JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 56, No. 1, Feb. 2021, pp. 93–121

doi:10.1017/S0022109019000814

Trading in Fragmented Markets

Markus Baldauf and Joshua Mollner*

Abstract

We study fragmentation of equity trading using a model of imperfect competition among exchanges. In the model, increased competition drives down trading fees. However, additional arbitrage opportunities arise in fragmented markets, intensifying adverse selection. Due to these opposing forces, the effects of fragmentation are context dependent. To empirically investigate the ambiguity in a single context, we estimate key parameters of the model with order-level data for an Australian security. According to the estimates, the benefits of increased competition are outweighed by the costs of multi-venue arbitrage. Compared with the prevailing duopoly, we predict the counterfactual monopoly spread to be 23% lower.

I. Introduction

Equity markets have become increasingly fragmented over the past 10 to 15 years. During that period, a surge of new exchanges and alternative trading systems entered and attracted volume that had previously been concentrated on primary venues. Illustrative is the decline in the New York Stock Exchange's (NYSE's) share of the volume of NYSE-listed stocks, from 82% in June 2004 to 27% in June 2018 (Angel, Harris, and Spatt (2015), CBOE (2018)).

This proliferation of trading venues and the accompanying dispersion of trades were actively encouraged by the U.S. Securities and Exchange Commission (SEC) through Regulation National Market System (NMS) to promote "vigorous

*Baldauf (corresponding author), baldauf@mail.ubc.ca, University of British Columbia Sauder School of Business; Mollner, joshua.mollner@kellogg.northwestern.edu, Northwestern University Kellogg School of Management. We are indebted to our advisors Timothy Bresnahan, Gabriel Carroll, Jonathan Levin, Paul Milgrom, and Monika Piazzesi. We also thank James Angel, Lanier Benkard, Philip Bond, Jean-Edouard Colliard (discussant), Carole Comerton-Forde, Alan Crawford, Darrell Duffie, Liran Einav, Thierry Foucault, Lorenzo Garlappi, Lawrence Glosten (discussant), Duncan Gilchrist, Terrence Hendershott, Christiaan Hogendorn, Yingxiang Li, Katya Malinova (discussant), Albert Menkveld, Peter Reiss, Alvin Roth, Georgios Skoulakis, Laura Tuttle, Brian Weller, Mao Ye, and seminar participants, as well as various industry experts, for valuable comments. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant OCI-1053575. Baldauf acknowledges financial support from the Stanford Institute for Economic Policy Research Kohlhagen Fellowship Fund and the Social Sciences and Humanities Research Council (SSHRC) grant 430-2018-00100. Mollner acknowledges financial support from the Stanford Institute for Economic Policy Research Kapnick Fellowship Program.

competition among markets" (SEC (2005)). The intuition that investors benefit from competition should resonate with any economist. But there is also a drawback to spreading out trade across markets: Fragmentation may create opportunities for fast traders to extract rents by arbitraging across venues. We capture this trade-off in a tractable and estimable multi-exchange model of limit order book trading. We show that depending on factors such as the arrival rate of information, the strength of the private transaction motives of investors, their arrival rate to the market, and the strength of market frictions, the introduction of new exchanges can either increase or decrease the transaction costs faced by investors.

We then employ the Australian market for an empirical application of the model. Using order-level data to estimate the parameters of our model, we find that investors are worse off in the prevailing duopoly than they would be under a monopoly exchange. We find support for this conclusion from a natural experiment in which a technical issue forced one of the two Australian exchanges to shut down for a day.

Section III develops the model. A security's shares are traded in limit order books on multiple exchanges. Its fundamental value is public information and evolves stochastically. Exchanges operate trading platforms and earn profits from trading fees. High-frequency traders (HFTs) may trade for profit through arbitrage or liquidity provision. Investors arrive stochastically with private trading motives and are differentiated along two dimensions. First, they differ in the strength of their private motive to transact. Second, they differ in their propensity to substitute among venues for a given price difference. In particular, we do not require every investor to trade at the exchange that offers the best price, which may be thought of as a reduced-form representation of some market friction.[1] Such frictions might stem from several sources, including the difficulty of monitoring prices in real time and agency problems between an investor and the broker who routes the investor's orders.

