

Aldrich, Eric M.; Friedman, Daniel

Working Paper

A theoretical model of the investors exchange

Working Paper, No. 738

Provided in Cooperation with:

University of California Santa Cruz, Economics Department

Suggested Citation: Aldrich, Eric M.; Friedman, Daniel (2017) : A theoretical model of the investors exchange, Working Paper, No. 738, University of California, Economics Department, Santa Cruz, CA

This Version is available at:

<https://hdl.handle.net/10419/197764>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A Theoretical Model of the Investors Exchange

Eric M. Aldrich*

Department of Economics

University of California, Santa Cruz

Daniel Friedman†

Department of Economics

University of California, Santa Cruz

July 8, 2017

Abstract

Investors' Exchange LLC (IEX) is a newly approved public exchange that is designed to discourage aggressive high-frequency trading. We explain how IEX differs from traditional continuous double auction markets and present summary data on IEX transactions by trader class and order type. Our primary contribution is a simple analytic model of IEX as a constrained version of the continuous double auction. The model predicts that IEX will generally improve price efficiency and lower transactions cost while increasing delay costs. A subset of the model's predictions are testable in the field or in a laboratory environment.

Keywords: Market design, IEX, lab experiments, high-frequency trading, continuous double auction.

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

Acknowledgements:

We are indebted to the Center for Analytical Finance (CAFIN) at UCSC for seed funding of the High Frequency Trading project of which this paper is a part, and to Dan Aisen for access to data and for insight into the IEX system. We benefited from helpful comments from Peter Cramton, Natalia Lazzati, Kristian Lopez Vargas, John Musacchio, and Nirvikar Singh. The conclusions and perspectives in this paper are the authors'; none of those individuals or organizations should be held responsible for remaining idiosyncrasies and shortcomings.

*Email: ealdrich@ucsc.edu.

†Email: dan@ucsc.edu.

1 Introduction

On June 17th, 2016, the United States Securities and Exchange Commission (SEC) approved Investors' Exchange LLC (IEX) to operate as a public stock exchange. Founded in 2012, the primary objective of IEX is to improve financial market operation for ordinary investors by reducing the advantages of trading firms that invest heavily in speed technology. To achieve that objective, all inbound and outbound messages to the IEX trading engine are delayed by 350 microseconds by routing them through a 38-mile cable that is coiled in a box. This innovation, commonly referred to as the “shoe box” or “speed bump”, is the focus of most previous discussions of the IEX market design (e.g. Lewis, 2015; Pisani, 2016).

However, on its own the IEX speed bump does little to promote market fairness. Rather, the function of the speed bump is to allow the IEX trading engine to have an accurate view of the National Best Bid and Offer (NBBO) across public exchanges at the time that it executes buy and sell orders. IEX's primary innovation is to allow orders that are pegged to the NBBO (a common order type on public exchanges) to adjust before they are exposed to orders routed through the shoe box. As a result, slow traders using pegged orders are largely protected from fast traders who attempt to “snipe” stale orders when new price information enters the market. Thus, while the speed bump is an essential feature of the IEX system, it only promotes fairness to the extent that it protects pegged orders.

Protection comes at a cost. All pegged orders at IEX are “hidden”, i.e., not visible in the limit order book. To encourage the submission of visible (“lit”) orders, IEX gives priority to lit orders at any given price, even those that arrive after hidden orders at the same price. Thus pegged orders face the implicit cost of always being queued behind visible orders.

Our goal in the present paper is to build a simple equilibrium model that (a) gives insight into the tradeoffs among different order types at IEX, and (b) facilitates comparison of IEX market outcomes to those in traditional continuous double auction (CDA) markets employed by most equities exchanges.

Our analysis begins in Section 2 by describing how the IEX system modifies the CDA format, and by reviewing summary data of transactions at the exchange, categorized by

trader type and order type.

Section 3 lists our simplifying assumptions and obtains closed-form equilibrium expressions for the usage rates of market orders vs. pegged orders and the prevalence of market makers and snipers. Section 4 obtains parallel results for CDA markets while also checking robustness by relaxing some of the more restrictive assumptions and Section 5 summarizes the impact of removing order protection.

Section 6 provides a concluding discussion, and highlights some empirically testable implications. Proofs, extended technical analysis, and additional institutional details can be found in Appendices A, B and C.

2 Institutional Background

A financial market format specifies how orders are processed into transactions. In this section we provide a general description of continuous double auctions (CDA) and a more specific description of the Investors' Exchange (IEX) format. We then present a summary of trading data from IEX and use the data to motivate elements of our model in Section 3.

2.1 Continuous Double Auction format

Most modern financial markets use variants of the *continuous double auction* (CDA) format, also known as the continuous limit order book (CLOB). A *limit order* is a message to the exchange 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). The CDA limit order book collects and sorts bids by (1) price and (2) time received (at each price), and likewise collects and sorts offers. The highest bid price and the lowest offer price are referred to, respectively, as the *best bid* and *best offer*, and the difference between them is called the *spread*.

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 SEC mandates that prices shown in the order book are discretized in units of pennies per share; fractions of a penny are not allowed in a standard limit order price. In contrast, time remains essentially continuous, which creates interesting complications for the CDA format (see Budish et al., 2015).

At present, there are 12 SEC-approved “national securities exchanges” in the United States that trade U.S. equities instruments. Under Regulation National Market System (Reg NMS) these exchanges are required to report transactions and quotations to a centralized processor, known as the Securities Information Processor (SIP). The SIP monitors all bids and offers at all 12 exchanges, and constantly updates the official National Best Bid and Offer (NBBO), consisting of the National Best Bid (NBB) and National Best Offer (NBO). However, since the speed of light is finite and the 12 exchanges have different physical locations, there is no “true” NBBO — at best there is an NBBO from the perspective of the SIP. For this reason, unlike the order books internal to an exchange, which never lock or cross, it is possible for the NBB or NBO to temporarily lock or cross with the best bid or offer at a specific exchange. These instances are fleeting, as Reg NMS requires other exchanges with less aggressive quotations to pass orders on to those with better bids or offers.

Most CDA exchanges recognize a variety of order types beyond simple limit orders. *Market orders* are the most common variation, specifying a very high bid or very low offer price and zero time in force. Most exchanges also recognize “hidden” orders which are not publicly displayed in the order book and which are given lower priority than ordinary “lit” (displayed) orders. The lexicographic priority system then is: price, display, time. For example, all hidden bid orders are prioritized after the lit orders at a specific price; among

themselves they are prioritized on a first-come, first-served basis, even if there are different types of hidden orders.

An important type of hidden order is a *pegged* limit order. An NBB peg is a limit order that enters the book at the current NBB and is automatically re-priced by the exchange when the NBB changes. An NBO peg is treated symmetrically at the NBO. The SEC also permits exchanges to offer hidden (but not lit) *midpoint pegs*: bids or offers that track the midpoint (and hence at sub-penny prices) of NBB and NBO.

2.2 The IEX format

The IEX market is a unique CDA variant that delays all inbound and outbound messages to its messaging server by 350 microseconds. This delay is long enough to allow the system a fresh view of the NBBO and to reprice pegged orders ahead of new messages that are coincident with changes in the NBBO. As a result, pegged orders are protected from fast traders who profit from transacting at stale prices when the NBBO changes.

Besides traditional (lit) limit and market orders, IEX offers the following types of (hidden) pegged orders:

- Midpoint peg. Limit orders pegged to the NBBO midpoint. By virtue of their more aggressive price, they have priority over traditional limit orders.
- Primary peg. Limit orders that are booked one price increment (typically \$0.01) away from NBB or NBO, but which are promoted to transact at NBB or NBO if sufficient trading interest arrives at those prices.
- Discretionary peg. Limit orders which first enter the (non-displayed) order book at the midpoint but, if not executed immediately, rest at either the NBB or NBO; see Appendix C for more details.

Unlike other US exchanges, IEX charges fees only for midpoint transactions and for nothing else.

2.3 Some Data

	Other Nonroutable				Primary Peg			
	Hidden		Lit		Hidden		Lit	
	BBO	Mid	BBO	Mid	BBO	Mid	BBO	Mid
Agency Remover	0.0349	0.0069	0.0395	0	0	0	0	0
Prop Remover	0.0474	0.0072	0.0255	0	0	0	0	0
Agency Adder	0.0078	0.0100	0.0691	0	0.0264	0	0	0
Prop Adder	0.0080	0.0021	0.0473	0	0.0219	0	0	0

	Midpoint Peg				Discretionary Peg			
	Hidden		Lit		Hidden		Lit	
	BBO	Mid	BBO	Mid	BBO	Mid	BBO	Mid
Agency Remover	0	0.1008	0	0	0	0.1004	0	0
Prop Remover	0	0.0719	0	0	0	0.0233	0	0
Agency Adder	0.0063	0.0716	0	0	0.0391	0.1968	0	0
Prop Adder	0.0002	0.0212	0	0	0.0009	0.0136	0	0

Table 1: IEX volume shares for December 2016 by order type and transaction price. Excludes routable orders and transactions in locked or crossed market conditions.

