are mixed: it is generally associated with informed price impact over short horizons, increased adverse selection costs for other traders, increased short-term volatility, and higher trading costs for institutional and retail traders. Examples from this literature include Zhang and Riordan (2011); Bershova and Rakhlin (2012); Breckenfelder (2013); Brogaard et al. (2014); Hasbrouck and Saar (2013); Hendershott and Riordan (2013); Hirschey (2013); Baldauf and Mollner (2015b,a); Brogaard and Garriott (2015) and Menkveld and Zoican (2017).

Despite the uncertainty surrounding the effects of high-speed trading, policy makers worldwide are already taking actions intended to discourage HFT. For example, in 2016, the U.S. Securities and Exchange Commission approved the Investors Exchange (IEX) to operate as a public securities exchange. A primary goal of the IEX market, which was founded in 2012 to provide a non-public alternative trading system (ATS) with delayed messaging (see Aldrich and Friedman, 2017), is to reduce potential advantages of HFT firms. The SEC decision to approve the IEX system as a public exchange had substantial network effects, primarily because of quote protection under Reg NMS Rule 611 (Hu, 2018). In addition to IEX, several other market formats have been proposed as alternatives to the CDA. These include the frequent batch auction (Budish et al. (2015),

Budish et al. (2014)) and the fully continuous exchange (Kyle and Lee, 2017). However, no scientific evidence exists on the relative performance of such market institutions.

The objective of this paper is not to address the net costs or benefits of HFT, but instead to compare the effects of different trading environments on market quality in the presence of HFT. Specifically, we use laboratory experiments to compare two leading financial market formats in the presence of high-frequency trading: the continuous double auction (CDA) and the frequent batch auction (FBA). The CDA (also known as the continuous limit order book) organizes trade in the majority of equities, futures and currency exchanges around the world. In this format, traders make publicly committed offers to buy and sell assets and are able to accept others' offers at any moment of time. In addition, traditional CDA markets employ a price-time priority system, which first ranks orders by price and then by the time they are received at the exchange (within price level). Because trading is continuous in time and orders are ranked by their submission time, communication speed is crucial in the CDA. Traders who can quickly react to new information have a substantial advantage over slower traders. This generates competition for speed technology, which is potentially socially inefficient. The FBA, on the other hand, does not allow trading to occur continuously. Instead, bids and asks are collected (batched) over discrete time intervals and call auctions are conducted at the end of each batching period. FBA therefore gives equal priority to orders received within a batching period, reducing the advantage of fast communication technology. Competition for speed becomes much less relevant than in the CDA and competition on price regains primacy.

The environment for our laboratory experiments is taken from the theoretical model of Budish et al. (2015) (hereafter referred to as BCS), where a single asset is traded on a single exchange. Traders submit commitments to buy or sell the asset in the form of *limit orders* and *market orders*. Two exogenous processes generate incentives to trade in this environment: changes in the publicly-observed fundamental value of the asset and the arrival of market orders from noise traders (*investors*) at random times. Although the two market formats price transactions differently, both assume that purchases (sales) are simultaneously liquidated (purchased back) at the fundamental value, allowing traders to book profits or losses instantaneously.

Human participants (acting as traders) choose among three broad strategies to earn real profits: to act as *market makers*, to act as *snipers*, or to not participate in the market. Their choices can be continuously revised throughout the experiment. Market makers and snipers may subscribe to a technology that reduces messaging latency to the exchange for a pre-specified flow cost. Additionally, market makers choose a spread

around the asset value at which they post bids to buy (below the value) and offers to sell (above the value). Market makers earn profit when the exchange matches them with a counterparty. Snipers attempt to exploit temporarily mispriced maker orders (stale quotes) by transacting with them at the time of a jump in the asset value.

To emulate modern financial markets, we develop an electronic architecture in which information and trading occur at millisecond time granularity. Specifically, for each format we deploy remote exchanges in an Amazon data center and utilize the Nasdaq OUCH protocol[1] for messaging. Human subjects make high-level strategic decisions (outlined above) by interacting with a computer interface in the laboratory. Their decisions are encoded into algorithms which act on their behalf by communicating with the exchange as market conditions change. The arrival of exogenous information to the market via investor orders and changes in the fundamental value is designed to be representative of the time-scale in which similar information arrives in the market for a liquid asset, such as a the S&P 500 exchange traded fund. Thus, human subjects take the role of analysts at trading firms that design algorithmic strategies rather than the algorithms themselves.

The BCS model makes sharp equilibrium predictions regarding the roles that subjects choose and their decisions to purchase fast communication technology. Under the CDA, market liquidity is minimal in equilibrium since the predatory strategy (sniping) is chosen by all but one trader. In contrast, under the FBA, all traders choose to provide liquidity (i.e. to be a market maker) in equilibrium. The BCS model also predicts that in the CDA every trading firm adopts the costly speed technology, while in the FBA no trader purchases that technology.

Although traders make zero profits in the equilibria of both formats, social welfare is superior under the FBA because retail investors pay a positive market spread in the CDA while in the FBA the spread is zero. This welfare difference is associated with a prisoner's dilemma that arises under CDA – the dominant strategy is for all traders to purchase fast communication technology, despite the fact that all would be better off without it – and which disappears in the FBA, as the gains of purchasing speed technology are minimal.

To understand the relationship of observed behavior to changes in the underlying environment, and hence predicted equilibria, we consider three treatments that vary the rate at which the asset value changes, the rate at which investors arrive, and the cost of fast communication technology. Regardless of treatment parameters, we find

---

[1]The OUCH protocol is a specification for sending binary messages of fixed length. It is designed for efficient, low-bandwidth communication between clients (traders) and a centralize host (exchange). For more information, please refer to the OUCH 4.2 specification (Nasdaq, 2018).

that subjects in the FBA market display substantially more passive liquidity, engage in less predatory behavior, and are less likely to purchase fast communication technology. Further, all FBA markets exhibit lower transactions costs, greater price efficiency, and less volatility. These results are all highly statistically significant and directionally support the predicted equilibria, although their magnitudes are somewhat attenuated relative to the precise equilibrium predictions.

Our work is as much a test of the BCS equilibrium model as a comparison of the CDA and FBA market formats. As such, our results are only relevant to real-world markets to the extent that the BCS model is a good characterization of actual trading behavior in those markets. However, testing the behavioral robustness of the BCS environment is interesting and useful in its own right. First, it is not clear, a priori, that human subjects can adequately learn, or that their behavior can converge in such a complicated and highly stochastic environment. Our results demonstrate that this is possible and can lead to useful insights. Additionally, the predicted equilibria are stylized and somewhat implausible due to their invariance over a wide range of the parameter space. Understanding behavioral divergences from these predictions will assist in the development of new theoretical models.

This paper contributes to the market design literature in finance and experimental finance. Related prior research on financial market design includes Roth and Xing (1994), who study the timing of transactions, Roth and Xing (1997), who study serial versus batch processing, Foucault (1999) and Roth and Ockenfels (2002), who introduce the idea of bid sniping, Ariely et al. (2005) who study how Internet auctions' ending rules shape the incentives for bid sniping, Du and Zhu (2017) and Fricke and Gerig (2015), who study the optimal frequency of double auctions, and Biais et al. (2014), who study "fast trading" and the externalities it generates. Haas and Zoican (2016) present a theoretical study of the impact of batch length on liquidity using an extension of the BCS model. Webb et al. (2007) empirically study the impact of the frequency of market clearing on volatility.

More broadly, this paper contributes to the literature that contrasts the performance of market institutions which has a long tradition in experimental economics.[2] Many of

---

[2]In a pioneer study, Smith (1962) finds that the CDA leads to competitive equilibrium outcomes within a short period of time and with only a few buyers and sellers. Plott and Smith (1978) compare two market institutions, the posted-offer and the CDA, and find that the latter is more efficient. Subsequently, a long literature in experimental markets has explored a diversity of questions, including information aggregation (e.g. Plott and Sunder, 1982, 1988; Forsythe and Lundholm, 1990; Copeland and Friedman, 1987, 1991; Theissen, 2000), convergence to rational expectations equilibrium (e.g. Forsythe et al., 1982, 1984; Friedman et al., 1984; Anderson et al., 1991; Morone and Nuzzo, 2016a), asset price bubbles (e.g. Smith et al., 1988; Van Boening et al., 1993; Noussair and Tucker, 2006), and information mirages (see e.g. Camerer and Weigelt, 1991; Morone and Nuzzo, 2016b).

these studies allow for comparisons between versions of continuous double auction and repeated call markets (FBA, in our paper). As reported in Friedman and Rust (1993), early laboratory experiments comparing the efficiency of the CDA to other formats (e.g., bilateral search, posted offer, and one-sided auctions) yielded results that were favorable to the CDA (Plott and Smith, 1978; Williams, 1980). Smith et al. (1982) compares the CDA with call markets and finds that CDA results in faster equilibrium convergence and more efficient outcomes.[3] In contrast, Friedman (1993) reports that a continuous-information version of a call market exhibits efficiency levels (informational and allocative) comparable to the CDA. Additionally, transaction prices in the call market are closer to the equilibrium price and market spreads are narrower. McCabe et al. (1992) also finds that the call auction exhibits smaller price variation than the CDA. Cason and Friedman (1996) compare several market formats including a repeated call market, CDA, a uniform-price CDA and a single call market. They find that FBA and CDA exhibit similar levels of trading and informational efficiency.[4] Van Boening et al. (1993) finds that bubbles and crashes are observed in both the CDA and call auctions and Schnitzlein (1996) reports that call auctions are equally efficient as CDA markets and outperform the CDA in market liquidity. Theissen (2000) finds that opening prices in a call market track the fundamental value of an asset better than the those of the CDA. However, the call market also underreacts to the arrival of new information. Using a model with randomly drawn redemption values, Kagel (1998) finds that both CDA and call markets display high efficiency levels and non-important price deviations from equilibrium level, but the CDA mechanism generates slightly higher allocative efficiency. Huber et al. (2008) study whether individuals with better information about the fundamental value earn more than those with less information in CDA and call markets. They report that, in both formats, only traders with significantly more information than others can achieve better payoffs. Morone and Nuzzo (2016b) study a market in which traders are uncertain about the presence of insiders. They find that information mirages occur more frequently under CDA relative to the call auction. Without market insiders, the call auction exhibits faster convergence to equilibrium and more stable prices. In the presence of market insiders, both formats exhibit similar informational efficiency performance.[5]

We highlight that existing research on financial auctions focuses on environments

---

[3] In this experiment, CDA also outperforms the simple implementation of call auction in allocative efficiency.

[4] Similarly, other studies comparing the CDA and FBA find that both formats are highly efficient in markets of perishables and independent private values (e.g Cason and Friedman, 1997, 1999, 2008).

[5] For surveys on experimental research in financial markets, see Friedman and Rust (1993), Holt (1995), Sunder (1995), Friedman (2008), Noussair and Tucker (2013) and Nuzzo and Morone (2017).

that are not specifically relevant to the study of HFTs. To our knowledge, this paper reports the first experimental study of the CDA and FBA in the presence of HFT.

The rest of the paper is organized as follows. Section 2 presents the experimental design and its implementation, Section 3 reports results of the experiments and Section 4 concludes. We collect discussions and proofs related to model calibration, off-equilibrium behavior and equilibrium market statistics in online Appendices A – D. Instructions for the experiments are included in online Appendix E.