In Section IV, we analyze the model's equilibrium to explore the effect of fragmentation on transaction costs as measured by the cum-fee spread: the quoted bid–ask spread plus twice the take fee (the take fee is levied by the exchange on the party that initiates the trade). In practice, this quantity is a significant component of the transaction costs of equity trading, and in the model, it is a sufficient statistic for welfare. Henceforth, we use "spread" as shorthand for the cum-fee spread unless otherwise specified.

Two forces give rise to a spread in this model: i) the market power of exchanges and ii) adverse selection stemming from a race to react to information. Regarding the second force, although information is publicly observable, adverse selection arises if a liquidity provider is unable to cancel mispriced quotes before they are exploited by arbitrageurs. A change in the number of venues affects the magnitude of each of these two forces and consequently affects the

[1]Our motivation for considering a model with this feature is that not every trade takes place at the best price in our data. Indeed, 8.9% of trades that occur when exchanges offer different prices take place at the exchange with the worse price (cf. Section V.D). Moreover, the estimated model implies a cross-price elasticity (percentage change in investor demand at one exchange given a 1% increase in the other exchange's spread) of only 0.8 (cf. Section VI.B). Both figures suggest the presence of significant market frictions.

equilibrium spread through two opposing channels. First, an increase in the number of exchanges reduces the spread through the "competition channel." Intuitively, exchanges have less market power when there are more of them. Consequently, each charges a lower trading fee, which results in a lower spread, ceteris paribus. Second, an increase in the number of exchanges raises the spread through the "exposure channel." Intuitively, investor order flow is more fragmented with more exchanges, forcing liquidity providers to deepen the aggregate book in response,[2] which exposes them to more adverse selection and results in a larger spread, ceteris paribus. Theory is silent on the net effect of fragmentation because either the competition channel or the exposure channel can dominate.

This theoretical ambiguity is consistent with the diverse findings that have been reached by an old and extensive empirical literature, of which some articles report a positive association between fragmentation and liquidity while others report the reverse. Most of those studies leverage variation in market structure to determine the effects of fragmentation in a given setting, but a different approach is needed to ascertain the effects of fragmentation in settings where identifying variation is unavailable, and to that end, our approach may be a viable alternative. To illustrate, we use an Australian security for an empirical application.

In Section V, we discuss the Australian market and our data. This setting is a natural fit for the model because it is particularly simple, featuring just two exchanges: the Australian Securities Exchange (ASX) and Chi-X Australia (Chi-X). For the analysis, we use order-level data on STW, an exchange-traded fund (ETF) tracking the Standard & Poor's (S&P)/ASX 200 index. The sample covers trading on both exchanges over 80 trading days in 2014.

In Section VI, we turn to estimation. The parameters are identified via the response of order flow to variation in prices at ASX and Chi-X, together with the average level of the spread. We then use the estimated model to conduct counterfactual analysis by comparing the current duopoly outcome to what would prevail under a monopoly. We find that the counterfactual monopoly spread would be 23% lower than the duopoly spread of 2.88¢. Thus, the exposure channel dominates the competition channel in the case of STW. Although these findings pertain to only one security in Australia, they do emphasize that unbridled competition among trading platforms can be harmful to investors.