Table 1 reports transaction volume statistics at IEX during the month of December 2016. The data exclude periods when markets were locked or crossed with NBBO (3.4% of volume) and exclude transactions involving orders routable to other exchanges (12.3% of volume; see Appendix C for a discussion of routable orders). The Table entries are normalized to sum to 100%, and so they are shares of the remaining 84.3% of all transactions.

IEX classifies traders into two broad types: (1) agencies (brokers), who provide services to and receive fees from external clients and who compete to offer rapid order execution at favorable prices, and (2) proprietary firms, who trade on their own account, maintaining net positions close to zero, and who earn revenue by buying at prices a bit lower on average than selling prices (either by adding liquidity at a spread or removing liquidity when stale quotes

persist in the order book). Firms that do both are classified as agencies.

Table 1 shows that Agency firms represent over 70% of volume at IEX; volume at other exchanges is typically more evenly split between agencies and proprietary traders. Agency volume has three main components:

1. Adding lit orders at BBO: 7.7% of transaction volume. Our model in the next section will attribute this to the proprietary arm of integrated agency firms.
2. Removing orders at BBO: 7.4% of volume. Our model will attribute this to impatient investor clients.
3. Midpoint and discretionary peg orders transacting at midpoint: 47.0% of volume. Our model will attribute this to less impatient clients.

Following is a similar breakdown for proprietary firms; again see Appendix C for more details.

1. Passive (adding) orders at BBO: 5.5% of volume. Our model attributes this to market making by proprietary firms.
2. Aggressive (removing) orders at BBO: 7.3% of volume. The model attributes this to proprietary “snipers,” who exploit unprotected stale limit orders when the NBBO changes.
3. Midpoint and discretionary peg orders transacting at midpoint: 13.0% of volume. For simplicity, and since they comprise only 25% of all midpoint and discretionary orders, our model will group this order flow with midpoint orders transmitted by agencies on behalf of their clients.

3 Baseline Model

Our baseline model is of a continuous double auction that protects midpoint pegs. The model highlights tradeoffs between order types under simplifying assumptions on the grid of

asset prices and on exogenous variables specifying investor arrival and changes in the asset's fundamental value. This baseline model also makes stark assumptions regarding who buys speed and which orders are protected from fast traders; the extended model in Section 4 will eliminate protection and will examine speed purchase decisions.

3.1 Assumptions

- A1.** The market consists of a single asset trading at a single exchange, one indivisible unit at a time.
- A2.** Prices lie on a discrete, uniform grid $\mathcal{P} = 1, 2, \dots, \hat{P}$. A price unit, i.e., the grid step size, represents half of the minimum price increment (e.g. a half penny per share).
- A3.** The fundamental value of the asset, V , follows a marked Poisson process on \mathcal{P} . The fundamental value changes to $V' \in \{V - 2, V + 2\}$ with equal innovation rate $\nu > 0$. That is, the total innovation rate is 2ν , with one-sided rate ν of a two-increment (e.g., one penny) upward jump and one-sided rate ν of a two-increment downward jump.
- A4.** An exogenous flow of impatient investors with unit demands arrive independently at Poisson rate $\rho > 0$ on each side of the market.
 - a. Investors have marginal gross surplus $\varphi > 1$ per unit of the asset.
 - b. An investor may have the broker transmit a market order, which executes immediately at midpoint and incurs execution fee $d \in (0, 1)$ if there is a contra-side midprice order resting in the (hidden) order book. If there is no such resting order, the market order executes immediately at the BBO and incurs trading cost of 1 (half of the bid/offer spread).
 - c. Alternatively, an investor may have the broker transmit a midpoint peg order, absorbing execution fee $d \in (0, 1)$. The pegged order executes immediately if there is a resting contra-side midprice order; otherwise it goes to the end of the same-side (hidden) order queue. If a midpoint peg order is not executed immediately, its net surplus is discounted at rate $\delta \geq 0$.

A5. The cost of speed, $c > 0$, and time lags in responding to innovations in V are such that

- a. Traders placing orders at BBO do not purchase speed and, when V jumps, are susceptible to sniping by other proprietary traders who do purchase speed.
- b. Snipers can reverse transactions immediately at V' .
- c. Pegged orders track V with so short a lag that they are protected from sniping.

Assumptions A1 and A2 are straightforward simplifications to sharpen the analysis. A3 and A4 are unrealistic¹ but together they capture the idea that the fundamental equilibrates supply and demand even while experiencing exogenous shifts. Making the price exogenous sharpens the comparison of market formats.

Assumptions A4a and A4c are intended to capture investors' impatience in a simple way and A4b implicitly assumes a deep order book at the BBO. Assumption A4c conflates discretionary pegs with midpoint pegs and ignores primary pegs; see Appendix C for a justification of this simplification. Assumption A5 is intended to streamline the first part of the analysis; later sections will analyze timing issues in more detail.

3.2 Action Space and State Space

The model presumes a streamlined set of just three order types:

- r : proprietary traders add single unit regular lit limit orders at the best bid ($V - 1$) and best offer ($V + 1$).
- p : brokers add single unit midpoint peg limit orders at price V . These orders are hidden, and are subject to a transaction fee d upon execution. Currently $d = \$0.0009$, less than one half spread (typically $\$0.005$). Since p orders execute against each other, there can be resting p orders on only one side of the market at any given time.

¹Indeed, at extreme prices ($V = 1, 2$ and $\hat{P} - 1, \hat{P}$), some jumps are infeasible so A3 must be modified. As a practical matter, the SEC permits the grid to be redefined in such extreme cases. Here, to keep the focus on matters of greater interest, we assume that such modifications are negligible because we are operating far away from the extremes.

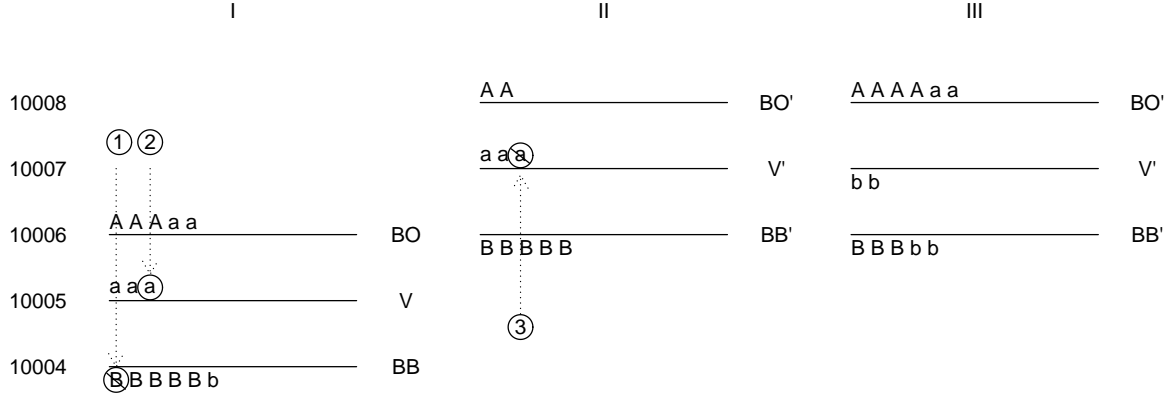


Figure 1: Example states in IEX market. Uppercase (lowercase) characters represent lit (hidden) orders at a particular price, and their position in the order book is represented by their horizontal position from the left. Character ‘A’ denotes an ask or offer; character ‘B’ denotes a bid. Panel I: initial state is $k = 2$ and $V = 10005$ half-pennies (i.e., $V = \$50.025$ per share); event (1) is a market sell order which ‘crosses the spread’ to transact at BB 10004 = \$50.02; event (2) is a midprice offer which rests at V implying a transition to $k = 3$. Panel II: V has jumped to 10007 = \$50.035; event (3) is a market buy order or midprice bid which transacts at V and triggers transition $k = 3 \rightarrow 2$. Panel III: V remains at 10007 but an excess of bids relative to offers drives k to -2.

m : brokers’ market orders remove liquidity at the midpoint if it is occupied by contra-side orders, in which case they also incur the fee d . When there are no contra-side midprice orders, an m order removes an r order at the best bid or best offer.

Since we assume that the order book is deep at BBO, and since there can be a positive number of midpoint orders on only one side of the market, the state of the market is described by the level of the fundamental value, $V \in \mathcal{P}$, and the *order imbalance* $k \in \mathbb{Z}$ at the midpoint price. By convention, $k < 0$ means that there are precisely $-k > 0$ midpoint peg buy orders (hidden) in the order book, $k > 0$ indicates k midpoint peg sell orders (hidden) in the order book, and $k = 0$ indicates an empty queue at the midpoint price V . See Figure 1.

New investor arrivals can trigger transitions in k . Let ω denote the fraction of arrivals that brokers transmit as midpoint peg orders, with the remaining $1 - \omega$ fraction transmitted

as market orders. Given the symmetry in Assumption A4, a new arrival generates a midpoint peg buy or sell order with probability $\omega/2$ each, or a market buy or sell order with probability $(1 - \omega)/2$ each. A new pegged sell (buy) order always generates a transition $k \rightarrow k + 1$ ($k \rightarrow k - 1$). A new market sell (buy) order generates a transition $k \rightarrow k + 1$ when $k < 0$ ($k \rightarrow k - 1$ when $k > 0$) and otherwise executes at BBO and generates no transition.