2   Experiment Design

We begin this section by describing the theoretical environment of the experiments and also providing detail on the two market formats that organize trade among participants. We then discuss the architecture of software and hardware used within the laboratory and for the remote exchange server. We conclude the section by describing implementation details related to treatments and the procedures of the experimental sessions.

2.1   Environment

Our laboratory environment is adapted from Budish et al. (2015) (BCS), where a single asset is traded on a single exchange. Traders express their willingness and commitment to buy or sell the asset by transmitting *limit orders* to the exchange. A limit order is a message comprised of four basic elements: (a) direction: buy (sometimes called a bid) or sell (sometimes called an ask or offer), (b) limit quantity (maximum number of units to buy or sell), (c) limit price (highest acceptable price for a bid, lowest acceptable price for an offer), and (d) time in force (indicating when the order should be canceled). A *market order* is a specialized limit order with the highest (lowest) possible limit price if it is a bid (offer). As a result, market orders transact immediately with any standing liquidity that expresses an opposing interest to buy or sell.

Two exogenous processes generate incentives to trade in this environment: (1) the fundamental value of the asset, $V(t)$, which is publicly observed and follows a compound Poisson process with arrival rate λ_V per second and jump distribution F_V and (2) a population of *investors* (noise traders) that arrive at random times with Poisson rate λ_I per second, placing unit market orders to buy and sell with equal probability. Since investors exclusively use market orders, they transact immediately as long as there is

7

contra-side interest. As described below, different market formats price transactions differently, but it is assumed that any purchase (sale) of the asset at any time is simultaneously liquidated (purchased) at the fundamental value, allowing traders to book profits or losses instantaneously.

The focus of the study is the behavior of N trading firms under differing market formats, and the outcomes this behavior generates. Human participants play the role of *trading firms* and, at any instant, can choose whether (a) to exit the market (*out*), or to participate either as (b) a *market maker* or (c) a *sniper*. In the latter two cases, traders also choose whether to invest in a technology that reduces round-trip messaging latency to the exchange from δ_{slow} to $\delta_{fast} < \delta_{slow}$ at a cost c_s per second. All traders are constrained to transact in unit shares of the asset.

Market makers are required to symmetrically post a buy order (*bid*) and sell order (*offer*) around the fundamental, $V(t)$. In practice, a maker chooses a spread, $s_i(t)$, which sets the price of her bid and offer to be, respectively, $V(t) - 0.5s_i(t)$ and $V(t) + 0.5s_i(t)$. Market makers earn profit when the exchange matches them with incoming investors, with the likelihood and magnitude of such transactions depending on the allocation mechanism of the market format.

When the fundamental value jumps from $V(t^-)$ to $V(t)$, a maker's orders will be temporarily mispriced at $V(t^-) \pm 0.5s$, where we assume that her choice of spread is fixed at $s_i(t) = s$. Immediately after the maker's algorithm learns about the new value $V(t)$, it submits a message to update her orders to $V(t) \pm 0.5s$. The update from $V(t^-) \pm 0.5s$ to $V(t) \pm 0.5s$ occurs at $t + \delta_{slow}$ by default, or earlier at $t + \delta_{fast}$, if the maker subscribes to the fast communication technology at a price of c_s per second.

Snipers attempt to exploit stale quotes at the time of a jump in the fundamental value. When a value jump results in temporarily mispriced (stale) makers' orders, a sniper will attempt to transact with one of those stale orders to make a profit. Specifically, when the value jumps up (down), snipers use market orders to quickly buy from (sell to) makers who have not yet updated their offers (bids). This is only possible when the jump is large enough and the market format allows it. As with makers, snipers' algorithms receive price information from the exchange and submit orders with default latency δ_{slow}, but can reduce their communication latency to δ_{fast} by paying c_s per second.

2.2 Market Formats

We now describe how the market formats separately handle limit orders that are transmitted by traders.

2.2.1 Continuous Double Auction

Most modern financial exchanges implement a variant of the continuous double auction (CDA), also known as the continuous limit order book. This format is characterized by a *limit order book* that sorts limit orders by (1) price and (2) time received (at each price). Bids are sorted from highest to lowest price and offers are sorted from lowest to highest price. The highest bid and the lowest offer are referred to, respectively, as the *best bid* and *best offer*, and the difference between them is called the *spread*.

Traders (whether human or automated) enter, replace, modify and cancel orders at any moment they choose and the CDA processes each limit order as it arrives. If the limit price locks (equals) or crosses (is beyond) the best contra-side price – e.g., if a new bid arrives with limit price equal to or higher than the current best offer – then the limit order immediately transacts ("executes" or "fills") at that best contra-side price, and the transacted quantity is removed from the order book. On the other hand, if the price is no better than the current best same-side price, then the new order is added to the order book, behind other orders at the same price. The exchange breaks ties by randomly ordering messages received at the same price and time. Time priority in the CDA favors traders that can send, modify and cancel limit orders quickly, which results in all traders acquiring fast communication technology in the BCS equilibrium.

In this framework, market makers balance the profit generated by trading with investors (only if they post the best bid and offer) with the potential cost of being sniped at the time of value jumps and with the potential cost of purchasing fast communication technology. Under the CDA, BCS investors' sell orders transact with the highest maker bid and buy orders transact with the lowest maker offer.[6] As a result, the maker with smallest spread, s, books profit $0.5s$ at a rate dictated by λ_I. However, at the moment of a sufficiently large positive (negative) change in the fundamental value, $J \equiv |\Delta V(t)| = |V(t) - V(t^-)| > 0.5s$, snipers will attempt to trade with the lowest (highest) maker offer (bid). A successful sniper earns profit $\Delta V(t) - 0.5s$ and the maker with smallest spread, s, takes a loss of the same size. The relative speed of other traders in the market dictates the probability that a maker is sniped when the fundamental

---

[6]Noise traders (investors) only submit market orders.

value changes, as the exchange breaks ties by randomly ordering messages received at the same time. The equilibrium probability of such events is described below.

Snipers, on the other hand, balance the profits from sniping with the potential cost of fast communication technology. As with makers, their probability of earning profits is closely related to the relative speed of other traders in the market.

Given N trading firms participating in the market, Budish et al. (2015) show that the equilibrium of their model under the CDA format consists of a single market maker, $N-1$ snipers, and all traders purchasing fast communication technology. Their equilibrium is characterized by two zero-profit conditions. For the maker,

$$\lambda_I \cdot \frac{s}{2} - \lambda_V \cdot \Pr\left(J > \frac{s}{2}\right) \cdot E\left[J - \frac{s}{2} | J > \frac{s}{2}\right] \cdot \frac{N-1}{N} = c_s, \tag{1}$$

which states that the expected profits from trading with investors, $\lambda_I \cdot \frac{s}{2}$, less the expected losses to snipers, $\lambda_V \cdot \Pr\left(J > \frac{s}{2}\right) \cdot E\left[J - \frac{s}{2} | J > \frac{s}{2}\right] \cdot \frac{N-1}{N}$, must be equal to the cost of buying speed services. For the sniper,

$$\lambda_V \cdot \Pr\left(J > \frac{s}{2}\right) \cdot E\left[J - \frac{s}{2} | J > \frac{s}{2}\right] \cdot \frac{1}{N} = c_s, \tag{2}$$

which says that the profits from sniping, $\lambda_V \cdot \Pr\left(J > \frac{s}{2}\right) \cdot E\left[J - \frac{s}{2} | J > \frac{s}{2}\right] \cdot \frac{1}{N}$, must be equal to expenditure on speed services. Note that sniping losses and profits are weighted by different fractions of N in Equations (1) and (2); this is a result of the fact that all traders are subject to identical communication latency δ_{fast}, causing the messages of all N players (the single maker and $N-1$ snipers) to be received by the exchange at the same time. With probability $\frac{N-1}{N}$, the maker loses $\lambda_V \cdot \Pr\left(J > \frac{s}{2}\right) \cdot E\left[J - \frac{s}{2} | J > \frac{s}{2}\right]$ to one of the snipers and with probability $\frac{1}{N}$ one of the snipers will earn the same amount.

Equations (1) and (2) endogenously determine both the maker's equilibrium spread, s^*, and the total number of trading firms, N^*. Budish et al. (2015) re-expresses the equilibrium as

$$\lambda_I \cdot \frac{s^*}{2} = \lambda_V \cdot \Pr\left(J > \frac{s^*}{2}\right) \cdot E\left[J - \frac{s^*}{2} | J > \frac{s^*}{2}\right] \tag{3}$$

$$\lambda_I \cdot \frac{s^*}{2} = N^* c_s. \tag{4}$$

Equation (3), which is the difference of Equations (1) and (2), determines s^* and Equation (4), which is the sum of Equation (1) and $N - 1$ times Equation (2), determines N^* for a given s^*. Budish et al. (2015) interpret Equation (4) as showing that the cost of speed, purchased by all traders, is borne entirely by investors via transactions with market makers.

All traders purchase fast communication technology in the CDA equilibrium to either prevent severe loss (maker) or to prevent exclusion from profits (snipers). Abstaining from fast communication is not a profitable deviation: a slow maker increases his chances of getting sniped by $1/N$ and a slow sniper reduces her sniping chances to zero. [7]

2.2.2 Frequent Batch Auction

In a Frequent Batch Auction (FBA), the trading day is divided into submission stages of equal length τ. These submission stages are referred to as *batching intervals* or *batches* and can be considered discrete time increments. Within a batch, traders communicate with the exchange in continuous, sealed-bid fashion in order to privately submit, modify and cancel limit orders, which are collected and held by the FBA matching engine.

At the conclusion of a batch, all orders are combined with unfilled orders from previous batches and the matching engine generates stair-step demand and supply curves from the aggregated bids and offers, respectively. If demand and supply do not intersect, no trade occurs and all orders carry over to the next batch, aside from those denoted as *immediate or cancel*. If demand and supply intersect, the market clears where supply equals demand, i.e., all infra-marginal bids and offers are executed at a uniform price p^* that clears the market. The FBA matching engine then publicly broadcasts information regarding executed trades, p^*, and the remaining order book.

Figure 1 diagrams the central features of an FBA batch. For a batch concluding at time t, traders subject to communication latency δ_i are not able to act on new information during time interval $(t - \delta_i, t]$ before the time t auction. For example, if the fundamental value of the asset changes in interval $(t - \delta_{slow}, t]$, slow makers will not have an opportunity to update their bids and offers and slow snipers will not have an opportunity to submit aggressive market orders. The same is true of fast

---

[7] Budish et al. (2015) focus entirely on static or instantaneous Nash Equilibrium. It is possible that an equilibrium with cooperation is sustainable. Although we do not attempt to derive such an equilibrium, we calculate profits under cooperation in our experiments in online Appendix C and also argue that the theoretical and empirical literature suggests that such an equilibrium is unlikely to be sustainable (due to the number of players interacting in our markets). We thank the editor for encouraging us to discuss the possibility of collusive play.
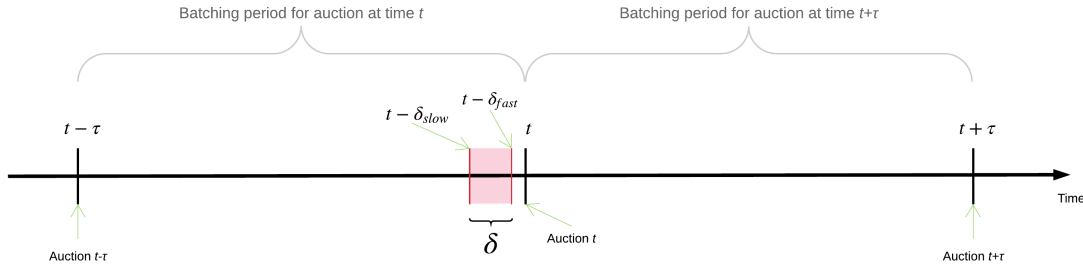
Figure 1: Timing in the FBA format (adapted from Budish et al. (2015)).

makers and snipers for asset value changes in the interval $(t - \delta_{fast}, t]$. The difference $\delta = \delta_{fast} - \delta_{slow}$, in relation to the total batch length τ, represents the relative advantage of fast communication technology: as the ratio δ/τ declines, one expects the value of speed technology to diminish.