In Section VII, we conduct a separate, out-of-sample analysis of a natural experiment in which Chi-X experienced a technical issue and halted its trading for the remainder of the day, leaving ASX as the only exchange in operation. The STW spread on ASX is significantly lower on the day of the Chi-X shutdown than on the surrounding days, as well as relative to an unaffected control group. This effect is consistent with the exposure channel postulated by the model, and its magnitude is in line with the estimates. Thus, this event provides additional support for the model's prediction that STW spreads would be lower under a monopoly than the prevailing duopoly. Moreover, the model's success in forecasting the result

[2]Fragmentation also increases aggregate depth in the models of Dennert (1993) and van Kervel (2015), due to similar forces. Moreover, that aggregate depth grows with the number of trading venues is a stylized fact that has been empirically detected both in general (Boehmer and Boehmer (2003), Fink, Fink, and Weston (2006), and Foucault and Menkveld (2008)) and in the specific context of Australia (He, Jarnecic, and Liu (2015), Aitken, Chen, and Foley (2017)).

of this natural experiment leads us to speculate that it may be a useful methodological tool for predicting the effects of fragmentation more broadly. Such a tool might be particularly useful to regulators, who often need to decide rule changes that will shape fragmentation. In addition to regulating the entry of new venues, a related issue currently in debate is whether issuers of thinly traded stocks should be able to suspend unlisted trading privileges for their securities, thereby reducing the number of exchanges at which they are traded (U.S. Treasury (2017), SEC (2019)).

II. Related Literature

A. Theory

An early model of competition among electronically linked limit order book markets is that of Glosten (1994). He demonstrates that in an idealized, frictionless setting, the liquidity of the aggregate market is invariant to the degree of fragmentation.[3] Our model, in contrast, allows for frictions that may prevent investors from trading at the best available price, and thus Glosten's invariance result does not apply. Rather, liquidity depends on the degree of fragmentation in two ways: i) the competition channel, via which fragmentation intensifies competition on trading fees, thereby improving liquidity, and ii) the exposure channel, via which fragmentation increases the number of arbitrage opportunities, thereby amplifying adverse selection and harming liquidity. In consequence, our article connects to two branches of the literature on the relationship between liquidity and fragmentation.

First, our article relates to models in which fragmentation may improve liquidity by enhancing competition. Earlier models focus on competition among market makers (Mendelson (1987), Dennert (1993), Bernhardt and Hughson (1997), and Biais, Martimort, and Rochet (2000)). However, because modern markets permit virtually free entry into market making, more recent work focuses on competition among venues themselves (Colliard and Foucault (2012), Pagnotta and Philippon (2018), and Chao, Yao, and Ye (2019)). Similar to our competition channel, they find that fragmentation may induce exchanges to lower their fees. Second, our article relates to models in which fragmentation may harm liquidity by altering the opportunities available to informed traders (Chowdhry and Nanda (1991), Dennert (1993)). Similar to our exposure channel, they find that adverse selection intensifies when more markets operate in parallel. Our primary theoretical contribution lies in developing a framework that combines these two forces into a single tractable and estimable model.

Also connected are other articles that, although they do not deal with fragmentation, model trading in the presence of adverse selection from privately informed traders (e.g., Copeland and Galai (1983), Glosten and Milgrom (1985), and Kyle (1985)), as well as work illustrating that similar forces arise in limit order books even when information is public (e.g., Foucault (1999), Budish, Cramton, and Shim (2015), Aït-Sahalia and Sağlam (2017a), (2017b)). We build

[3]Budish, Lee, and Shim (2019) prove a similar result in a setting where exchanges strategically set trading fees and charge for speed technology (i.e., co-location and data access).

primarily upon the model of Budish et al. (2015), which we extend in several ways. For instance, exchanges in our model strategically set trading fees in competition for investors, and these fees constitute a source of transaction costs in addition to adverse selection.

B. Empirics[4]

The effects of fragmentation in financial markets have been studied using a diverse set of empirical strategies, including cross-sectional and panel regression; matched-sample analysis; and studies of entry events, consolidation events, cross-listing events, and rule changes. Taken together, the findings of this literature are extremely heterogeneous:

- Many of these articles find positive associations between fragmentation and liquidity (Branch and Freed (1977), Hamilton (1979), Neal (1987), Cohen and Conroy (1990), Battalio (1997), Mayhew (2002), Weston (2002), Boehmer and Boehmer (2003), De Fontnouvelle and Harris (2003), Fink et al. (2006), Nguyen, Van Ness, and Van Ness (2007), Foucault and Menkveld (2008), Chlistalla and Lutat (2011), O'Hara and Ye (2011), and Menkveld (2013)). Particularly relevant to our article are the articles by He et al. (2015) and Aitken et al. (2017), who also study the Australian market, finding liquidity to improve on average in conjunction with the 2011 entry of Chi-X.