The following proposition, proved in Appendix A, characterizes the stationary distribution of k .

Proposition 3.1. *Let $\omega \in (0, 1)$ be the probability that an investor arrival on either side of the market results in a midprice peg order. Given Assumptions A1-A5, there is a unique steady state distribution $q : \mathbb{Z} \rightarrow (0, \infty)$ of the order imbalance,*

$$q_k = \left(\frac{1 - \omega}{1 + \omega} \right) \omega^{|k|}, \quad k \in \mathbb{Z}. \quad (3.1)$$

Equation (3.1) tells us that the steady state distribution is symmetric around a sharp peak at zero, and drops off at exponential rate as $|k|$ increases. The distribution has a single parameter, the fraction ω of investor orders that are transmitted as midpoint peg orders rather than as market orders. The fraction is determined endogenously in equilibrium, as shown below.

3.3 Investor Surplus

Following assumption A4, investors choose between midpoint peg orders and market orders.

Peg. A midpoint peg order generates surplus φ less the execution charge d . These orders transact immediately with contra-side midpoint peg orders when such orders exist, but are otherwise placed at the back of the midpoint queue (e.g., to position $k + 1$ when the state is $k \geq 0$), and thus incur delay costs expressed in terms of the given discount rate δ . The relevant discount factor is $\beta = \exp\left(-\frac{\delta}{\rho}\right) \leq 1$ when contra-side orders arrive at rate $\rho > 0$. Thus, by A4, the conditional expected net surplus is $(\varphi - d)\beta^{k+1}$ for a pegged sell order when $k \geq 0$.

Using steady state probabilities (3.1) for order imbalance k , the unconditional expected net surplus for a midpoint peg sell order is

$$\begin{aligned}
\pi_p &= (\varphi - d) \left[\sum_{k=-\infty}^{-1} q_k + \sum_{k=0}^{\infty} q_k \beta^{k+1} \right] \\
&= (\varphi - d) \left[\frac{1 - \omega}{1 + \omega} \right] \left[\frac{\omega}{1 - \omega} + \frac{\beta}{1 - \beta\omega} \right] \\
&= \left(\frac{\varphi - d}{1 + \omega} \right) \left[\omega + \frac{\beta(1 - \omega)}{1 - \beta\omega} \right]
\end{aligned} \tag{3.2}$$

The model's symmetry ensures that (3.2) also applies to pegged buy orders.

Market order. Like a pegged order, with probability $\sum_{k=-\infty}^{-1} q_k$ a market order will execute immediately against a contra-side midpoint peg and earn $\varphi - d$. With probability $\sum_{k=0}^{\infty} q_k$, there will be no contra-side midpoint orders and, unlike a pegged order, a market order will execute immediately against an r order at BBO. In that case, since the price is 1 half-spread away from V , it earns $\varphi - 1$. Thus the expected net surplus for a market order (either buy or sell) is

$$\pi_m = (\varphi - d) \sum_{k=-\infty}^{-1} q_k + (\varphi - 1) \sum_{k=0}^{\infty} q_k = \left(\frac{\varphi - d}{1 + \omega} \right) \omega + \frac{\varphi - 1}{1 + \omega}. \tag{3.3}$$

3.4 Proprietary Trader Profits

Proprietary traders divide themselves into two groups: market makers and snipers.

Snipers trade off the flow cost, c , of buying speed against profits from sniping stale r orders following a jump in the fundamental value V . The number of potential targets, N_r , is the equilibrium number of regular limit orders at best bid and best offer. Thus, when an opportunity arises, each of N_s snipers uses fast market orders to obtain on average $\frac{N_r}{N_s}$ successful snipes.² By assumptions A3 and A5, each successful snipe of a resting r order

²Assumption A1 can be construed as limiting each sniper to at most one snipe per V jump. Under that interpretation (which we do not adopt), the factor $\frac{N_r}{N_s}$ in equation (3.4) is replaced by $\min\{1, \frac{N_r}{N_s}\}$, and (3.5) requires a similar modification. In that case, when $2\nu < c$, snipers necessarily earn negative profit, so in equilibrium there are $N_s = 0$ snipers and $N_r = +\infty$ market makers. However, in the less expensive sniping case $2\nu \geq c$, the formulas below are unaffected by the alternative interpretation of A1.

involves buying (or selling) a single share at $V + 1$ (or $V - 1$) and reversing the transaction at $V' = V + 2$ (or $V' = V - 2$), yielding a profit of 1 half-spread. Since opportunities arrive at both sides of the market at rate ν , the expected flow profit for a sniper is

$$\pi_s = 2\nu \frac{N_r}{N_s} - c. \quad (3.4)$$

Market makers place r orders at the BBO, trading off the single half-spread gain of transacting with a market order against a possible half-spread loss to a sniper. Market orders arrive at each side of the market at rate $(1 - \omega)\rho$ and jumps to each side occur at rate ν so, by maintaining both a bid and ask at BBO, a market maker obtains expected flow profit

$$\pi_r = \frac{2(1 - \omega)\rho}{N_r} - 2\nu. \quad (3.5)$$

3.5 Equilibrium

Definition 3.1. *Given an exogenous flow cost of speed $c > 0$, midpoint transaction fee $d \geq 0$, investor gross surplus $\varphi \geq 1$ and discount factor $\beta = \exp\left(-\frac{\delta}{\rho}\right) \in (0, 1]$, and arrival rates $\rho, \nu > 0$ for investors and fundamental value innovations, the vector (ω^*, N_r^*, N_s^*) constitutes a market equilibrium if*

1. *at $\omega^* \in (0, 1)$ (resp. $\omega^* = 0$, $\omega^* = 1$), a midpoint peg order has (resp. no more than, no less than) the same expected net surplus as a market order, and*
2. *with $N_s^* \geq 0$ snipers and $N_r^* \geq 0$ market makers, proprietary traders earn zero expected profit from either activity.*

The idea behind the first equilibrium condition is that investors and brokers will increase the fraction $\omega \in (0, 1)$ of pegged orders whenever the expected surplus differential $\pi_p - \pi_m$ is positive, and decrease ω when the differential is negative. Hence expected (net discounted) surplus should be equal at an interior steady state, while at corner solutions $\omega^* = 0$ (1) we should have $\pi_p \leq (\geq) \pi_m$. The second condition arises from the reasonable assumption that there are no substantial barriers to entry or exit for either of the two proprietary activities.

Under current assumptions, equilibrium takes a simple form.

Proposition 3.2. *Under assumptions A1 - A5 and parameter restrictions $\varphi \geq 1 > d \geq 0$, $\beta = \exp\left(-\frac{\delta}{\rho}\right) \in (0, 1]$, and $\rho, \nu > 0$, there is a unique market equilibrium (ω^*, N_r^*, N_s^*) . If $\omega^* \in (0, 1)$, then*

$$\omega^* = \frac{\varphi - d - \beta^{-1}(\varphi - 1)}{1 - d} \quad (3.6)$$

is the equilibrium fraction of brokers/investors choosing midpoint pegs vs. market orders, and the equilibrium numbers of proprietary traders choosing to act as market makers and snipers are, respectively,

$$N_r^* = \frac{\rho}{\nu}(1 - \omega^*) \quad (3.7)$$

$$N_s^* = \frac{2\rho}{c}(1 - \omega^*). \quad (3.8)$$

Proof. Applying the first market equilibrium condition we obtain equation (3.6) as follows:

$$\begin{aligned} \pi_p = \pi_m &\iff \left(\frac{\varphi - d}{1 + \omega}\right) \left[\frac{\beta(1 - \omega)}{1 - \beta\omega}\right] = \frac{\varphi - 1}{1 + \omega} \\ &\iff (\varphi - d)\beta(1 - \omega) = (\varphi - 1)(1 - \beta\omega) \\ &\iff \omega^* = \frac{\varphi - d - \beta^{-1}(\varphi - 1)}{1 - d}. \end{aligned} \quad (3.9)$$

If the last expression in (3.9) is negative, then it is straightforward to show that $\pi_p(0) \leq \pi_m(0)$ and that $\omega^* = 0$; if that last expression is greater than 1.0, then $\pi_p(1) \geq \pi_m(1)$ and $\omega^* = 1$.

We obtain Equations (3.7) and (3.8) by applying the second market equilibrium condition $\pi_r = \pi_s = 0$ to Equations (3.5) and (3.4) and solving for N_r and N_s . \square

4 Extensions

The equilibrium in Section 3 was derived under the assumptions that pegged orders are protected from sniping, that a sufficient number of r orders always rest at the BBO, and that market makers do not purchase speed technology. We now explore what happens when some of those assumptions are relaxed.