Under certain conditions, the BCS equilibrium for FBA consists of all traders choosing to act as market makers with zero spread and none of them purchasing fast communication technology. Specifically, Budish et al. (2015) show that if:

$$\frac{\delta}{\tau} \cdot \lambda_V E\left[J | J > 0\right] < c_s \qquad (5)$$

there is no incentive for traders to take the role of a fast sniper in the FBA. Intuitively, the left-hand-side of Equation (5) is the profit attributed to a fast sniper who successfully exploits a unit share posted by a slow market maker when a jump occurs in the interval δ: $\frac{\delta}{\tau} \cdot \lambda_V$ represents the probability of a value jump in the short interval δ and $E\left[J | J > 0\right]$ represents the magnitude of the gain (when $s^* = 0$). Since there are no fast snipers in the market, the need for high-speed communication technology vanishes. Competition focuses purely among makers on price (à la Bertrand) and dictates that all market makers undercut each other until $s^* = 0$ (all bids and offers are set to equal the observed fundamental value, V).

It is important to note that, as in the CDA, makers earn profits in the FBA by transacting with investors. However, since investors submit the best bids and offers within a batching period, the FBA auction often pairs their orders together as transactions. As a result, makers only transact with investors when the number of buying investors within a batch differs from the number of selling investors. Such transactions occur at the market clearing price, p^*, which is weakly lower (higher) than the best bid

(offer). Thus, for a given set of player strategies in the market, there are two reinforcing channels by which makers earn lower transaction profits in the FBA relative to the CDA: (1) fewer investor transactions and (2) transactions prices that are closer to the fundamental value of the asset. This reduction in profits is partially offset by smaller losses to sniping.[8]

2.3 Architecture

The experimental architecture consists of two components: (1) a laboratory software interface that implements the BCS model and which is specific to each of the market formats (CDA and FBA) and (2) a remote exchange that is specific to each of the formats.

The laboratory software is a high-level computer interface that allows human subjects to continuously tune algorithmic trading strategies during a trading session. As market conditions evolve, these algorithms interact with the remote exchange server at millisecond time granularity.

Communication latency and the speed with which traders realize their strategies is a central feature of the BCS equilibrium. However, these latencies are technological constraints not related to human physiological limitations. Consequently, in our experiment, the speed at which subjects trade should not depend on their human-reaction time, but solely on the instructions they provide to their algorithms and whether or not they use a costly technology that reduces messaging latency between their algorithms and the exchange.

2.3.1 Laboratory Interface

The laboratory component of the architecture is implemented using the Redwood 2 platform. Redwood 2 is a Django-based experiment platform designed for games in continuous time.[9] The laboratory layer is comprised of the following modules:

1. A graphical user interface that displays payoffs, recent history and current state of the market, and trader controls.

---

[8]Similar to the CDA case, there is the possibility of cooperative equilibria in the FBA. In the online Appendix C, we argue that collusive play is unlikely due to the number or players in each market.

[9]To our knowledge, Redwood 2 the only open-source, browser-based experimental platform that can support all three: continuous-time games, arbitrary graphical interfaces and connection to external market engines as required by our architecture (https://github.com/RedwoodAdmin/RedwoodFramework).

2. Individual algorithms (one per subject) that keep track of subjects' choice states and implement their strategies by automatically sending messages to the remote exchange (using Nasdaq's OUCH 4.2 protocol) as market conditions change.

3. A market manager that handles OUCH messages composed by subjects' algorithms and passes them on to the remote exchange. In return, the market manager receives confirmation messages from the exchange, which are forwarded to individual algorithms. The market manager also implements value jumps and investor arrivals from prespecified draws of the Poisson processes describe above.

The user interface for the CDA format is shown in panel (a) of Figure 2. The *info box* (top) displays basic information about the market state such as the current bid-offer spread, the number of active traders and the cost of speed c_s. Traders use the *choice box* (bottom right) to adjust choices at any moment during the trading period. To (re)enter as a market maker, a subject clicks the "Maker" button, or selects a spread (s_i) by clicking in the white area of the *choice box*. To (re)enter as a sniper, a subject clicks the "Snipe" button. By clicking the "Out" button, a subject cancels any limit orders, deactivates any algorithms trading on her behalf, and unsubscribes from speed services. The *profit box* (bottom left) displays subjects' accumulated profit at each moment of time. A subject's profit jumps if there is an investor arrival that executes against her buy or sell orders, if she snipes, or if she gets sniped.

The interface for the FBA format is essentially the same as in the CDA and is shown in panel (b) of Figure 2. The FBA choice box includes a time marker that indicates the length and the moment of the auction and the profit box includes markers that indicate the instant of each auction. In addition, the FBA info box displays aggregated information on the roles and speed choices of other traders in the market. This difference from the CDA info box was a programming oversight (they were intended to be identical), but does not constitute a large discrepancy as the information is directly (in the case of number of makers) or indirectly (number of snipers and speed choices) observable from the choice and profit boxes. Further, the interim summary screens in both formats provided the same aggregated information between each four-minute trading period.

2.3.2 Remote Exchange

Each of the market formats we study utilizes a general purpose market engine that acts as a remote exchange. The remote exchanges are patterned after the architecture
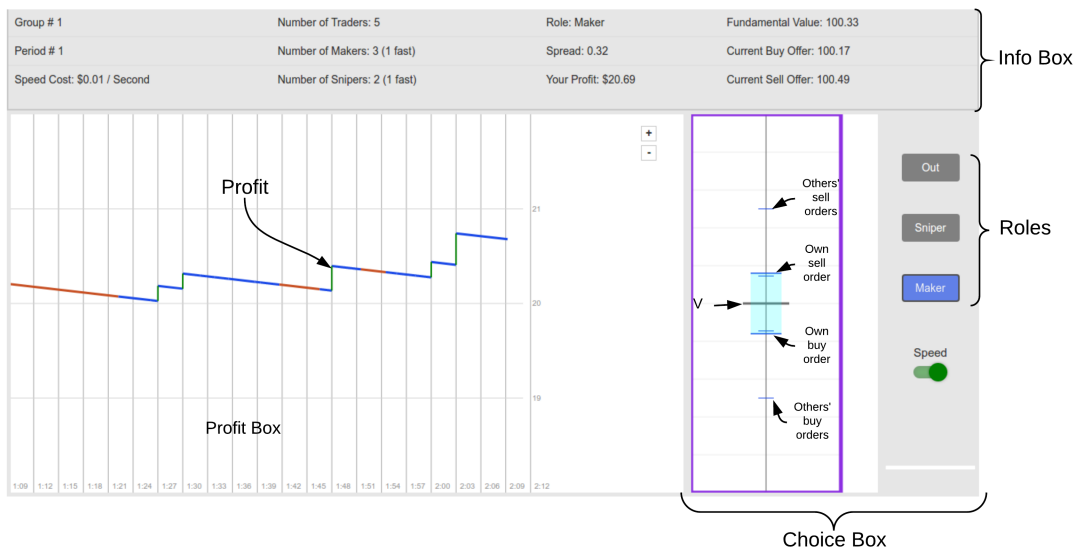
14

(a) CDA
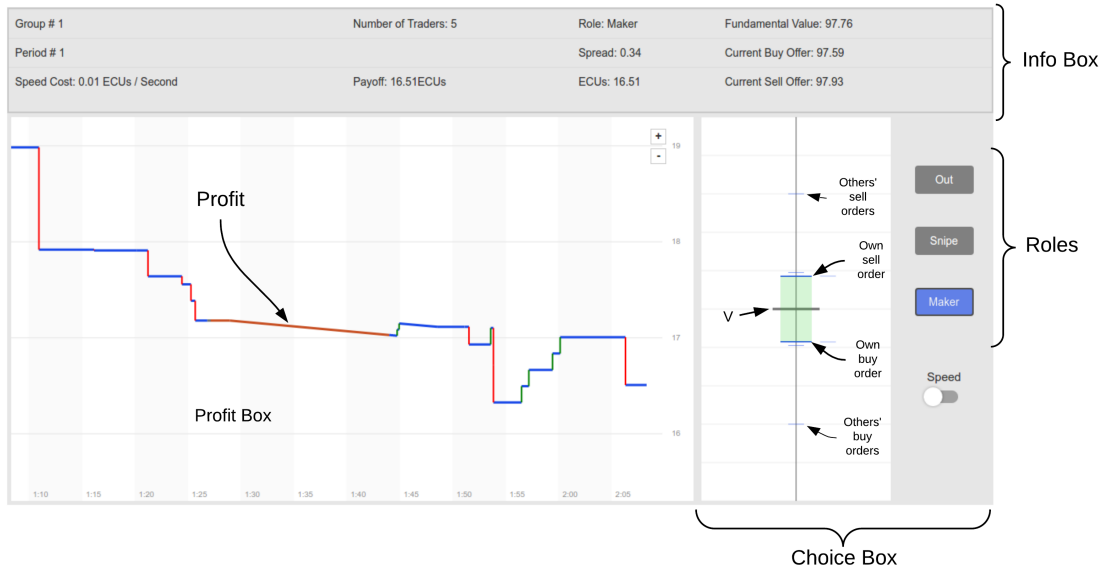
(b) FBA

Figure 2: Experimental user interfaces for CDA (panel a) and FBA (panel b).

15

of exchanges under the Nasdaq OMX umbrella and were developed with the goal of reusing the software infrastructure in future experiments and allowing other researchers the same access.

The remote exchanges are each comprised of two components: a *messaging server* and a *matching engine*. The messaging server follows the Nasdaq OUCH 4.2 specification, which strictly defines a set of incoming and outgoing private messages and their format (the order and byte length of each field within a message). The most common incoming messages (from participant to exchange) sent via OUCH are new order submissions, as well as updates and cancellations of existing orders. Outgoing OUCH messages (from exchange to participants) consist of private confirmations of incoming messages. In addition, the messaging server follows the Nasdaq ITCH 4.1 protocol for public broadcasts of market information.

The matching engine processes orders relayed from the messaging server, executes transactions according to a specified format (see section 2.2), and reports the results to the messaging server. In both formats we consider, the matching engine maintains a limit order book – the set of unexecuted, actionable limit orders in the market at a given point in time. Both components are hosted on an Amazon's Elastic Compute Cloud (EC2) infrastructure located in California, enabling tight control over physical accessibility and communication latencies.[10]

2.4   Treatment Design

A treatment in our experiment consists of a combination of *market format* and a *market configuration*. We study two market formats (CDA and FBA) and three market configurations. In our baseline configuration (C1) we set $\lambda_I = 1/3$, $\lambda_V = 1/4$ and $c = 0.01$. As argued in online Appendix A, these values are suggestive of actual dynamics observed in the S&P 500 exchange traded fund. Substituting these values in the BCS equilibrium Equations (3) and (4), results in $s^* = 0.324$ and $N^* = 5.4$ under CDA. Under FBA, $s^* = 0$ and $N^* = 6$, since everyone enters the market in that equilibrium and since there are a total of 6 traders in each market (group).