- Others find negative associations between fragmentation and liquidity (Bessembinder and Kaufman (1997), Arnold, Hersch, Mulherin, and Netter (1999), Amihud, Lauterbach, and Mendelson (2003), Hendershott and Jones (2005), Bennett and Wei (2006), Gajewski and Gresse (2007), Nielsson (2009), and Bernales, Riarte, Sagade, Valenzuela, and Westheide (2017)).

- Some others find an inverted-U relationship, in which liquidity is maximized under moderate degrees of fragmentation (Degryse, de Jong, and van Kervel (2015), Boneva, Linton, and Vogt (2016)). Similarly, Haslag and Ringgenberg (2017) also find a nuanced association: Fragmentation benefits liquidity for large stocks but harms it for small stocks.

The diversity of these findings indicates that the effects of fragmentation are highly context dependent. We contribute by proposing a model to explain the role of context. Moreover, the model may also serve as a tool for predicting the effect of fragmentation in a given context.

III. Model

A single security is traded at one or more exchanges by two categories of traders: investors and HFTs. The timing is as follows: First, exchanges set make and take fees. Second, trading occurs over an interval of continuous time $[0, T]$.

[4]See Appendix B.A in the Supplementary Material for a more detailed description of the empirical articles cited here.

A. Security

The fundamental value of the security at time t, v_t, is public information, and it evolves as a compound Poisson jump process with arrival rate $\lambda_j \in \mathbb{R}_+$. Positive and negative jumps of size $\sigma \in \mathbb{R}_+$ occur with equal probability.

B. Exchanges

X exchanges each operate a separate limit order book with continuous prices and divisible shares. Order types include limit, cancellation, immediate-or-cancel, and market orders. Orders are processed sequentially in the usual way. If multiple orders arrive simultaneously, then ties among traders are broken uniformly at random (e.g., by random latency).

Exchanges are horizontally differentiated, which we model by assuming that they are metaphorically located at equally spaced points around a circle of unit length, as in Salop (1979).[5] The location of exchange x is denoted l_x. Exchange x sets make and take fees $\tau_{x,\text{make}} \in \mathbb{R}$ and $\tau_{x,\text{take}} \in \mathbb{R}$, which are collected, respectively, from the passive and aggressive parties of each trade that occurs on the exchange. Negative fees correspond to rebates. Trading fees are chosen once and for all before trading commences at time 0.

C. Investors

Investors arrive at a Poisson rate $\lambda_i \in \mathbb{R}_+$ with a two-dimensional type $(\tilde{l}, \tilde{\theta})$. The first component, $\tilde{l}$, is an independent draw from $U[0,1]$ and denotes a position on the aforementioned circle. The second component, $\tilde{\theta}$, is an independent draw from $U[-\theta, \theta]$ and denotes a private value for trading one share of the security.

Let $b_{x,t}$ and $a_{x,t}$ denote, respectively, the cum-fee bid and ask at exchange x at time t. Investors are restricted to market orders. By submitting a market order for $y \in \{-1,0,1\}$ shares to exchange $x \in \{1,\ldots,X\}$, an investor who arrives at time t obtains utility

$$(1) \qquad u_t(y,x|\tilde{\theta}) \quad = \quad \mathbb{1}(y=1)(v_t + \tilde{\theta} - a_{x,t}) + \mathbb{1}(y=-1)(b_{x,t} - v_t - \tilde{\theta}).$$