Timing notation. Jumps in the fundamental value V are registered at the Securities Information Processor (SIP) and resting p orders automatically adjust in parallel with latency

$\tau_{SIP} > 0$. By default, traders' messages to the exchange are subject to round-trip latency τ_{slow} , however, for flow cost c , traders may reduce their latency to $\tau_{fast} < \tau_{slow}$. The exchange imposes an additional fixed delay $\eta \geq 0$ so that $\tilde{\tau}_{fast} = \tau_{fast} + \eta$ and $\tilde{\tau}_{slow} = \tau_{slow} + \eta$. In traditional CDA markets, $\eta = 0$, while at IEX it is chosen so that $\tilde{\tau}_{fast} = \tau_{fast} + \eta > \tau_{SIP}$. To compress notation, let

$$\xi = \begin{cases} 1 & \text{if } \tilde{\tau}_{fast} = \tau_{fast} + \eta \leq \tau_{SIP} \\ 0 & \text{if } \tilde{\tau}_{fast} = \tau_{fast} + \eta > \tau_{SIP}, \end{cases} \quad (4.1)$$

which is interpreted as an order protection parameter. When $\xi = 0$, pegged orders are fully protected from sniping and jump in tandem with V before any other messages reach the exchange, as in assumption A5c. When $\xi = 1$, however, they may be sniped for profit by fast traders.

4.1 Unprotected Midpoint Pegs: $\xi = 1$

We continue to assume that liquidity adders do not have access to speed technology, so resting p orders are vulnerable when V jumps and $\xi = 1$. A successful sniper gains (and the liquidity adder loses) $|V' - V| = 2$ half-spreads on a midpoint peg, rather than the usual 1 half-spread on an order resting at BBO.

When a midpoint peg offer is queued behind k other pegged offers, it will be sniped if and only if a positive jump in V occurs before $k + 1$ buy orders arrive from brokers. Thus, the conditional probability of not being sniped is $\left(\frac{\rho}{\rho + \xi\nu}\right)^{k+1}$, with expected profit $(\varphi - d)\beta^{k+1}$, where the discount factor is still $\beta = \exp\left(-\frac{\delta}{\rho}\right)$. With complementary probability, the offer is sniped, resulting in a 2 half-spread loss discounted in the same manner. Midpoint peg bids are treated the same way as offers.

The steady state distribution of order imbalance, $k \in \mathbb{Z}$, is different than in the protected case, because sniping now induces transitions $k \rightarrow 0$. The following proposition, proved in Appendix A, generalizes the stationary distribution of k to cover the unprotected case.

Proposition 4.1. *Let $\omega \in (0, 1)$ be the probability that an investor arrival on either side of the market results in a midpoint peg order. Given assumptions A1-A5b, there is a unique*

steady state distribution $\tilde{q} : \mathbb{Z} \rightarrow (0, \infty)$ of the order imbalance, with

$$\tilde{q}_k = \left(\frac{1-\lambda}{1+\lambda} \right) \lambda^{|k|}, \quad k \in \mathbb{Z}, \quad (4.2)$$

where

$$\lambda = \frac{1}{2} \left(1 + \frac{\xi\nu}{\rho} + \omega \right) - \frac{1}{2} \sqrt{\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega} \in (0, \omega]. \quad (4.3)$$

The composite parameter ξ indicates whether pegged orders are protected from sniping ($\xi = 0$) or not ($\xi = 1$).

Note that that (4.3) collapses to $\lambda = \omega$ in the protected case $\xi = 0$.

We proceed as before to obtain the equilibrium value of λ , and thus ω .

Peg. Following Proposition 4.1 and the preceding discussion, one readily verifies that Equation (3.2), the expected net surplus for a midpoint peg order, generalizes to:

$$\begin{aligned} \pi_p &= (\varphi - d) \left[\sum_{k=-\infty}^{-1} \tilde{q}_k + \sum_{k=0}^{\infty} \tilde{q}_k \left(\frac{\beta\rho}{\rho + \xi\nu} \right)^{k+1} \right] - 2 \sum_{k=0}^{\infty} \tilde{q}_k \beta^{k+1} \left[1 - \left(\frac{\rho}{\rho + \xi\nu} \right)^{k+1} \right] \\ &= (\varphi - d) \left[\frac{\lambda}{1+\lambda} + \frac{1-\lambda}{1+\lambda} \frac{\beta\rho}{(\rho + \xi\nu - \beta\rho\lambda)} \right] \\ &\quad - \frac{1-\lambda}{1+\lambda} \frac{2\beta}{(1-\beta\lambda)} + \frac{1-\lambda}{1+\lambda} \frac{2\beta\rho}{(\rho + \xi\nu - \beta\rho\lambda)}. \end{aligned} \quad (4.4)$$

Market order. An investor choosing a market order will earn the same expected net surplus as in Equation (3.3) with \tilde{q}_k replacing q_k :

$$\pi_m = (\varphi - d) \sum_{k=-\infty}^{-1} \tilde{q}_k + (\varphi - 1) \sum_{k=0}^{\infty} \tilde{q}_k = (\varphi - d) \frac{\lambda}{1+\lambda} + (\varphi - 1) \frac{1}{1+\lambda}. \quad (4.5)$$

As in the protected case, the first terms in (4.4) and (4.5) represent execution against a contraside midpoint peg. Since they are identical, they again cancel in the equal surplus condition.

Sniper profit. Snipers now have $N_r + \xi N_p$ potential targets: the regular orders plus unprotected pegged orders. Since the profit is 2 half-spreads on the latter, Equation (3.4) becomes

$$\pi_s = 2\nu \frac{N_r + 2\xi N_p}{N_s} - c. \quad (4.6)$$

Market maker profit. The conditions for market makers are unchanged, so (3.5) still characterizes their profitability. The only difference is in the equilibrium fraction of investors choosing pegged orders, as we now show.

Proposition 4.2. *Given assumptions A1-A5b and parameter restrictions $\varphi \geq 1 > d \geq 0$, $\beta = \exp\left(-\frac{\delta}{\rho}\right) \in (0, 1]$, and $c, \rho, \nu > 0$, there is a unique market equilibrium. If the equilibrium fraction of investors choosing midpoint pegs is interior, i.e., $\tilde{\omega}^* \in (0, 1)$, then*

$$\tilde{\omega}^* = \tilde{\lambda} + \xi \left(\frac{\tilde{\lambda}}{1 - \tilde{\lambda}} \right) \frac{\nu}{\rho} \quad (4.7a)$$

$$N_p^* = \frac{\tilde{\lambda}}{1 - \tilde{\lambda}^2} \quad (4.7b)$$

$$N_r^* = \frac{cN_s^*}{2\nu} - 2\xi N_p^* = \frac{c(1 - \omega^*)\rho}{2\nu^2} - \frac{2\xi\tilde{\lambda}}{1 - \tilde{\lambda}^2} \quad (4.7c)$$

$$N_s^* = (1 - \tilde{\omega}^*) \frac{\rho}{\nu}, \quad (4.7d)$$

where

$$\begin{aligned} \tilde{\lambda} = \frac{1}{2\beta^2\rho(1-d)} & \left[\beta^2\rho(\varphi - d) - \beta\xi\nu(\varphi + 1) - \beta\rho(\varphi + d - 2) \right. \\ & - \left((\beta\xi\nu(\varphi + 1) + \beta\rho(\varphi + d - 2) - \beta^2\rho(\varphi - d))^2 \right. \\ & \left. \left. - 4\beta^2\rho(1-d)(\beta\rho(\varphi - d) - \rho(\varphi - 1) - \xi\nu(\varphi - 1 + 2\beta).) \right)^{1/2} \right] \end{aligned} \quad (4.8)$$

Proof. Equating expected surplus $\pi_p = \pi_m$ for pegged and market orders in Equations (4.4) and (4.5) yields the following quadratic expression in λ :

$$\frac{\beta\rho(1-\lambda)(\varphi-d)}{\rho+\xi\nu-\beta\rho\lambda} - \frac{2\beta(1-\lambda)}{1-\beta\lambda} + \frac{2\beta\rho(1-\lambda)}{\rho+\xi\nu-\beta\rho\lambda} = \varphi - 1. \quad (4.9)$$

Solving for λ via the usual quadratic formula results in two solutions. Appendix A shows that the condition $\lambda < 1$ requires the larger (smaller) solution to satisfy

$$(1 - \beta)(\varphi - 1)(\nu + \rho(1 + \beta)) < 0 (> 0) \quad (4.10)$$

The parameter restrictions ensure that the left-hand-side of Equation (4.10) is nonnegative, so the relevant solution involves the negative discriminant, which is written in Equation (4.8).

Equation (4.7a) follows from Corollary A.1.

Equation (4.7b) gives the expected number N_p^* of pegged orders vulnerable to sniping:

$$N_p^* = \sum_{k=-\infty}^0 0\tilde{q}_k + \sum_{k=1}^{\infty} k\tilde{q}_k = \frac{1-\tilde{\lambda}}{1+\tilde{\lambda}} \sum_{k=1}^{\infty} k\tilde{\lambda}^k = \left(\frac{1-\tilde{\lambda}}{1+\tilde{\lambda}} \right) \frac{\tilde{\lambda}}{(1-\tilde{\lambda})^2} = \frac{\tilde{\lambda}}{1-\tilde{\lambda}^2}. \quad (4.11)$$

Equations (4.7c) and (4.7d) follow by substituting (4.7a) and (4.7b) into Equations (3.5) and (4.6), setting them equal to zero in accordance with the equilibrium condition, and solving for N_r and N_s . \square

Remark. By the same logic as in the previous proposition, for parameters such that the left-hand side of (4.7a) is less than zero (or greater than 1), the reader can show that profit inequalities imply that $\tilde{\omega}^* = 0$ (or 1). Then the remaining formulas collapse; e.g., $N_p = 0$ by definition when $\tilde{\omega}^* = 0$, and it can be seen that $N_r^* = 0$ when $\tilde{\omega}^* = 1$.