For the remaining configurations, we consider variations in the relative jump and

---

[10]While there were some alternative options to implement our remote exchange (such as Flex-E-Markets), we opted for developing our own exchange server for two reasons. First, we wanted our platform to be open source so it can be available to other researches for their use and contribution. Second, this exchange was built to exhibit negligible latencies at the messaging and matching engines level. In this way, the laboratory interface can be used to control and implement the menu of different communication technologies (each with its own latency) that is made available to traders.

investor arrival intensities, as well as the cost of speed, in order to induce larger equilibrium spreads while maintaining the same endogenous population of traders, $N^* \approx 5.4$. In configuration 2 (C2), we use $\lambda_I = 1/5$, $\lambda_V = 1$ and $c = 0.0104$. This could be loosely interpreted as a volatile market, with many asset value changes (once per second) and few investor arrivals. In this case the CDA equilibrium spread is much higher: $s^* = 0.566$. Finally, in configuration 3 (C3) we consider $\lambda_I = 1/2$, $\lambda_V = 1$ and $c = 0.022$, which doubles the cost of speed technology while maintaining the rapidly moving market, but with relatively more investor arrivals. Under this parameterization, the CDA equilibrium spread is $s^* = 0.475$. As with C1, $s^* = 0$ and $N = 6$ under FBA for C2 and C3. In all configurations, we use a Gaussian jump distribution, $\Delta V(t) \sim \mathcal{N}(0, \sigma = 0.5)$, and the initial value for V is set at 100 $ECUs$. Panel (a) of Table 2 shows the parameters for each market configuration. Equilibrium predictions for these configuration parameters are shown in Panel (b) of the same table.

In addition to the parameters above, we set $\delta_{slow} = 0.5$ seconds, $\delta_{fast} = 0.1$ seconds for both CDA and FBA in all configurations. We fix the length of a batching period to be $\tau = 3$ seconds, which means that the fraction of time that snipers can exploit slow makers under FBA is $\frac{\delta_{slow} - \delta_{fast}}{\tau} = \frac{0.4}{3} = 0.133$.

We used a between-subjects design to collect data from two sessions for each of the six treatments $\{CDA, FBA\} \times \{C1, C2, C3\}$. Each session consisted of 12 traders, divided into two independent groups (markets) of six participants and a single treatment was implemented in a session. A session consisted of eight consecutive trading periods of four minutes, with fixed groups across all periods (partner matching). To obtain data for 24 markets (6 treatments $\times$ 4 groups), we conducted a total of 12 sessions. Within a trading period, participants were able to change strategies at any moment. However, as highlighted above, the study and architecture were designed to test traders' strategic behavior rather than human reaction times.

Traders were informed in the instructions about the potential size of the market (six traders), latencies (δ_{slow}, δ_{fast}), arrival rates (λ_I, λ_V) and the cost of speed (c). See online Appendix E for experiment instructions.

2.5   Procedures

Sessions were conducted at the LEEPS Laboratory of the University of California, Santa Cruz. Recruitment was implemented through LEEPS' ORSEE instance (`econlab.`

ucsc.edu; Greiner, 2015). Instructions were provided on the computer screen.[11] After 15 to 20 minutes of reading instructions, a trial round of 90 seconds was launched during which subjects tested the interface with the understanding that their actions had no consequences for their earnings and would not comprise part of the formal experiment. Subjects were then given the option to publicly ask questions to the experimenters. The remainder of the session consisted of eight trading periods of four minutes. In between trading periods, subjects received a summary screen displaying the distribution and profits of different roles in the corresponding trading period. Each participant received 20 experimental currency units (ECUs) as initial endowment at the beginning of each trading period and final payments were based on the final wealth of a randomly chosen period, using an exchange rate of two ECUs per dollar, plus a $7 participation fee. The information was provided to subjects in the written instructions.

Although ending a trading period with negative profits was possible (i.e., losing 20 ECUs in the period), this happened only in 1.3% of the cases and subjects never left the laboratory with less than $7. Members of a group (market) were not able to communicate with each other before, during or after the experiment, nor did they learn the identity or characteristics of other participants. Payments were implemented following standard procedures of confidentiality.

3   Results

We compare the CDA and FBA market institutions with the following metrics: (1) market liquidity, measured by the fraction of traders who choose to be *market makers*, which is a surrogate for the depth of the order book in this environment; (2) the prevalence of predatory behavior, measured by the fraction of traders who choose to be *snipers*; (3) the penetration of high-speed communication technology in the market, measured by the fraction of traders that choose to subscribe to this service; and (4) transaction costs for investors, measured by the minimum spread among market makers. We also study other auxiliary metrics associated with volatility (e.g., the standard deviation of the minimum spread) and informational efficiency (such as the root mean squared deviation between realized prices and fundamental value). Measures of allocative efficiency are less relevant for this paper because we are dealing with a common value asset.

---

[11]Instructions were extensively discussed and piloted (with students and colleagues) to achieve a balance between clarity and reasonable length.

Prior to reporting our experimental results, several observations are in order. First, one of the CDA sessions (configuration 3) was terminated after 7 (rather than 8) periods because the 1.5-hour session time limit was reached. This affected two of the four markets for that configuration. Second, in certain circumstances, order repricing for market makers resulted in maker-to-maker sniping events. Online Appendix B describes these transactions in more detail and shows that they do not affect the BCS equilibrium. We discovered maker-to-maker snipes during initial pilot sessions and decided not to eliminate them, since doing so would violate typical order book rules. However, we did not explicitly acknowledge this feature in the experiment instructions, instead allowing subjects to observe the consequences of these events on their profit and decision screens. Finally, after collecting our experimental data, we discovered that the FBA exchange did not randomize transaction allocation for orders received within the same batch and at the same price. Under FBA, time priority should not exist for orders received by the exchange during the same batching period; such priority should only exist for orders received in prior batches and which have not been filled. Although within-batch time priority is a violation of the FBA matching mechanism, it did not have a material impact on market outcomes in our experiment for two reasons. First, market maker protection from snipers in the FBA is primarily attributed to a batch length that is substantially longer than the latency differential between fast and slow traders, and not from the random priority of orders at the same price within a batch. Second, order priority affected a very small fraction of orders (only those with identical prices within a single batch) and, at worst, provided an increased incentive to purchase fast communication services, which is not the equilibrium strategy and which was not the prevalent behavior in our experiments. Revising the system to allow for randomization of orders during auctions would potentially yield results that are even more congruent with predicted equilibria.

3.1 Subject Choices

To quantify treatment effects on choices made by subjects, we estimate the following model:

$$y_{g,t} = \sum_{j=1}^{3} [\alpha_j Cj_{g,t} + \gamma_j Cj \times FBA_{g,t}] + \epsilon_{g,t}, \tag{6}$$

	(1) Maker (%)	(2) Sniper (%)	(3) Speed (%)	(4) Min. Spread	(5) RMSD
Configuration 1	54.13***	30.89***	56.12***	0.226***	0.347***
	(1.640)	(2.959)	(2.972)	(0.0438)	(0.00474)
Configuration 2	30.27***	58.11***	69.20***	0.678***	0.512***
	(4.507)	(3.294)	(1.594)	(0.114)	(0.00918)
Configuration 3	40.03***	49.51***	69.02***	0.706***	0.460***
	(2.368)	(2.888)	(4.498)	(0.0374)	(0.0170)
FBA × Config. 1	24.88***	-10.95*	-36.15***	-0.123**	-0.135***
	(5.022)	(5.507)	(5.399)	(0.0439)	(0.0119)
FBA × Config. 2	49.08***	-44.23***	-37.82***	-0.502***	-0.102***
	(6.494)	(6.214)	(3.968)	(0.118)	(0.0307)
FBA × Config. 3	33.10***	-35.55***	-48.37***	-0.560***	-0.0791***
	(5.255)	(4.309)	(5.024)	(0.0422)	(0.0246)
Observations	10934	10934	10934	10934	142

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1: OLS estimates for Equation (6). The dependent variables in models 1–3 are, respectively, the average percentage of traders choosing to be makers, snipers, and subscribers to speed services. The dependent variables in models 4 and 5 are, respectively, the average minimum spread and the market-level RMSD. The regressions for all dependent variables except RMSD are conducted with data at the three-second resolution. We exclude the first ten seconds of each period and the first two periods of each session to account for starting effects. Group-level clustered standard errors are reported in parentheses.

where $y_{g,t} \in \{Maker_{g,t}, Sniper_{g,t}, Speed_{g,t}, MinSpread_{g,t}\}$ is indexed by group and time, Cj is a dummy variable for market configuration $j \in \{1, 2, 3\}$, and $Cj \times FBA_{g,t}$ is the dummy variable indicating the interaction between configuration j and the FBA format. In this context, since there is no constant in the regression, α_j captures the mean of y in CDA configuration j, and γ_j captures the FBA - CDA mean difference for y in configuration j. We estimate this model using a three-second time resolution, excluding the first ten seconds of each period to account for starting effects, and the first two periods of each session to account for possible learning. The estimation was done by combining data for all formats and configurations. The coefficients are OLS estimates and standard errors are cluster-robust at the group level. The regression results are reported in Table 1.

Result 1: More subjects choose to be market makers under FBA.

Model 1 of Table 1 shows that the fraction of market makers is between 25% and 50% higher under FBA than CDA, with higher values associated with configurations 2 and 3. Recall that relative to configuration 1, the market under configuration 2 experiences jumps in the fundamental value roughly four times more often, has 40% fewer investors, and has nearly identical cost of speed. Configuration 3, on the other hand, also has four times as many jumps in the fundamental value, but receives 1.5 times the number of investors, and has double the cost of speed. Thus, both configurations 2 and 3 can be regarded as volatile markets, which is costly to market makers, but configuration 3 partly compensates makers with additional income through investors, whereas configuration 2 provides even lower investor income, relative to the baseline calibration. In addition, under configuration 3, all subjects bear a higher cost of speed (makers and snipers). Thus, the results in Table 1 suggest that increased market making under FBA is positively related to volatile environments and fewer investors.

Since market makers are constrained to limit orders of unit size, the fraction of subjects acting as market makers is a direct measure of liquidity or market depth. Hence, the regression results demonstrate that FBA has a positive effect on liquidity that is statistically significant. Moreover, separate tests of the hypotheses $\gamma_2 > \gamma_1$ and $\gamma_3 = \gamma_1$ cannot be rejected at the 1% level, suggesting the liquidity improving effect of the FBA is highest in the most volatile regime with few investor arrivals.

The first row of plots in Figure 3 depict time series of the fraction of market makers, averaged across all groups within a given treatment. The faint, solid lines represent average values computed at five-second intervals and bold, solid lines represent exponentially-weighted moving averages with ten-second half lives. Dashed lines represent equilibrium values under the BCS model. Each panel of the plot depicts the full time series for a session, concatenating the metrics across the eight four-minute periods, which are visually separated by alternating gray and white vertical bands. To smooth start-up effects within periods, in which players' default state is "Out", we do not include the first ten seconds of each period in the plot.