Yet investors do not necessarily act to maximize utility, although the model permits that case. Instead, the aforementioned investor chooses y and x to maximize

$$(2) \qquad \hat{u}_t(y,x|\tilde{l},\tilde{\theta}) \quad = \quad u_t(y,x|\tilde{\theta}) - 2\alpha \cdot d(\tilde{l}, l_x)^2,$$

where $d(l,l') = \min(|l-l'|, 1-|l-l'|)$ yields the distance between two points on the circle.[6] That investors do not always act to maximize their utility could be thought of as the result of an unmodeled market friction. Investors are heterogeneously affected by this friction, owing to their different locations on the circle. The parameter α governs the extent of this friction. One possibility is $\alpha = 0$,

[5] An exchange's location should be interpreted metaphorically as representing its nonprice characteristics (e.g., available order types, latency, ownership). Note that we model neither the entry game of exchanges nor their location game but, rather, solve for equilibrium under a fixed number of equally spaced exchanges.

[6] For the baseline analysis, we assume that investors do not split their orders, with each investor trading at a single exchange and a single point in time. Nevertheless, Appendix F.D in the Supplementary Material demonstrates that the model's equilibrium remains intact even if investors can split their orders across exchanges.

in which case frictions vanish and investors are utility maximizers. In the other extreme, as α grows large, investors become increasingly likely to choose the exchange closest to them on the circle, without regard for the terms of trade.

Such market frictions might stem from several sources: difficulties associated with monitoring prices in real time, the agency problem between an investor and the broker who routes his orders, and so forth. Of course, the strength of such frictions could be modulated by regulation (e.g., trade-through protection, as in the United States) and enforcement thereof.

As noted, investor/broker conflicts of interest constitute a potential source of frictions. For example, a broker might be tempted to route an order to an exchange with an inferior price if it would pay him (the broker) a rebate.[7] This explanation is less applicable to this article's eventual empirical application in Australia, where there are no such rebates. More germane to that setting are the following possibilities: i) Checking prices at all exchanges might require a broker to make technological investments or costly effort; ii) a broker might have an ownership stake in a certain exchange, which could introduce a financial incentive to route orders there;[8] or iii) a broker might have a nonpecuniary preference for a certain exchange, perhaps due to its speed and infrastructure, order types offered, or relationship capital. These latter explanations appear especially plausible in the context of Australia, which, unlike the United States, does not have an order protection rule (cf. Appendix B.B in the Supplementary Material).

D. High-Frequency Traders

There is an infinite number of HFTs. Each trades to maximize her own profits.

Assumptions. The parameters of the model satisfy the following conditions:

A1. $\lambda_j \leq \lambda_i / X$

A2. $\sigma > \theta$.

A3. $\lambda_i \left(1 - \frac{1}{\theta}\frac{\Sigma}{2}\right)\frac{\Sigma}{2} \geq \lambda_j X \left(\sigma - \frac{\Sigma}{2}\right)$, where Σ is defined in terms of the parameters as follows:[9]

$$\Sigma \equiv \begin{cases} \theta\left(1 + \frac{\lambda_j}{\lambda_i}\right) & \text{if } X = 1 \\ \theta + \frac{4\alpha}{X^2} - \sqrt{\theta^2 + \frac{16\alpha^2}{X^4} - \frac{8\alpha\theta\lambda_j}{X\lambda_i}} & \text{if } X \geq 2. \end{cases}$$

[7] A related note is that rebates influence how brokers route retail nonmarketable limit orders (Battalio, Corwin, and Jennings (2016)) and institutional orders (Battalio, Hatch, and Sağlam (2018)), sometimes at the cost of execution quality.

[8] Consistent with the operation of such frictions, Anand, Samadi, Sokobin, and Venkataraman (2019) show that brokers obtain a lower execution quality when routing client orders to alternative trading systems that they own.