Corollary 4.1. *In the limiting case $\delta/\rho \rightarrow 0$ ($\beta \rightarrow 1$), the steady-state value of λ is*

$$\hat{\lambda} = 1 - \xi \frac{(\varphi + 1)\nu}{(1-d)\rho}, \quad (4.12)$$

which is valid for $\rho \geq \frac{\varphi+1}{(1-d)}\nu$.

Proof. When $\beta = 1$, Equation (4.9) simplifies to

$$(\varphi - d)(1 - \lambda)\rho + 2(1 - \lambda)\rho = (\varphi + 1)(\rho + \xi\nu - \rho\lambda) \quad (4.13)$$

from which (4.12) follows. \square

5 Impact of order protection

Proposition 4.2 predicts, using a simplified model, the precise impact of removing midpoint peg protection. Given any admissible parameter vector, the equilibrium expressions can be evaluated separately for $\xi \in \{0, 1\}$; the difference represents the impact at that parameter configuration.

As explained in Appendix B, the vector $(c, d, \delta, \varphi, \rho/\nu) = (10, 0.18, 3, 1.8, 0.5)$ is our baseline calibration of the model parameters. With $\xi = 0$ (midpoint pegs protected) the

formulas in Proposition 4.2 give equilibrium values $(\tilde{\omega}^*, N_p^*, N_r^*, N_s^*) = (0.74, 1.66, 0.51, 0.66)$, and with protection removed we achieve a corner solution, $(\tilde{\omega}^*, N_p^*, N_r^*, N_s^*) = (0, 0, 2, 2.56)$.

Performance metrics. Does protection enhance market performance? Taking the investor's perspective, the main issues are price (which should be as close as possible to fundamental value), fees, and delay. Specifically, we propose the following three performance metrics.

1. Price efficiency. The mean absolute deviation, DP, of transaction price from fundamental value should be as small as possible. In our model, DP is 0 for investor orders transacting at the midprice, 1 for orders transacting at BBO, and 2 for midpoint orders that are sniped (only in the case of no protection). We define

$$Q = \sum_{k=1}^{\infty} \tilde{q}_k = \frac{\lambda}{1 + \lambda} \quad (5.1)$$

$$\begin{aligned} S &= \sum_{k=0}^{\infty} \tilde{q}_k \left[1 - \left(\frac{\rho}{\rho + \xi\nu} \right)^{k+1} \right] \\ &= \frac{1}{1 + \lambda} - \left(\frac{1 - \lambda}{1 + \lambda} \right) \frac{\rho}{\rho + \xi\nu - \rho\lambda} \\ &= \frac{\xi\nu}{(1 + \lambda)(\rho(1 - \lambda) + \xi\nu)}, \end{aligned} \quad (5.2)$$

to be (respectively) the probabilities that an investor order encounters contra-side interest at the midprice and that a midpoint peg is sniped. Thus, conditional on choosing a market order, an investor encounters no contra-side midpoint order with probability $1 - Q$, and conditional on choosing a pegged order, an investor is sniped with probability S . As a result,

$$\begin{aligned} DP &= (1 - \omega)(1 - Q) + 2\omega S \\ &= \frac{1 - \omega}{1 + \lambda} + \frac{2\xi\nu\omega}{(1 + \lambda)(\rho(1 - \lambda) + \xi\nu)}. \end{aligned} \quad (5.3)$$

When $\xi = 0$ and pegged orders are protected, $\lambda = \omega$ and Equation (5.3) simplifies to $DP = \frac{1 - \omega}{1 + \omega}$.

2. Transaction cost. Investors pay brokerage fee b , which is generally in the \$0.003–\$0.005 range (see Appendix B), or 0.6 to 1.0 half-spreads. With probability Q , a market order

executes immediately at midprice and is charged an additional explicit fee of d , while with probability $1 - Q$ it executes at BBO and pays an additional implicit fee of 1 half spread in the form of worse execution price. In equilibrium the cost to an investor should be the same for a pegged order as for a market order. Thus the per-share mean transaction cost is

$$TC = b + dQ + (1 - Q) = b + \frac{1 + d\lambda}{1 + \lambda}. \quad (5.4)$$

3. Delay cost. The expected fractional loss of surplus due to discounting is zero except for orders transmitted as midprice pegs that go to the back of the queue. By the logic of the previous section, conditional on same-side imbalance $k \geq 0$ the expected discount factor is $\left(\frac{\beta\rho}{\rho + \xi\nu}\right)^{k+1}$, implying a proportional loss $\left[1 - \left(\frac{\beta\rho}{\rho + \xi\nu}\right)^{k+1}\right]$ of net surplus. We define DC as the unconditional expected proportional loss,

$$\begin{aligned} DC &= \omega \sum_{k=0}^{\infty} \tilde{q}_k \left[1 - \left(\frac{\beta\rho}{\rho + \xi\nu}\right)^{k+1}\right] \\ &= \frac{\omega}{1 + \lambda} - \frac{\omega\beta\rho}{\rho + \xi\nu} \left(\frac{1 - \lambda}{1 + \lambda}\right) \sum_{k=0}^{\infty} \left(\frac{\beta\rho\lambda}{\rho + \xi\nu}\right)^k \\ &= \frac{\omega}{1 + \lambda} \left(\frac{\rho + \xi\nu - \beta\rho}{\rho + \xi\nu - \beta\rho\lambda}\right). \end{aligned} \quad (5.5)$$

Using the parameter vector above, Table 2 reports performance impact measures.

Measure	$\xi = 0$	$\xi = 1$	Difference
ω^*	0.743	0	0.743
DP	0.148	1	-0.852
TC	1.45	1.8	-0.35
DC	0.216	0	0.216

Table 2: Performance measures at market equilibrium with and without order protection, given baseline parameters $(c, d, \delta, \varphi, \rho/\nu) = (10, 0.18, 3, 1.8, 0.5)$.

Comparative statics. An important advantage of an analytic solution is that one can compute equilibrium values of all outcome variables and performance metrics for *any* admissible

parameter vector. What can we say about the general impact of order protection?

Clearly the equilibrium fraction ω^* of investor orders transmitted as midpoint pegs is the main driver for performance metrics; higher values imply lower DP and TC, albeit at the cost of higher DC. It seems reasonable to conjecture that midpoint pegs are more common when they are protected. As shown in Appendix A, the conjecture is true in the sense that removing protection decreases N_p^* . It is also true in the sense that, conditional on order imbalance k , removing protection impairs the profitability of midpoint peg orders more than that of market orders and thus tends to reduce their equilibrium share. However, there is a subtle indirect effect that goes in the other direction: the order imbalance distribution shifts towards smaller imbalances, resulting in faster fills for midpoint peg orders. This reduces the sniping hazard and makes pegs more attractive.

It turns out that for some extreme parameter values, the indirect effect more than offsets the direct effect. For example, with one modification to our baseline calibration, $\nu = 0.01$, the equilibrium values of ω with and without order protection, are respectively $\omega^* = 0.74277$ and $\tilde{\omega}^* = 0.74285$. Other parameter configurations generate even larger differences in the equilibrium fractions, but we highlight this particular case as it is only requires one modification of our baseline parameters. The proof of Proposition 4.1 in Appendix A shows that the solution to Equation (4.3) obeys the restriction $\lambda < \omega$. As observed in Equation (4.3), this relationship is a direct result of the term $\frac{\nu}{\rho}\xi$. Thus, when $\nu \ll \rho$, as in the parameters proposed above, the impact of ξ (order protection) is minimal and investors may decrease their flow of midpoint pegs even when offered protection from sniping.

6 Discussion

The ultimate source of profits for both proprietary traders and brokers in our model is the exogenous order flow from investors. Investor orders provide fee income to brokers, whose orders subsequently provide income to proprietary traders who make markets via lit resting orders at the best bid and best offer. Some of that income is diverted to snipers, who transact with stale BBO orders immediately following a jump in the fundamental value. Intuitively,

we have a food chain, with impatient investors' market orders sustaining regular limit orders, which sustain sniping.

For a small fee $d \ll 1$, IEX offers investors/brokers a better option: a hidden midpoint peg that is protected from snipers and executes at a (half spread) better price. However, pegged orders incur an expected delay that increases with the fraction, ω , of investors that choose pegs. When that expected delay³ is sufficiently disadvantageous, investors (or their brokers) will resort to the standard market order which executes against market makers' lit best bids and offers.

Our model lays out the equilibrium consequences of these tradeoffs, providing predictions that can be tested in laboratory and field data.