The blue lines in Figure 3 correspond to equilibrium and observed values under CDA and the green lines correspond to those of the FBA. Under BCS, the equilibrium number of makers in the CDA is predicted to be $\frac{1}{N} * 100 = 100/6 = 16.67\%$ for all configurations, whereas the FBA equilibrium is 100% for all configurations. In magnitude, the fraction of makers fails to achieve the equilibrium values, but the time series show a separation in the direction of the predicted equilibria: for all configurations the fraction of makers is higher under FBA, and the magnitude of the difference is larger under the volatile

21
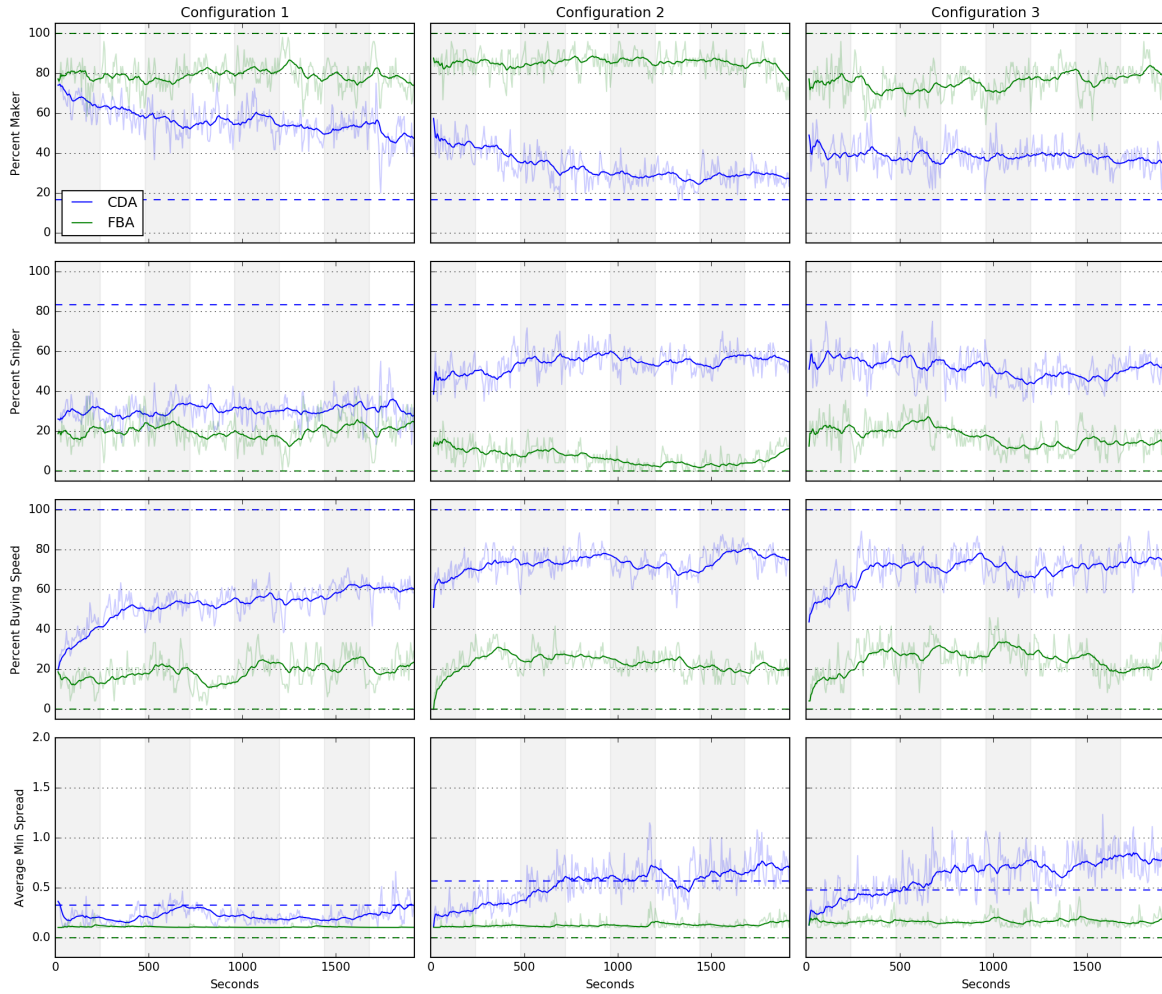
Figure 3: Time series of subjects' actions: strategies, speed subscriptions and minimum spread. In each box, we plot traders' strategies in CDA (blue lines) and FBA (green lines). Each column of the figure represents a market configuration and each row represents a variable related to the average of the outcome across groups. The dashed lines correspond to theoretical equilibrium values predicted by the BCS model. Faint solid lines correspond to observed time series sampled at five-second intervals and bold solid lines are their exponentially-weighted moving averages with a ten-second half life.

configurations, 2 and 3. Interestingly, the fraction of FBA market makers is relatively constant across configurations, and the increased difference in market making under those configurations is largely related to a diminished fraction of makers in the CDA. To the extent that the five-second averages exhibit accurate variation in the data, the time series show a clear statistical separation.

We note that the time series provide mild evidence of learning. Specifically, for several of the metrics, the data exhibits a trend during the first two periods, which subsequently stabilizes. This learning behavior is slightly different between formats. In FBA, for example, the average number of role changes in the first two periods is 11.4 and then falls 7.33 for the remaining periods. In the CDA, the number of role changes is higher (consistent with the coordination issues of the predicted equilibrium) and stable in all periods. The prevalence of market making, on the other hand, is declining in the first periods of the CDA markets, but stable throughout the whole FBA session.

Result 2: Fewer subjects choose to be snipers under FBA.

Predatory behavior (the fraction of traders acting as snipers), reported in model 2 of Table 1, is substantially lower for all FBA configurations relative to the CDA, with the difference falling in the range $[10\%, 44\%]$. Notably, the total population of snipers is much higher in the CDA for the volatile configurations (C2 and C3) , and lower in the FBA for those same configurations. Again, tests of the hypotheses $\gamma_2 > \gamma_1$ and $\gamma_3 = \gamma_1$ cannot be rejected at the 1% level, suggesting that the FBA has a greater impact on discouraging sniping in the most volatile regime with few investors.

The BCS model predicts the equilibrium percentage of snipers to be $\frac{5}{N} * 100 = 500/6 = 83.33\%$ under CDA and 0% under FBA, for all configurations. These values are depicted as blue and green (respectively) dashed lines in the second row of plots in Figure 3. As with the percentage of market making, the time series in those plots do not achieve the equilibrium values, but exhibit a clear separation in the direction of the equilibria, with only marginal separation for configuration 1. Again, similar to the percentage of market makers, the difference in the FBA - CDA percentages of snipers is larger in volatile environments with fewer investors.

Result 3: Fewer subjects purchase speed under FBA.

Model 3 of Table 1 shows that purchases of speed services are consistently lower under FBA. Specifically, the total fraction of traders purchasing speed services under

23

CDA falls in the range $[56\%, 69\%]$, while the FBA - CDA differences are in the range $[-48\%, -36\%]$. Equilibrium predictions, that $\alpha_j = 100$ and $\gamma_j = -100$ for $j \in \{1, 2, 3\}$, are rejected at the 1% level.

The BCS model predicts the equilibrium percentage of players purchasing speed services to be 100% under CDA and 0% under FBA, for all configurations. These values are depicted as blue and green (respectively) dashed lines in the third row of plots in Figure 3. Similar to the metrics discussed above, the time series in those plots do not achieve the equilibrium values, but exhibit a clear separation in the direction of the equilibria, and the magnitudes of the differences are slightly larger for configurations 2 and 3.

Result 4: Minimum spreads are lower under FBA.

Transactions costs, measured by the average minimum spread, are predicted by the BCS model to be $s^* = 0.324, 0.566$ and 0.475 for configurations 1, 2 and 3, respectively, in the CDA (see Section 2.4) and $s^* = 0$ for all configurations in the FBA. Model 4 in Table 1 shows that observed values are close to the minimum increment of 0.1 ECUs in FBA, while the average CDA spreads are much higher and closer to their theoretical counterparts.[12] The FBA - CDA differences in spread were $-0.123, -0.502$, and -0.560 ECUs, respectively for C1, C2 and C3, and serve as upper bounds for the differences, since they exclude the many investor-to-investor transactions at the fundamental value (zero spread) that occur in the FBA when an equal number of buy and sell investors arrive within one batch. In general, these differences closely track equilibrium values: for C2 we do not reject the null hypothesis that $\alpha_1 = .566$ and $\gamma_1 = -0.46$, while for C1 and C3, the minimum spread values are close to, but statistically distinct from, equilibrium levels. Despite this, a key comparative static across CDA configurations holds in the data: CDA-C2 and CDA-C3 both have statistically higher minimum spread relative to CDA-C1.

The BCS equilibrium minimum spreads are depicted as blue (CDA) and green (FBA) dashed lines in the fourth row of plots in Figure 3. The observed time series are largely congruent with those values, both in direction and magnitude. Perhaps most interesting is the fact that observed FBA spreads consistently remained at or near the floor of 0.1 ECUs implemented in the lab interface.

Panel (b) of Table 2 reports summary statistics for the metrics outlined above, sampled at one-second intervals, averaged across both time and groups and within

---

[12]Our lab interface implemented a minimum spread of 0.1 ECUs.

treatment. As with Figure 3, we exclude the first ten seconds of each four-minute period in order to eliminate start-up effects, and we also exclude the first two periods of each session, to account for learning. Congruent with the results above, those values demonstrate that under the FBA (1) more traders choose to act as makers, (2) fewer choose to act as snipers, (3) fewer choose to purchase speed services, and (4) minimum spreads are smaller.

Although Budish et al. (2015) focused on the "naive" extension of the static Nash equilibrium at each instant (batch) for the CDA (FBA), there is the possibility of collusive behavior in our experimental markets. In online Appendix C we discuss the manner in which collusive play would exhibit in our experiments and also reference the relevant literature. Our data indicates, however, that human traders in both market formats exhibited behavior more congruent with competitive, rather than collusive, play. Collusive play in these undercutting games (both CDA and FBA) would imply prevalence of market making at maximum spreads and no purchase of speed. Instead, we observe that most markets operated closer to competitive play: market spreads were always near the competitive equilibrium predictions (far away from that of collusive play). In the case of the FBA, the market spread was almost always at the lower bound of the admissible spread in the computer interface. In the case of CDA, although market making was significantly more common than the static Nash equilibrium prediction, the majority of traders subscribed to speed services and spreads were very close to the competitive equilibrium. See online Appendix C for further discussion. In terms of profits, collusive play in FBA would predict profits equal to $(27.9, 20.3, 35.2)$ ECUs per period/player for configurations 1–3, respectively. In contrast, the observed profits of FBA (reported in Table 2) were much lower: $(0.435, 0.372, 1.52)$ ECUs for configurations 1–3, respectively. Similarly, for the CDA, while predicted collusive profits were $(26.7, 16.0, 40.0)$ ECUs per period/player, observed profits were $(0.0869, 0.603, 4.31)$ ECUs. This evidence is consistent with the theoretical and empirical literature, suggesting that cooperation is rarely observed when games have three or more players.