[9] As Lemma 1 i) (in the Supplementary Material) establishes, A1 implies $\Sigma \in \mathbb{R}$, so this is a well-defined inequality.

https://doi.org/10.1017/S0022109019000814 Published online by Cambridge University Press

A1 requires the per-exchange arrival rate of investors to exceed the arrival rate of jumps in the value of the security. Thus, it ensures that episodes of adverse selection are sufficiently infrequent. A2 requires the magnitude of information about the security to exceed the magnitude of investors' private trading motives. Thus, it implies that any increase in the spread will crowd out more liquidity-based trades from investors than information-based trades from HFTs.[10] A3 requires that if Σ is the cum-fee spread prevailing at all exchanges, then the resulting payments by investors exceed the arbitrage profits of HFTs. To understand why it is needed, an analogy is to oligopolistic price competition with fixed costs, where a symmetric pure strategy equilibrium exists only if i) fixed costs are not too large, ii) price competition is not too intense, and iii) the market is not too small. In the current setting, adverse selection plays the role of fixed costs. The parameters σ, λ_j, and X determine the magnitude of adverse selection, and A3 prevents them from being too large; α determines the intensity of competition among exchanges, which cannot be too small; and λ_i and θ determine the available gains from trade, which also cannot be too small.

IV. Equilibrium

This section describes the subgame perfect Nash equilibrium (SPNE) of the model, then discusses how the equilibrium depends on the parameters. Throughout, the focus is on the *cum-fee spread*: the spread plus twice the take fee. This is the appropriate quantity because it measures the transaction costs borne by investors. As such, it is a sufficient statistic for the welfare implications of the model; an increase affects welfare in two ways: i) gains from trade decline because some marginal investors may cease to trade, and ii) there are transfers away from the inframarginal investors who continue to trade.

A. Equilibrium Description

We solve the model by backward induction: first characterizing equilibrium trader behavior for given fees, then taking those outcomes as given to identify equilibrium fee choices. An intuitive description of equilibrium strategies follows, and we defer to the proofs of Propositions 1 and 2 in Appendix A of the Supplementary Material for a complete treatment.

We first discuss the case of a monopoly exchange, extending to the oligopoly case later in the section. In equilibrium, HFTs sort into two roles, as in Budish et al. (2015). One plays the role of "liquidity provider," establishing quotes of one share at both the bid and the ask and maintaining them so that the midprice tracks the value of the security. The remaining HFTs play the role of "sniper," attempting to trade whenever information arrivals create mispricings in the quotes of the liquidity provider. They also ensure the spread is such that the liquidity provider earns zero profits, as in Bertrand competition.

After each jump, HFTs race to react: the liquidity provider to cancel her mispriced quotes and the snipers to exploit them. Each race results in a tie, which is broken uniformly at random. Because an infinite number of HFTs assume the

[10]Similar assumptions also appear in Glosten (1994) (Assumption 2) and Biais et al. (2000) (that $v'(\theta) \geq 0$).

sniper role in equilibrium, the liquidity provider loses each race.[11] Thus, sniping is one cost in the liquidity provider's zero-profit condition. A second cost is the make fee of the exchange. Revenue derives from investors: Each transacts if his private value $\tilde{\theta}$ exceeds half the spread, and the liquidity provider earns that half spread from every such trade.

This zero-profit condition determines the spread. Inducting backward, the monopolist exchange sets make and take fees to maximize expected revenue. The key trade-off is that although higher fees generate more revenue per trade, they also induce larger spreads, which crowd out some investor trades and reduce volume. Proposition 1 characterizes the (cum-fee) spread induced by the optimal fee choice.

Proposition 1. With a single exchange ($X=1$), there exists an SPNE with spread

$$s^* = \theta\left(1+\frac{\lambda_j}{\lambda_i}\right). \tag{3}$$

Equation (3) illustrates the two sources of the spread in this model: adverse selection and exchange market power. First, liquidity providers use the spread to offset the costs of adverse selection, which arises when they lose the race to react to public information. Thus, the spread depends on the relative arrival rates of information and investors, λ_j/λ_i, which governs the degree of adverse selection. Second, the exchange takes into account that fewer investors will trade at a wider spread. Thus, the spread also depends on θ, which governs the price elasticity of demand. Indeed, absent adverse selection, it satisfies the classic Lerner condition, which equates the markup to the inverse of the demand elasticity.