6.1 Testable predictions

The simplest version of the model is intended to capture the functioning of IEX in calm conditions. It assumes a thick order book of slow (unprotected) regular orders at BBO, and assumes that midpoint pegged orders are protected from sniping. Key predictions include

1. Both the number of active market makers N_r and snipers N_s will increase when the flow ρ of investors increases. Indeed, if the discount rate δ is proportional to ρ ,⁴ so that β and ω^* remain constant, then the equilibrium quantities of both types of proprietary traders are directly proportional to ρ , as seen in equations (3.7) - (3.8).
2. Equations (3.7) - (3.8) also show that an increase in ν (i.e., an increase in sniping opportunities) will proportionately decrease the population size N_r of market makers but (perhaps surprisingly) have no impact on the population size N_s of snipers.
3. An increase in the cost of speed, c , will proportionately reduce the population size N_s of snipers, but will have no effect on the population size N_r of market makers.

³Since pegged orders are hidden, traders can not observe the actual delay in advance.

⁴This might be the case if impatience arises mainly from concerns about preemption by other investors.

4. The fraction ω^* of impatient investors that transmit pegged orders is an increasing function of the discount factor $\beta = \exp\left(-\frac{\delta}{\rho}\right) \in (0, 1]$.

What happens when midpoint orders are not protected from sniping ($\xi = 1$)? According to Propositions 4.1 and 4.2,

1. For a given positive fraction ω of orders transmitted as pegs, Equation (4.3) tells us that $\lambda < \omega$, i.e., the order imbalance is more tightly distributed around zero.
2. The equilibrium value of ω^* is smaller when $\xi = 1$ for a wide range of parameter values, including baseline parameters. However, the inequality can go the other way for certain extreme parameter values.
3. Given the usual case that $\tilde{\omega}^* < \omega^*$, Equations (3.7), (3.8), (4.7c), and (4.7d) show that N_r and N_s will be larger for $\xi = 1$.
4. Most importantly, the performance metrics show that at baseline parameters with midpoint peg order protection, prices are far more efficient and transactions costs are slightly lower, but delay costs are somewhat higher.

6.2 Future work

To isolate the impact of order protection, the comparisons above assume that IEX's unique fee structure is maintained throughout. Future work with small variants on the present model could investigate the impact of fee structure, with and without order protection.

How well do current results stand up when key simplifying assumptions are relaxed? In particular, what happens when we relax Assumption A5 and allow market makers adding lit orders at BBO, and traders adding midpoint pegs, to also purchase speed at flow cost $c > 0$? Those who purchase speed have some chance of escaping snipers. Closed form analytic solutions no longer seem possible, but one can use recursion techniques (in particular, the Erlang B model) to solve numerically for equilibrium. Preliminary work so far indicates no qualitative changes to current results, but more work is needed.

A more ambitious extension of the present model would consider what happens when the fundamental value follows a more complicated rule than Assumption A3, or when it is unobservable and tied to the investor arrival process. Thus, one could replace A3 and A4 by an exogenous and time-varying process of investor arrivals in which V is implicitly defined by balancing buy and sell orders in expectation.

Another important theoretical extension is to relax Assumption A1 and to model competing exchanges, and possibly multiple securities. The NBBO and the fundamental values would be endogenous, given some appropriately specified overall investor demand that would distribute itself across exchanges, assets and order types.

Empirical work need not wait for these theoretical extensions. In the laboratory, one could investigate whether human subjects in the broker role track $\hat{\omega}$ when we vary δ, φ, d , and whether human subjects in the proprietary trader role follow the comparative static predictions with regard to the impact of ν, ρ, c on N_r, N_s . Using field data, one might examine the order balance distribution and the present model's comparative statics.

Appendices

A Distribution of Order Imbalance

A.1 Protected Pegged Orders

Proposition 3.1. *Let $\omega \in (0, 1)$ be the probability that an investor arrival on either side of the market results in a midprice peg order. Given Assumptions A1-A5, there is a unique steady state distribution $q : \mathbb{Z} \rightarrow (0, \infty)$ of the order imbalance,*

$$q_k = \left(\frac{1 - \omega}{1 + \omega} \right) \omega^{|k|}, \quad k \in \mathbb{Z}. \quad (3.1)$$

Proof. As noted in the text, an investor arrival generates a midpoint peg buy or sell order,

or a market buy or sell order, with respective probabilities $\omega/2, \omega/2, (1-\omega)/2, (1-\omega)/2$. Recall also that a new pegged sell (resp buy) order generates a transition $k \rightarrow k+1$ (resp. $k \rightarrow k-1$), while a new market sell (resp buy) order generates a transition $k \rightarrow k+1$ when $k < 0$ (resp. $k \rightarrow k-1$ when $k > 0$) and otherwise no transition.

Thus an arrival updates a negative imbalance probability $p(k|k < 0)$ to $p'(k|k < 0) = \frac{\omega}{2}p(k+1|k < 0) + \frac{1-\omega}{2}p(k|k < 0) + \frac{1}{2}p(k-1|k < 0)$; the first two terms arise from p and m buy orders respectively, and the third term from any sell order. In steady state $p'(k|k < 0)$ is equal to the pre-update value $p(k|k < 0)$. Thus we obtain the following steady state equation for negative imbalance, together with analogous equations for a zero imbalance and a positive imbalance:

$$p(k|k < 0) = \frac{\omega}{2}p(k+1|k < 0) + \frac{1-\omega}{2}p(k|k < 0) + \frac{1}{2}p(k-1|k < 0) \quad (\text{A.1})$$

$$p(0) = \frac{1}{2}p(1) + (1-\omega)p(0) + \frac{1}{2}p(-1) \quad (\text{A.2})$$

$$p(k|k > 0) = \frac{1}{2}p(k+1|k > 0) + \frac{1-\omega}{2}p(k|k > 0) + \frac{\omega}{2}p(k-1|k > 0). \quad (\text{A.3})$$

Substituting $p(k|k < 0) = q_k$ in (3.1) for all $k < 0$, and writing $b = \left(\frac{1-\omega}{1+\omega}\right)$ to reduce notation, we verify directly that equation (A.1) holds:

$$\begin{aligned} b\omega^{-k} &= \frac{\omega}{2}b\omega^{-(k-1)} + \frac{1-\omega}{2}b\omega^{-k} + \frac{1}{2}b\omega^{-(k+1)} \\ &= \left(\frac{1}{2} + \frac{1}{2}\right)b\omega^{-k} + \left(\frac{1}{2} - \frac{1}{2}\right)b\omega^{-(k+1)} \\ &= b\omega^{-k}. \end{aligned} \quad (\text{A.4})$$

Verifying that $q_k = p(k|k \geq 0)$ in (3.1) satisfies (A.2) and (A.3) is similarly straightforward. Any other solution of the steady state equations (A.1) - (A.3) must be proportional to q , so to prove the proposition it suffices to verify that q is a probability distribution on \mathbb{Z} :

$$\sum_{k \in \mathbb{Z}} q_k = b \left(1 + 2 \sum_{k=1}^{\infty} \omega^k \right) = b \left(1 + 2 \frac{\omega}{1-\omega} \right) = b \frac{1+\omega}{1-\omega} = bb^{-1} = 1. \quad \square$$

A.2 Unprotected Pegged Orders

When midpoint peg orders are not protected from sniping, the set of possible events increases to six: the four types of investor arrivals, in addition to increasing and decreasing jumps

in the fundamental. The stationary distribution then has the same general form as in the protected case, but is a more complicated function of the exogenous parameters.

Proposition 4.1. *Let $\omega \in (0, 1)$ be the probability that an investor arrival on either side of the market results in a midpoint peg order. Given assumptions A1-A5b, there is a unique steady state distribution $\tilde{q} : \mathbb{Z} \rightarrow (0, \infty)$ of the order imbalance, with*

$$\tilde{q}_k = \left(\frac{1 - \lambda}{1 + \lambda} \right) \lambda^{|k|}, \quad k \in \mathbb{Z}, \quad (4.2)$$

where

$$\lambda = \frac{1}{2} \left(1 + \frac{\xi\nu}{\rho} + \omega \right) - \frac{1}{2} \sqrt{\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega} \in (0, \omega]. \quad (4.3)$$

The composite parameter ξ indicates whether pegged orders are protected from sniping ($\xi = 0$) or not ($\xi = 1$).

Proof. The relative probabilities of the four investor events are unchanged from the previous proposition, but to accommodate the directional jumps in the fundamental value, those four probabilities all shrink by the factor $\frac{\rho}{\rho + \xi\nu}$. An upwards jump has probability $\frac{1}{2} \frac{\nu}{\rho + \nu}$ and causes the transition $k \rightarrow 0$ when $k > 0$ and $\xi = 1$, and otherwise has no effect. A downwards jump has the same probability and causes the transition $k \rightarrow 0$ when $k < 0$ and $\xi = 1$, and otherwise has no effect.