3.2   Market Outcomes

Panel (c) of Table 2 reports summary statistics that measure volatility of prices, volatility of strategy choices (changes of roles), pricing deviations from fundamental value, number of transactions, and trader profits. The first two rows of panel (c) provide a measure of transaction price volatility via the standard deviation of transaction price differences, $Std(P_t - P_{t-1})$. We derive the equilibrium equations that govern these val-

| | | Configuration 1 | | Configuration 2 | | Configuration 3 | |
		CDA	FBA	CDA	FBA	CDA	FBA
(a) Market conditions							
λ_I		1/3		1/5		1/2	
λ_V		1/4		1		1	
c_{speed}		0.01		0.01		0.022	
(b) Choices							
Making (%)	Experiment	54	78.1	30.2	78.8	40.1	72.9
	Equilibrium	16.7	100	16.7	100	16.7	100
Sniping (%)	Experiment	31	20.8	58.1	14.5	49.5	14
	Equilibrium	83.3	0	83.3	0	83.3	0
Speed (%)	Experiment	56.1	19.7	69	31.7	69.2	20.7
	Equilibrium	100	0	100	0	100	0
Min. Spread	Experiment	0.226	0.103	0.677	0.179	0.709	0.147
	Equilibrium	0.324	0	0.566	0	0.475	0
(c) Market stats							
$Std(P_t - P_{t-1})$	Experiment	2.51	0.561	4.62	1.00	6.68	1.11
	Equilibrium	0.241	0.289	0.276	0.327	0.235	0.430
$Std(MinSpread)$	Experiment	0.204	0.0235	0.536	0.144	0.394	0.127
	Equilibrium	0	0	0	0	0	0
Status Changes	Experiment	20.5	6.26	31.6	6.26	17.0	7.34
	Equilibrium	N/A	0	N/A	0	N/A	0
$RMSD(P_t - V_t)$	Experiment	0.347	0.212	0.512	0.410	0.460	0.381
	Equilibrium	0.223	0.136	0.329	0.211	0.372	0.276
# Transactions	Experiment	156	85.2	172	99.3	248	134
	Equilibrium	106	80	100	48	147	120
Period Profits	Experiment	.0869	.435	.603	.372	4.31	1.52
	Equilibrium	0	0	0	0	0	0

Table 2: Summary statistics for experimental data. Panel (a) Shows the parameters of the three configurations we use in the experiments. Panel (b) reports the average percentage of subjects acting as market makers, snipers, average percentage of subjects purchasing speed services, and the average minimum spread posted by market makers. Values are sampled at three-second intervals and averaged across time and subjects within each treatment. Predicted equilibrium values are reported below observed averages. Panel (c) reports summary measures related to volatility (standard deviation of price changes and minimum spread, number of strategy/role changes), informational efficiency (root mean squared deviation from price to fundamental value), volume (number of transactions), and profits. For both panels, we exclude the first ten seconds of each period and the first two periods of each session to account for starting effects.

ues under CDA and FBA in Propositions D.1 and D.2 of online Appendix D, and report those values for each experimental configuration in the second row of panel (c). The equilibrium CDA values are quite constant across configurations, ranging from 0.241 to 0.276 ECUs. The FBA values are all larger, primarily due to the fact that they measure volatility over 3-second time (batch) intervals, and they are substantially larger for the more volatile configurations, ranging from 0.289 to 0.430 ECUs. The data, on the other hand, show that observed price differences under FBA are roughly two or three times more volatile than their theoretical counterparts, and are between five and 7.5 times more volatile in the CDA configurations, with price volatility higher (for both CDA and FBA) in configurations 2 and 3.

Rows 3-6 of panel (c) report volatility measures related to traders' choices: the standard deviation of the minimum spread and the average number of changes in trading strategy (status) by subjects within a trading period. In equilibrium, the standard deviation of minimum spread should be zero since makers always choose the fixed equilibrium spread, s^*. In the data, the standard deviation of minimum spread is nonzero but very low under FBA, and between three and five times greater under CDA, with higher values under configurations 2 and 3. The model, however, makes no theoretical statement about frequency of strategy switching in the CDA since it only demands a fixed strategy profile in aggregate, while under the FBA there should be no switching since all traders play an identical strategy. Empirically, status changes are between two and 3.5 times greater under the CDA configurations, relative to FBA. The coordination difficulties that can only arise in the CDA are compatible with the CDA exhibiting a higher number of strategy changes than the FBA.

Rows 7 and 8 of panel (c) report the root mean squared deviation (RMSD[13]) of transaction prices relative to the contemporaneous fundamental value. We derive the equilibrium equations that govern these values under CDA and FBA in Propositions D.3 and D.4 of online Appendix D, and report those values for each experimental configuration in the eighth row of panel (c). The theoretical CDA values range from 0.223 to 0.372 ECUs, while the FBA values are between 1.3 and 1.6 times smaller, ranging between 0.136 and 0.276 ECUs. Empirically, the RMSDs are between 1.2 and 1.9 times larger than their theoretical counterparts, with the CDA values consistently larger than those under FBA.

To additionally measure the effect of treatments on RMSD, Model 5 of Table 1

---

[13] $RMSD = \sqrt{\frac{\sum (P_t - V_t)^2}{N_{trans}}}$, where N_{trans} represents the number of transactions within a trading period.

reports the following regression:

$$RMSD_h = \sum_{j=1}^{3} \left[\alpha_j Cj_h + \gamma_j Cj_h \times FBA_h \right] + \epsilon_h. \tag{7}$$

The regression is fit at the level of trading period and group (h), and therefore utilizes 142 observations.[14] Consistent with the equilibrium predictions outlined in above, RMSD is lower under FBA. Specifically, RMSD under CDA is in the range $[0.347, 0.512]$, while the FBA - CDA differences are in the range $[-0.135, -0.0791]$.

Rows 9 and 10 of panel (c) report the number of transactions and trading profits earned by makers and snipers during the experiment. In equilibrium, the number of CDA transactions should be

$$N_{trans} = 240\lambda_I + 240\lambda_V \Pr\left(J > \frac{s^*}{2} \right) \frac{5}{6}, \tag{8}$$

where 240 is the length of each trading period in seconds. The respective equilibrium values of N_{trans} for configurations 1, 2 and 3 are 106, 100 and 147, which are much lower than the observed values reported in row 9 of panel (c). Under the FBA, the number of equilibrium transactions is $240\lambda_I$, or 80, 48 and 120 for configurations 1, 2 and 3, which, aside from configuration 2, are close to the observed values in row 9 of panel (c). The FBA counts include transactions that occur between investors, and aside from configuration 2, the theoretical values are close to the observed values. The discrepancy in C2 emerges from the high frequency of value jumps and transactions of makers sniping makers, described in online Appendix B.

The last two rows of Table 2 report net profits for makers and snipers in ECUs. Equilibrium profits for both formats are identically zero. Observed values are calculated by summing all gains and losses in each trading period and averaging the aggregated values across periods and subjects (or by subtracting the endowment from the end-of-period profits). In all configurations, observed profits are positive, ranging from 0.0869 to 4.31 ECUs, with configuration 3 having the highest profits for both CDA and FBA formats. For the markets included in this calculation (period > 2), only 0.5% of the 840 cases (140 markets times six traders) resulted in a trader finishing the period with a negative total profit.

---

[14]Six treatments with four groups each, trading in periods 3-8, less two group-periods because of the session that stopped after period seven.

Although the BCS environment is not suited for studying measures of allocative efficiency, such measures are of tangential interest. The percentage of unfilled investor orders is such a measure. Both formats predict that 100% of investors orders are filled in equilibrium, while in practice, temporary coordination failure in the CDA (such as all traders acting as snipers) can result in a lower percentage. Indeed, while only 0.14% of investors' orders were not filled in the FBA, the same value was 2.57% in the CDA.

It is also important to note that the purchase of speed is a measure of deviation from Pareto efficiency. Since traders improve profits off by collectively abstaining from speed purchases, increased use of fast communication technology is associated with social waste. This is due to the fact that $V(t)$ is publicly observable: in such an environment, faster communication technology serves no purpose in acquiring more or better information.

In summary, the measures in panel (c) of Table 2 show that the FBA reduces the volatility of transaction prices and minimum spreads, enhances price efficiency, results in more stable trader choices, and is more closely aligned with equilibrium profits.[15][16]

3.3   Transitory Market Dynamics

Haruvy et al. (2007) find evidence that traders' beliefs are adaptive and heavily influenced by market trends. Chan et al. (1996) further document empirical evidence and rationale for momentum in asset market returns. To determine if subjects in our experiments were likewise influenced by transitory, off-equilibrium trends in observable market information, we fit a vector autoregression of the form:

$$\boldsymbol{y}_t = \boldsymbol{a} + \boldsymbol{\Phi}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{9}$$

$$\boldsymbol{y}_t' = [\%Sniper_t, \%Speed_t, MinSpread_t, Turbulence_t], \tag{10}$$

where, as before, $\%Sniper_t$ is the average fraction of subjects choosing to be snipers during time interval t, $\%Speed_t$ is the average fraction of players choosing to purchase speed services during time interval t, and $MinSpread_t$ is the average minimum spread in the market over time interval t. $Turbulence_t$ is defined as the ratio of number of price changes to the number of investor arrivals, $\frac{N_{V,t}}{N_{I,t}}$, during time interval t [17]. Since N_V and

---

[15]None of the findings on individual and market behavior are sensitive to excluding the initial ten seconds and the first two periods of data from analysis.

[16]Online Appendix D provides a detailed calculation of the equilibrium market statistics.

[17]We do not include the fraction of market makers in the VAR since it is highly co-linear with the

N_I are exogenously determined by calibrated Poisson processes, we constrain the VAR so that the last element of $\boldsymbol{a}$ and the last row of $\boldsymbol{\Phi}$ are equal to zero, causing the last element of $\boldsymbol{\varepsilon}_t$ to be equal to $Turbulence_t$. We set the time interval of the regression to be 3 seconds, as this is the natural interval of the FBA treatment (the length of the batch), and is reasonable interval over which to measure transitory effects in the CDA [18].

From the perspective of a market maker, turbulence measures the countervailing forces of price volatility (increased sniping costs) with rate of investor arrivals (increased market making income). Further, the theoretical counterpart λ_V/λ_I plays a direct role in the determination of the BCS equilibrium for the CDA, as seen in Equation (3). Equation (5) shows that only λ_V has an impact on the FBA equilibrium. The constrained VAR in Equations (9) and (10) not only allows us to measure the transitory effects of players' choices in one period on subsequent behavior, it also captures the transitory effects of turbulence on each outcome variable, net of related effects on other choices.

Table 3 reports parameter estimates of the constrained VAR(1) for both CDA and FBA. The first row of panel (b) shows that the fraction of snipers, the fraction of subjects purchasing speed and turbulence all have a positive and significant (at least at the 5% level) relationship with subsequent decisions to snipe. The minimum spread, on the other hand, has a negative relationship, at the 5% level. Interestingly, the only lagged variable to be significantly related to speed purchases, is the fraction of agents purchasing speed, whereas minimum spread is positively and significantly (again, at least at the 5% level) related to all variables except speed purchases. Panel (b), however shows that the only statistically significant relationships under FBA are variables lagged with themselves. Specifically, turbulence has no impact on subsequent behavior in the FBA. Altogether, the results show that very-short term, innovations in market conditions impact behavior in the CDA, while such effects of transient market changes do not exist in the FBA.