The oligopoly case is similar to the monopoly case, with the main difference being that investors choose not only whether to trade but also where. As before, HFTs sort into two roles. One per exchange plays the role of liquidity provider, maintaining quotes of one share at the bid and one share at the ask, and the remaining HFTs play the role of sniper.

Whereas a monopolist exchange is constrained only by the own-price elasticity of investors (higher fees might lead investors not to trade), an oligopolist exchange must also consider cross-price elasticities (higher fees might lead investors to trade at other exchanges instead). Taking these trade-offs into account, exchanges set their trading fees in a simultaneous-move game. We focus our analysis on the symmetric equilibrium of this game, and Proposition 2 characterizes the resulting equilibrium (cum-fee) spread.

Proposition 2. With multiple exchanges ($X\geq 2$), there exists an SPNE with spread

$$s^* = \theta+\frac{4\alpha}{X^2}-\sqrt{\theta^2+\frac{16\alpha^2}{X^4}-\frac{8\alpha\theta\lambda_j}{X\lambda_i}}. \tag{4}$$

[11]In the model, the liquidity provider races against an infinite number of other HFTs and therefore loses the race with probability one. If there were instead N HFTs, the liquidity provider would lose with probability $(N-1)/N$, so that her zero-profit condition would be approximately the same provided N is large. And indeed, HFTs are quite numerous in practice. The Australian market is representative: 550 HFTs were active in 2012 (ASIC (2013)).

As in the monopoly case, the oligopoly spread is influenced by θ and λ_j/λ_i. Yet in addition, there is now also a role for X, the number of exchanges, as well as α, which governs the strength of market frictions affecting the exchange choices of investors. One reason for the presence of these two parameters is that they interact with the prevailing spreads to determine the number of investors who choose to trade at a given exchange: i) If X increases, then each exchange obtains a smaller share of investor trades at equal spreads, and ii) if α increases, then each exchange's share of investor trades becomes less responsive to its spread. As a result, these parameters affect the trade-offs exchanges face when they set trading fees and, in turn, the equilibrium spread as well.

The remainder of this section describes the derivation of the monopoly and oligopoly spreads via backward induction from the zero-profit condition of the liquidity provider. The discussion focuses on the ask side of the book, but the bid-side version is analogous. At any instant, one of two things may affect the ask-side profits of a liquidity provider: An investor may arrive with a motive to buy, or the value of the security may jump upward. We begin with investor arrivals. Investors with a motive to buy arrive at the rate $\lambda_i/2$. In the case of an oligopoly, when the cum-fee ask is a_x on an exchange x and a_{-x} on the others, x is the preferred exchange of such an investor with a probability of $\left[(X/2\alpha)(a_{-x}-a_x)+1/X\right]_0^1$;[12] in the case of a monopoly, it is always the preferred exchange. Conditional on the investor preferring exchange x, the investor trades with a probability of $\left[1-(a_x-v)/\theta\right]_0^1$. Thus, investors buy at exchange x at the rate $(\lambda_i/2)\left[1-(a_x-v)/\theta\right]_0^1$ in the case of a monopoly and

$$(5)\qquad \frac{\lambda_i}{2}\left[\frac{X}{2\alpha}(a_{-x}-a_x)+\frac{1}{X}\right]_0^1\left[1-\frac{a_x-v}{\theta}\right]_0^1$$

in the case of an oligopoly. From each of these trades, the liquidity provider earns a_x-v and must pay the make fee $\tau_{x,\text{make}}$ to the exchange. Next, we consider jumps in the value of the security. Upward jumps arrive at the rate $\lambda_j/2$. Conditional on such a jump occurring, the liquidity provider at exchange x will lose $\sigma+v-a_x$ to a sniper and must also pay the make fee $\tau_{x,\text{make}}$ to the exchange. Combining all this, the zero-profit conditions that determine the liquidity provider's ask are, respectively for the monopoly and oligopoly cases, as follows:

$$(6)\qquad \frac{\lambda_i}{2}\left[1-\frac{a_x-v}{\theta}\right]_0^1(a_x-v-\tau_{x,\text{make}}) \quad = \quad \frac{\lambda_j}{2}(\sigma+v-a_x+\tau_{x,\text{make}}),$$

$$(7)\qquad \frac{\lambda_i}{2}\left[\frac{X}{2\alpha}(a_{-x}-a_x)+\frac{1}{X}\right]_0^1\left[1-\frac{a_x-v}{\theta}\right]_0^1(a_x-v-\tau_{x,\text{make}}) \quad = \quad \frac{\lambda_j}{2}(\sigma+v-a_x+\tau_{x,\text{make}}).$$