With those modifications, the equations parallel to (A.1) - (A.3) that define the steady state distribution become:

$$p(k|k < -1) = \frac{\omega}{2} \frac{\rho}{\rho + \xi\nu} p(k + 1|k < -1) + \left(\frac{1}{2} - \frac{\omega}{2} \frac{\rho}{\rho + \xi\nu} \right) p(k|k < -1) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(k - 1|k < -1) \quad (A.5)$$

$$\begin{aligned} p(0) &= \sum_{k \neq 0} \frac{1}{2} \frac{\xi\nu}{\rho + \xi\nu} p(k) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(1) + \left(1 - \omega \frac{\rho}{\rho + \xi\nu} \right) p(0) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(-1) \\ &= \frac{1}{2} \frac{\xi\nu}{\rho + \xi\nu} - \frac{1}{2} \frac{\xi\nu}{\rho + \xi\nu} p(0) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(1) + \left(1 - \omega \frac{\rho}{\rho + \xi\nu} \right) p(0) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(-1) \\ &= \frac{1}{2} \frac{\xi\nu}{\rho + \xi\nu} + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(1) + \left(1 - \frac{\rho\omega + (1/2)\xi\nu}{\rho + \xi\nu} \right) p(0) + \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(-1) \end{aligned} \quad (A.6)$$

$$p(k|k > 1) = \frac{1}{2} \frac{\rho}{\rho + \xi\nu} p(k+1|k > 1) + \left(\frac{1}{2} - \frac{\omega}{2} \frac{\rho}{\rho + \xi\nu} \right) p(k|k > 1) + \frac{\omega}{2} \frac{\rho}{\rho + \xi\nu} p(k-1|k > 1). \quad (\text{A.7})$$

By symmetry, $p(1) = p(-1)$, so the equation for $p(0)$ implies

$$p(1) = p(-1) = \left(\omega + \frac{\xi\nu}{2\rho} \right) p(0) - \frac{\xi\nu}{2\rho}. \quad (\text{A.8})$$

Solving Equations (A.5) - (A.8) for $p(k|k < 0)$ and $p(k|k > 0)$, we find

$$p(k+1|k > 0) = \left(1 + \frac{\xi\nu}{\rho} + \omega \right) p(k|k > 0) - \omega p(k-1|k > 0) \quad (\text{A.9})$$

$$p(k-1|k < 0) = \left(1 + \frac{\xi\nu}{\rho} + \omega \right) p(k|k < 0) - \omega p(k+1|k < 0). \quad (\text{A.10})$$

Equations (A.9) and (A.10) are linear second order homogeneous difference equations, whose general solution takes the form

$$p(k) = a_1 \lambda_1^{|k|} + a_2 \lambda_2^{|k|}, \quad (\text{A.11})$$

where

$$\lambda_1 = \frac{1}{2} \left(1 + \frac{\xi\nu}{\rho} + \omega \right) + \frac{1}{2} \sqrt{\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega} \quad (\text{A.12})$$

$$\lambda_2 = \frac{1}{2} \left(1 + \frac{\xi\nu}{\rho} + \omega \right) - \frac{1}{2} \sqrt{\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega}, \quad (\text{A.13})$$

are the roots of the quadratic equation

$$\lambda^2 - \left(1 + \frac{\xi\nu}{\rho} + \omega \right) \lambda + \omega = 0. \quad (\text{A.14})$$

The discriminant $\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega$ is bounded above by $\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2$ and bounded below by $\left(1 + \frac{\xi\nu}{\rho} + \omega \right)^2 - 4\omega(1 + \frac{\xi\nu}{\rho}) = \left(1 + \frac{\xi\nu}{\rho} - \omega \right)^2$ for all $\nu, \rho > 0$ and $\xi, \omega \in [0, 1]$. As a result, $\lambda_1 \geq 1$ and, as required by equation (4.3) of the proposition, $\lambda \equiv \lambda_2 \in (0, \omega)$. It is easily seen that $\lambda = \omega$ when $\xi = 0$.

From the boundary condition $p(k) \rightarrow 0$ as $k \rightarrow \infty$, we see that $a_1 = 0$ since $\lambda_1 \geq 1$. Consequently, Equation (A.11) implies that $p(0) = a_2$. Enforcing the summability constraint

for a probability distribution, we find:

$$1 = \sum_{k=-\infty}^{\infty} p(k) = a_2 + 2a_2 \sum_{k=1}^{\infty} \lambda^k = a_2 \left[1 + 2\frac{\lambda}{1-\lambda} \right] = a_2 \left[\frac{1+\lambda}{1-\lambda} \right], \quad (\text{A.15})$$

Hence, $a_2 = p(0) = \frac{1-\lambda}{1+\lambda}$ and from Equation (A.11) we obtain the desired expression (3.1). \square

Corollary A.1. *Given parameters ν , ρ and ξ , the steady-state fraction of brokers choosing to place midpoint peg orders is*

$$\omega = \lambda + \xi \left[\frac{\lambda}{1-\lambda} \right] \frac{\nu}{\rho}, \quad (\text{A.16})$$

where λ is the steady state value determined in Proposition 4.1.

Proof The result is obtained by solving for ω in Equation (A.14).

Remark. Clearly ω is strictly increasing in λ for the relevant parameter values, so its inverse function $\lambda(\omega|\xi = 1, \nu, \rho)$ exists and is also strictly increasing.

B Baseline Model Calibration

B.1 Investor Fraction ω^*

Recall that ω is the fraction of investor orders transmitted as midpoint pegs and for $\xi = 0$

$$Q = \sum_{k=1}^{\infty} q_k = \frac{\omega}{1+\omega} \quad (\text{B.1})$$

is the probability that there is a contra-side order resting at midprice (see, e.g., Equation (3.3)). A new investor order is represented in Table 1 in one of three ways:

1. With probability Q , an agency will remove liquidity at midprice.
2. With probability $\omega(1 - Q)$, and agency will add liquidity at midprice.
3. With probability $(1 - \omega)(1 - Q)$, an agency will remove liquidity at BBO.

Conditional on an agency order, the probabilities of these sum to unity; unconditionally (given the presence of proprietary traders) the sum of probabilities (.2081, .2784, .0744, respectively) is 0.5609. As a result,

$$\frac{0.2081}{0.5609} = Q = \frac{\omega}{1 + \omega} \implies \omega \approx 0.59 \quad (\text{B.2})$$

$$\frac{0.2784}{0.5609} = \omega(1 - Q) = \frac{1}{1 + \omega} \implies \omega \approx 1.01 \quad (\text{B.3})$$

$$\frac{0.0744}{0.5609} = (1 - \omega)(1 - Q) = \frac{1 - \omega}{1 + \omega} \implies \omega \approx 0.77. \quad (\text{B.4})$$

$$(\text{B.5})$$

The average of these values is 0.79, so we will choose baseline parameters that yield $\omega \approx .75$.

B.2 Midprice Transaction Fee

The IEX fee for transacting at the mid price is \$0.0009. As a single price unit in our model is equivalent to \$0.005, we set $d = 0.18$ price units.

B.3 Investor Surplus

We define φ as the surplus for the marginal investor with impatience β^* (defined below). Such an investor is willing to transmit a market order at unit cost (0.5 spreads or pennies) in addition to the direct fee of \$0.003 – \$0.005 (an approximation reported to us by practitioners) per share. The direct fee is equivalent to 0.6 – 1 half-spreads, so $\varphi \approx 1 + 0.8 = 1.8$ half-spreads.

B.4 Discount Factor

Suppose each investor i has private impatience parameter $\beta_i \in [0, 1]$, drawn independently from a given distribution $F(\beta)$. In practice, investors choose from a long menu of broker algorithms for placing and canceling orders, and their choices partially reveal their values of β_i .

In our model, investors choose just between midpoint pegs and market orders, implying a threshold, $\tilde{\beta}$, such that more patient investors (those with $\beta_i > \tilde{\beta}$) choose pegs and less patient investors choose market orders. Thus, given $\tilde{\beta}$, a fraction $\omega = 1 - F(\tilde{\beta})$ of the orders are transmitted as pegs.

Our steady state distribution of order imbalances (Proposition 3.1) implies a distribution of waiting times, and thus expected investor profits $\pi_i(\theta|\omega, \beta_i)$, for order types $\theta \in \{\text{peg}, \text{mkt}\}$. By maximizing over θ (choosing the preferred order type) we obtain a new threshold $\tilde{\beta}'$. The result is a map $M : [0, 1] \rightarrow [0, 1]$, $\hat{\beta} \mapsto \hat{\beta}'$.

Lemma B.1. *If the distribution F is continuous, then the mapping M , defined above, has a unique fixed point $\beta^* \in [0, 1]$.*

Proof sketch. Show that M is continuous and monotone decreasing, so the conclusion follows from the intermediate value theorem.

This result allows us to infer β^* from our calibration of ω^* and the other parameters: given vector (ω^*, φ, d) we use the equal profit condition for the marginal investor, Equation (3.9), to solve

$$\begin{aligned} \beta^* &= \frac{\varphi - 1}{\varphi - d - \omega^*(1 - d)} \\ &= \frac{1.8 - 1}{1.8 - 0.18 - 0.75 \times (1 - 0.18)} \\ &= 0.79. \end{aligned} \tag{B.6}$$

Substituting $\rho = 12.82$ (determined below) into the relation $\beta^* = \exp(-\delta/\rho)$ we arrive at $\delta = -\rho \log(\beta^*) \approx 3$.