Figure 4 depicts impulse responses for the estimated VARs reported in Table 3 for a horizon of 20 three-second periods, or 1 minute of clock time. In each case, the impulse responses measure the effect of a one-unit increase in the impulse variable, holding all other variables constant. The responses in both figures are congruent with the parameter estimates reported in Table 3. Specifically, under FBA shocks to each of the

---

fraction of snipers.

[18]We also evaluated the VAR for longer time intervals, with diminishing effects in the interval length, suggesting that subjects' reactions to the changing environment are short term. This is corroborated in the impulse responses that we report below.
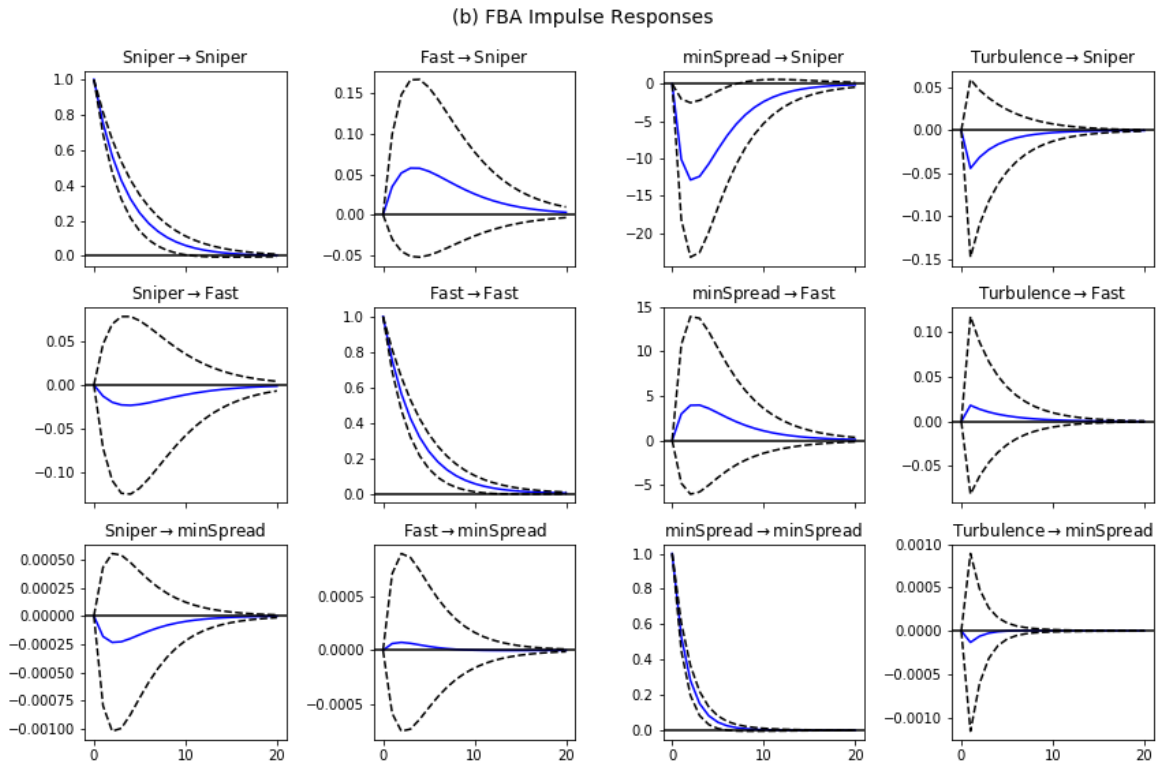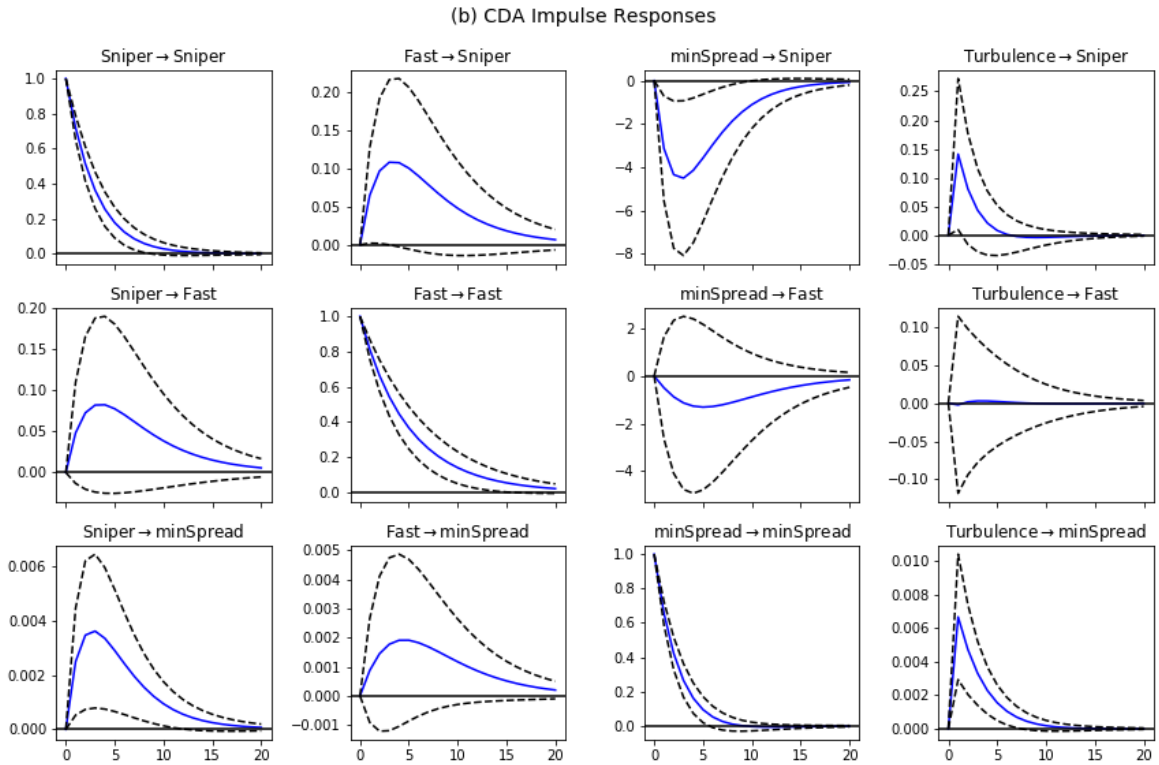
Figure 4: Unit impulse responses for the estimated VAR under CDA (panel a) and FBA (panel b).

	(a) CDA				
	Constant	$\%Sniper_{t-1}$	$\%Speed_{t-1}$	$MinSpread_{t-1}$	$Turbulence_{t-1}$
$\%Sniper_t$	10.1***	0.720***	0.0651**	-3.14**	0.141**
	(2.15)	(0.0354)	(0.0320)	(1.24)	(0.0670)
$\%Speed_t$	10.3***	0.0480	0.813***	-0.494	-0.00220
	(1.90)	(0.0313)	(0.0283)	(1.09)	(0.0593)
$MinSpread_t$	-0.00389	0.00250**	0.000877	0.656***	0.00665***
	(0.0611)	(0.00101)	(0.000911)	(0.0352)	(0.00191)
	(b) FBA				
	Constant	$\%Sniper_{t-1}$	$\%Speed_{t-1}$	$MinSpread_{t-1}$	$Turbulence_{t-1}$
$\%Sniper_t$	4.66***	0.752***	0.0347	-10.1**	-0.0442
	(1.03)	(0.0317)	(0.0331)	(4.13)	(0.0529)
$\%Speed_t$	5.73***	-0.0129	0.752***	2.98	0.0182
	(0.995)	(0.0305)	(0.0319)	(3.97)	(0.0509)
$MinSpread_t$	0.0683***	-0.000182	0.0000620	0.529***	-0.000132
	(0.0102)	(0.000313)	(0.000327)	(0.0408)	(0.000523)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Parameter estimates for the constraint VAR(1) in Equations (9) and (10). Standard errors are reported in parentheses. Panel (a) reports CDA estimates and panel (b) reports FBA estimates.

variables only have autocorrelative effects and no cross-correlative effects in subsequent time periods. Under CDA (panel (a) of Figure 4), (1) a transitory 1% increase in the fraction of snipers leads to a small (no more than 0.003 ECUs), positive and significant increase in the minimum spread for about 40 seconds, (2) a transient 1% increase in the fraction of subjects purchasing speed services results in an increase in the fraction of snipers by as much as 0.1% for no more than 10 seconds, (3) a transitory increase in the minimum spread by 1 ECU leads to a significant decline in the fraction of snipers by as much as 4% for nearly 30 seconds, and (4) a unit increase in turbulence results in a very short (less than 5 seconds), but significant increase in the fraction of snipers and a longer (a little over 20 seconds) and significant increase in the minimum spread, by as much as 0.006 ECUs.

We repeated the VAR analysis using only the final four periods of each session and found that the results were nearly identical to those reported above. This suggests that the short-term dynamics are a permanent feature of the respective formats and not caused by different starting effects or learning patterns. As a whole, our VAR results show that relative to the CDA, the FBA attenuates behavioral responses to short-term shocks in market conditions. We highlight that these are behavioral patterns not

32

predicted by the model, which is silent regarding market dynamics. Understanding this dynamic behavior is an important reason to conduct experiments in the lab and field.

4   Conclusions

We use laboratory experiments to empirically study and compare the performance of continuous double auctions and frequent batch auctions as financial market allocation mechanisms. The environment for our experiments follows the model of Budish et al. (2015), where a single asset is traded on a single exchange and two exogenous processes generate incentives to trade: changes in a publicly-observed fundamental value of an asset and the arrival of market orders from noise traders (*investors*). In our experiment, human participants (acting as traders) tune algorithms that trade on their behalf. Traders have access to a costly technology that reduces latency of messaging with the (remote) exchange. We emulate modern financial markets by developing an electronic architecture in which information and trading occur at millisecond time granularity and which follows the Nasdaq OUCH messaging protocol. Each of the market formats (CDA and FBA) is studied under three market conditions, varying the degree of volatility in the market and the cost of technology.

We find that, compared to the CDA, the FBA exhibits higher levels of liquidity, less predatory behavior, less investment in communication technology, lower transactions costs, and higher informational efficiency. Specifically, there are between 25-50% more market markers under the FBA than the CDA, 10-44% fewer aggressive traders seeking to trade on stale information, and 36-49% fewer traders purchasing speed technology. Further, the average minimum spread among market makers is substantially lower in the FBA, volatility of transactions prices and spreads is lower, and deviations in transactions prices from the underlying asset value are lower as well. As an additional measure of interest, we estimate the sensitivity of traders' behavior to transitory shocks in several market environment variables and find that the sensitivity is statistically absent in the FBA, despite the fact that such shocks have statistically significant short-run effects on behavior in the CDA.

We also find that market behavior is closer to BCS predictions in more turbulent markets. There are two possible explanations for this finding. First, it could be that the mechanisms outlined in Budish et al. (2015) become stronger in stressed markets. Second, markets with greater turbulence might provide more opportunities for learning since information arrives more frequently. Although our evidence suggests that learn-

ing cannot explain the finding, our experiment is not designed to discriminate clearly between those two explanations. We leave the study of this question to future research.

Taken together, our results suggest that the frequent batch auction is a welfare improving mechanism for allocating trade within a financial market. Such market designs are of interest to policy makers and regulators, as they reduce both volatility and dead-weight loss, and shift lost surplus to investors with fundamental portfolio needs. Further, our results suggest that exchange operators may find such alternative market designs promising, as they will inherently attract liquidity from fundamental investors and encourage fast traders to forgo expenditure on costly communication technology and instead focus their attention on liquidity provision.

We recognize that our results are as much a test of the BCS model as a comparison of the FBA and CDA formats, and that they are only relevant to the extent that the model broadly characterizes actual strategies and interactions by agents in the market (which we think it does). Beyond the relatively simple strategy space for agents in this environment, it would be useful to test the robustness of the FBA mechanism to more sophisticated behavior and complex preferences. Such behavior might include traders that place multiple-unit orders (which either deepen liquidity or have substantial price impact), investors that intelligently shred orders through time, and reactive algorithms (for both makers and snipers) that respond to other traders' strategies rather than to changes in the environment alone. We hope the approach of this paper will generate insights for such research program, and for the development of new theory.

References

Aldrich, E. M. and Friedman, D. (2017), "Order Protection through Delayed Messaging," *Working Paper*, 1–43.

Anderson, S., Johnston, D., Walker, J., and Williams, A. (1991), "The efficiency of experimental asset markets: Empirical robustness and subject sophistication," in *Research in Experimental Economics*, ed. R. Mark Isaac, Greenwich, CT: JAI Press, pp. 107–190.

Ariely, D., Ockenfels, A., Ariely, D., Ockenfels, A., Roth, A., and Roth, A. (2005), "An Experimental Analysis of Ending Rules in Internet auctions," *RAND Journal of Economics*, 36, 890–907.

Baldauf, M. and Mollner, J. (2015a), "High-Frequency Trading and Market Performance," *Working Paper*.

— (2015b), "Trading in Fragmented Markets," *Working Paper*.

Bershova, N. and Rakhlin, D. (2012), "High-Frequency Trading and Long-Term Investors: A View from the Buy Side," *Working Paper*.

Biais, B., Foucault, T., and Moinas, S. (2014), "Equilibrium Fast Trading," *Working Paper*.

Breckenfelder, J. (2013), "Competition Between High-Frequency Traders, and Market Quality," *Working Paper*.

Brogaard, J. and Garriott, C. (2015), "High-Frequency Trading Competition," *Working Paper*.

Brogaard, J., Hendershott, T., and Riordan, R. (2014), "High Frequency Trading and Price Discovery," *Review of Financial Studies*, 27, 2267–2306.

— (2015), "Price Discovery without Trading: Evidence from Limit Orders," *Working Paper*, 1–51.

Budish, E., Cramton, P., and Shim, J. (2014), "Implementation details for frequent batch auctions: Slowing down markets to the blink of an eye," *American Economic Review: Papers & Proceedings*, 104, 418–424.

— (2015), "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response," *The Quarterly Journal of Economics*, 130, 1547–1621.

Camerer, C. and Weigelt, K. (1991), "Information Mirages in Experimental Asset Markets," *The Journal of Business*, 64, 463–493.

Cason, T. N. and Friedman, D. (1996), "Price formation in double auction markets," *Journal of Economic Dynamics and Control*, 20, 1307–1337.

— (1997), "Price Formation in Single Call Markets," *Econometrica*, 65, 311–345.

— (1999), "Price formation and exchange in thin markets: A laboratory comparison of institutions," in *Money, Markets and Method: Essays in Honour of Robert W. Clower*, eds. Howitt, P., de Antoni, E., and Leigonhufvud, A., Elgar, no. February, chap. 8, pp. 155–179.

— (2008), "A Comparison of Market Institutions," in *Handbook of Experimental Economics Results*, eds. Plott, C. R. and Smith, V. L., Amsterdam: North-Holland, vol. 1, chap. 33, pp. 264–272.

Chan, L. K. C., Jegadeesh, N., and Lakonishok, J. (1996), "Momentum Strategies," *The Journal of Business*, 51, 1681–1713.

Copeland, T. E. and Friedman, D. (1987), "The Effect of Sequential Information Arrival on Asset Prices: An Experimental Study," *The Journal of Finance*, 42, 763–797.

— (1991), "Partial Revelation of Information in Experimental Asset Markets," *The Journal of Finance*, 46, 265–295.

Du, S. and Zhu, H. (2017), "What is the Optimal Trading Frequency in Financial Markets?" *Review of Economic Studies*.

Forsythe, R. and Lundholm, R. (1990), "Information Aggregation in an Experimental Market," *Econometrica*, 58, 309–347.

Forsythe, R., Palfrey, T. R., and Plott, C. R. (1982), "Asset Valuation in an Experimental Market," *Econometrica*, 50, 537–567.

— (1984), "Futures Markets and Informational Efficiency: A Laboratory Examination," *The Journal of Finance*, 39, 955–981.

Foucault, T. (1999), "Order flow composition and trading costs in a dynamic limit order market," *Journal of Financial Markets*, 2, 99–134.

Fricke, D. and Gerig, A. (2015), "Too Fast or Too Slow? Determining the Optimal Speed of Financial Markets *," *Working Paper*.

Friedman, D. (1993), "How Trading Institutions Affect Financial Market Performance: Some Laboratory Evidence," *Economic Inquiry*, 31, 410–435.

— (2008), "laboratory financial markets," in *The New Palgrave Dictionary of Economics*, eds. Durlauf, S. N. and Blume, L. E., Palgrave Macmillan, 2nd ed.

Friedman, D., Harrison, G. W., and Salmon, J. W. (1984), "The Informational Efficiency of Experimental Asset Markets," *Journal of Political Economy*, 92, 349–408.

Friedman, D. and Rust, J. (eds.) (1993), *The Double Auction Market: Institutions, Theory, and Evidence*, New York: Routledge, 1st ed.

Greiner, B. (2015), "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125.

Haas, M. and Zoican, M.-A. (2016), " Discrete or Continuous Trading? HFT Competition and Liquidity on Batch Auction Markets," *Working Paper*.

Hagströmer, B. and Nordén, L. (2013), "The diversity of high-frequency traders," *Journal of Financial Markets*, 16, 741–770.

Haruvy, E., Lahav, Y., and Noussair, C. N. (2007), "Traders' Expectations in Asset Markets: Experimental Evidence," *The American Economic Review*, 97, 1901–1920.

Hasbrouck, J. and Saar, G. (2013), "Low-latency trading," *Journal of Financial Markets*, 16, 646–679.

Hendershott, T. and Riordan, R. (2013), "Algorithmic trading and the market for liquidity," *Journal of Financial and Quantitative Analysis*, 48, 1001–1024.

Hirschey, N. (2013), "Do High Frequency Traders Anticipate Buying and Selling Pressure?" *Working Paper*.

Holt, C. A. (1995), "Industrial Organization: A Survey of Laboratory Research," in *The Handbook of Experimental Economics*, eds. Kagel, J. H. and Roth, A. E., Princeton: Princeton University Press, vol. 1, chap. 5, pp. 349–444.

Hu, E. (2018), "Intentional Access Delays, Market Quality, and Price Discovery: Evidence from IEX Becoming an Exchange," *Working Paper*.

Huber, J., Kirchler, M., and Sutter, M. (2008), "Is more information always better? Experimental financial markets with cumulative information," *Journal of Economic Behavior and Organization*, 65, 86–104.

Jovanovic, B. and Menkveld, A. J. (2015), "Middlemen in Limit Order Markets," *Working Paper*.

Kagel, J. H. (1998), "Double Auction Markets with Stochastic Supply and Demand Schedules: Call Markets and Continuous Auction Trading Mechanisms," *Working Paper*, 1–38.

Kyle, A. S. and Lee, J. (2017), "Toward a Fully Continuous Exchange," *Working Paper*.

Lewis, M. (2015), *Flash Boys: A Wall Street Revolt*, New York: W. W. Norton & Company.

Malinova, K., Park, A., and Riordan, R. (2014), "Do retail traders suffer from high frequency traders?" *Working Paper*.

McCabe, K. A., Rassenti, S. J., and Smith, V. L. (1992), "Designing Call Auction Institutions: Is Double Dutch The Best?" *The Economic Journal*, 102, 9–23.

Menkveld, A. J. and Zoican, M. A. (2017), "Need for speed? Exchange latency and liquidity," *Review of Financial Studies*, 30, 1188–1228.

Morone, A. and Nuzzo, S. (2016a), "Asset Markets in the Lab: A Literature Review," *Working Paper*, 1–30.

— (2016b), "Market Efficiency, Trading Institutions and Information Mirages: Evidence from an Experimental Asset Market," *Working Paper*, 1–39.

Narang, M. (2010), "Tradeworx Inc. Public Commentary on SEC Market Structure Concept Release," *Technical Report*.

Nasdaq (2018), "OUCH Version 4.2," *Technical Specification*.

Noussair, C. and Tucker, S. (2006), "Futures Markets and Bubble Formation in Experimental Asset Markets," *Pacific Economic Review*, 11, 167–184.

Noussair, C. N. and Tucker, S. (2013), "Experimental Research on Asset Pricing," *Journal of Economic Surveys*, 27, 554–569.

Nuzzo, S. and Morone, A. (2017), "Asset markets in the lab: A literature review," *Journal of Behavioral and Experimental Finance*, 13, 42–50.

Plott, C. R. and Smith, V. L. (1978), "An Experimental Examination of Two Exchange Institutions," *The Review of Economic Studies*, 45, 133–153.

Plott, C. R. and Sunder, S. (1982), "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models," *Journal of Political Economy*, 90, 663–698.

— (1988), "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085–1118.

Roth, A. E. and Ockenfels, A. (2002), "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review*, 92, 1093–1103.

Roth, A. E. and Xing, X. (1994), "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *The American Economic Review*, 84, 992–1044.

— (1997), "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," *Journal of Political Economy*, 105, 284–329.

Schnitzlein, C. R. (1996), "Call and Continuous Trading Mechanisms Under Asymmetric Information: An Experimental Investigation," *The Journal of Finance*, 51, 613–636.

Smith, V. L. (1962), "An Experimental Study of Competitive Market Behavior," .

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988), "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56, 1119–1151.

Smith, V. L., Williams, A. W., Bratton, W. K., and Vannoni, M. G. (1982), "Competitive Market Institutions: Double Auctions vs . Sealed Bid-Offer Auctions," *The American Economic Review*, 72, 58–77.

Sunder, S. (1995), "Experimental Asset Markets: A Survey," in *The Handbook of Experimental Economics*, eds. Kagel, J. H. and Roth, A. E., Princeton: Princeton University Press, chap. 6, pp. 445–500.

Theissen, E. (2000), "Market structure, informational efficiency and liquidity: An experimental comparison of auction and dealer markets," *Journal of Financial Markets*, 3, 333–363.

Van Boening, M. V., Williams, A. W., and LaMaster, S. (1993), "Price bubbles and crashes in experimental call markets," *Economics Letters*, 41, 179–185.

Webb, R. I., Muthuswamy, J., and Segara, R. (2007), "Market Microstructure Effects on Volatility at the TAIFEX," *The Journal of Futures Markets*, 27, 1219–1243.

Williams, A. W. (1980), "Computerized Double-Auction Markets: Some Initial Experimental Results," *The Journal of Business*, 53, 235–258.

Zhang, S. and Riordan, R. (2011), "Technology and Market Quality: The Case of High Frequency Trading," *ECIS 2011 Proceedings*.