B.5 Arrival Intensities

Table 3 reports the S&P 500 exchange traded fund (ticker SPY) transaction volume at IEX for the month of December, 2016. As we consider the asset in our model to be similar to a highly liquid asset such as the SPY, we use the volume statistics in Table 3 to calibrate the investor arrival intensity parameter, ρ . The total number of SPY shares traded by

	Other Nonroutable				Primary Peg			
	Hidden		Lit		Hidden		Lit	
	BBO	Mid	BBO	Mid	BBO	Mid	BBO	Mid
Agency Remover	489370	80279	2506659	0	0	0	0	0
Prop Remover	2262796	230826	1201430	0	0	0	0	0
Agency Adder	146504	24141	2030384	0	263913	0	0	0
Prop Adder	14655	1041	5978539	0	1882565	0	0	0
	Midoint Peg				Discretionary Peg			
	Hidden		Lit		Hidden		Lit	
	BBO	Mid	BBO	Mid	BBO	Mid	BBO	Mid
Agency Remover	0	1644581	0	0	0	591604	0	0
Prop Remover	0	565548	0	0	0	2200	0	0
Agency Adder	99915	1228941	0	0	1800154	2057015	0	0
Prop Adder	0	4802	0	0	3100	4490	0	0

Table 3: IEX SPY volume for December 2016 by order type and transaction price. Excludes routable orders and transactions in locked or crossed market conditions.

Agency Removers (across all order types) is 5,312,493 and the shares traded by Agency Adders via only Midpoint and Discretionary pegs is 5,186,025, resulting in a total volume of 10,498,518 shares by the equivalent of investors in our model. Since our model considers an investor arrival to be a unit transaction, and since a unit transaction at IEX is 100 shares, there were a total of $10,498,518/100 \approx 105,000$ investor arrivals during the 21 trading days, or $21 \times 6.5 \times 60 = 8190$ trading minutes of December, 2016, resulting in $\rho = 105,000/8190 = 12.82$ investor arrivals per minute or roughly 1 arrival every 4.68 seconds.

To calibrate ν we utilize SPY quotation data at Nasdaq, which, given its liquidity and overall market share, is a good surrogate for the SPY NBBO. Our sample covers the period 16 June – 11 September, 2014. There are 26,216,524 quotations in the 62-day period, which

comprises 1,450,800,000 milliseconds during trading hours, or approximately 1 quote every 55 milliseconds. Defining a jump as any midpoint price change of magnitude at least \$0.01 over the period of four quotations, or 220 milliseconds, resulted in an average of approximately 2,500 jumps per day, $\nu = 6.41$ jumps per minute, or one jump every 9.36 seconds.

Combining the values of ρ and ν , our baseline measures suggest $\frac{\nu}{\rho} \approx 0.5$, or that the intensity of value jumps is about half that of investor arrivals. In our model we typically work with the inverse ratio, $\frac{\rho}{\nu} \approx 2$.

B.6 Cost of Speed

At the time of this writing, one of the premier microwave transmission services, McKay Brothers LLC, offers low latency data services for 8 select ETFs (such as SPY) for \$3,100 per month. This translates to $\$3100/(8 \times 8190) = \0.047 or approximately $c = 10$ half-spreads per symbol, per minute.

C Institutional Information

C.1 Order Routing

In accordance with Regulation National Market System (Reg NMS), IEX routes orders to protected quotations at other exchanges when those quotations offer price improvement. The IEX router does this both at initial receipt of an order, and at periodic intervals for orders resting on the book. The latter feature is referred to as resweep. To be eligible for such protection, orders must be designated as “routable”, whereas “nonroutable” orders are sent directly to the IEX book and are not eligible for resweep.

The order book and router are distinct components of the IEX system. After passing through the initial 350 microsecond point-of-presence delay, nonroutable orders are sent directly to the IEX order book, whereas routable orders are sent to the router. The IEX order router then disseminates these latter orders to all national market systems (including

their own) following a proprietary routing table. Messages that are passed between the IEX order book and router are subject to an additional one-way 350 microsecond delay. As a result, routable orders that are sent to the IEX order book experience a cumulative delay of 700 microseconds before queuing behind other orders in the system. No additional delay is enforced between the IEX router and external exchanges.

As described in Section 2, routable orders constitute only 15% of IEX trading volume and represent traders that use the IEX router as an access point to the national market system. The remaining, nonroutable volume, represents trading interest intended to capture incentives of the IEX market design.

C.2 Pegged Order Types

Section 2 lists the three types of pegged orders at IEX. Midpoint pegs rest at the midpoint of NBBO, whereas primary pegs are booked in the hidden order queue one price increment (typically \$0.01) below (above) NBB (NBO), and are promoted to transact at NBB or NBO if sufficient trading interest arrives at those prices. Discretionary pegs synthesize the benefits of these first two: when entering the order book, they check the NBBO midpoint for contra-side interest, but in the absence of such interest, are pegged to NBB or NBO and are queued behind other hidden orders at those prices. Further, in the event that contra-side interest subsequently arrives at the NBBO midpoint, discretionary peg orders can be promoted to transact at the midpoint. If no such interest arrives, discretionary pegs are treated as typical hidden NBBO orders.

Table 1 shows that midpoint trading constitutes a little more than 60% of volume, discretionary peg trading accounts for 37% of volume and 89% of discretionary pegs are transacted at the midpoint. The implication is that midpoint volume is nearly evenly split between midpoint and discretionary pegs. Primary pegs and discretionary pegs transacted at BBO each account for 5% or less of reported volume. Thus, while there is a distinction between midpoint and discretionary peg orders in practice, nearly all discretionary peg orders transact at midpoint. For this reason, we reduce the decision space for order types in our model to a

simple midpoint peg.

Table 1 also reports small volume statistics for seemingly incongruous trades: (1) midpoint orders that transact at BBO and (2) hidden nonroutable orders (not pegs) that transact at midpoint. The first case occurs when midpoint pegs are booked with a limit price constraint which binds after subsequent movements in the NBBO. In such instances, an order that originally rested at midpoint might later rest and transact at BBO. The second case occurs under nuanced conditions where the NBBO is more than a single price increment wide or when the IEX BBO is wider than the NBBO (which may be a single increment). In such instances, the NBBO may coincide with the IEX midpoint or the hidden order at IEX may be subject to a special midpoint price constraint⁵ and later transact with contra-side orders at midpoint.

C.3 Crumbling Quote

The volume statistics for midpoint pegs in Table 1 show that proprietary firms are three times more likely to act as liquidity removers at midpoint (7.16% of volume) than as liquidity adders (2.12% of volume). This is indicative of opportunistic stale-quote arbitrage in advance of movements in the NBBO. Despite the fact that the IEX delay is intended to combat such exploitative activities, the company has reported an increase in anticipatory trading: midpoint quotes being removed at unfavorable prices immediately prior to changes in the NBBO (Bishop, 2017). This trading is almost certainly a result of improved probabilistic modeling of NBBO liquidity shifts by fast traders.

In an effort to further protect pegged orders from adverse selection, IEX has developed the

⁵When the IEX BBO is wider than the NBBO and a nonroutable hidden order enters the order book with a limit that would otherwise be passed on to another exchange displaying NBBO, the order is booked at the NBBO midpoint and may be promoted to transact at the NBBO at a later time. For example, suppose the NBBO is $\$10.00 \times \10.01 and the IEX order book is $\$10.00 \times \10.02 when a nonroutable hidden buy order arrives with a limit of $\$10.01$. The order will be booked at $\$10.005$ and will later transact at $\$10.01$ if a sell limit arrives at that price. Alternatively, it may transact with midpoint pegs, discretionary pegs, or market orders at midpoint.

“crumbling quote signal”: a model that forecasts changes in the NBBO (the crumbling quote) and temporarily prevents primary and discretionary peg orders from exercising discretion at their potentially more aggressive prices in order to minimize their exposure to anticipatory traders. That is, when the crumbling quote signal is on, discretionary pegs do not transact at midpoint and primary pegs do not transact at BBO. Midpoint pegs do not receive protection from the crumbling quote signal.

While we view the crumbling quote signal as an important innovation to the IEX market design, we have excluded it from our model in order to focus attention on the primary role of the speed bump and its interaction with pegged order types. We consider study of the crumbling quote signal, however, to be a valuable direction for future work.

References

- Bishop, A. (2017), “The Evolution of the Crumbling Quote Signal,” *IEX White Paper*, 1–30.
- Budish, E., Cramton, P., and Shim, J. (2015), “The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response,” *The Quarterly Journal of Economics*, 130, 1547–1621.
- Fox, M. B., Glosten, L. R., and Rauterberg, G. V. (2015), “The New Stock Market: Sense and Nonsense,” *Duke Law Journal*, 65, 191–277.
- Lewis, M. (2015), *Flash Boys: A Wall Street Revolt*, New York: W. W. Norton & Company.
- Pisani, B. (2016), “SEC gives its blessing to the IEX’s ‘speed bump’ trading,” <http://www.cnbc.com/2016/06/17/sec-gives-its-blessing-to-the-iexs-speed-bump-trading.html>.
- Wah, E., Hurd, D. R., and Wellman, M. P. (2015), “Strategic Market Choice: Frequent Call Markets vs. Continuous Double Auctions for Fast and Slow Traders,” *Working paper*.