Conditional on exchange x setting the fees $\tau_{x,\text{make}}$ and $\tau_{x,\text{take}}$, the liquidity provider on exchange x quotes to satisfy the appropriate zero-profit condition. That is, if a_x is the zero-profit cum-fee ask, then the liquidity provider sets the quoted ask $\hat{a}_x=a_x-\tau_{x,\text{take}}$.

[12] We use the notation $[\cdot]_0^1$ to denote truncation to the unit interval: $[x]_0^1\equiv\max(0,\min(1,x))$.

https://doi.org/10.1017/S0022109019000814 Published online by Cambridge University Press

Taking this behavior of liquidity providers as given, each exchange x sets fees to maximize its profits, which are the product of $\tau_{x,\text{make}} + \tau_{x,\text{take}}$ and the volume it processes. The resulting equilibrium spread is as described in Proposition 1 for the case of a monopoly. For oligopolies, we focus on symmetric equilibria, in which the same total fee is set by each exchange. In such equilibria, cum-fee quotes are identical across exchanges, and trades at prices inferior to those available elsewhere never occur on path. Nevertheless, such trades might occur off path if exchanges were to set different total fees, so that the resulting spreads would differ as well. The parameter α determines the rate at which such trades would take place in those off-path events, and through that, it plays an important role in determining the equilibrium fee and, in turn, the equilibrium spread, which is characterized by Proposition 2.

B. The Effect of Fragmentation

In the model, it is theoretically ambiguous how the spread (and hence welfare) depends on the number of exchanges. The ambiguity is caused by two opposing channels.

Fragmentation may reduce the spread through the *competition channel*. Intuitively, exchanges have less market power when they have more competitors and must therefore reduce their trading fees to retain investors. All else equal, lower fees induce a lower spread.[13]

But fragmentation may raise the spread through the *exposure channel*. First observe that frictions fragment the order flow of investors: For a given profile of quotes, some investors might trade at one exchange, whereas others would opt for another. Liquidity providers cannot predict where the next investor will seek to trade, and to cater to him, they must therefore quote at every exchange, and thus by aggregating across venues, they quote more depth than the investor will actually demand. Quoting less than one share at any given exchange would mean foregoing some profitable trades with investors. Next, consider an increase in the number of exchanges. Investor order flow becomes even more fragmented, and aggregate depth increases further. In the model, depth in fact grows linearly. Therefore, whenever the security value moves away from current prices, more shares are exposed to the resulting mispricing, amplifying adverse selection. All else equal, the spread rises as liquidity providers quote wider to compensate.

We illustrate with two cases of the model. First consider the limit as α diverges to infinity, which is to say that investors do not condition their choice of exchange on prices. For both the monopoly and oligopoly cases, the spread is $s^* = \theta\left(1 + X\lambda_j/\lambda_i\right)$, which is increasing in X. The reason is that if investors do not respond to prices, then multiple exchanges are a collection of isolated monopolists. Yet additional exchanges provide snipers with more opportunities to trade on a given piece of information, which increases adverse selection. Intuitively, the competition channel is shut down, so the exposure channel dominates.

[13]Note that equation (4) in Proposition 2 suggests that the spread converges to 0 as X diverges. But this should *not* be interpreted to mean that the competition channel always dominates in the limit. The reason is that A3 may be violated at large values of X, and we would not expect the symmetric equilibrium characterized by the proposition to prevail in such cases. Rather, we might expect an asymmetric outcome in which a subset of the exchanges exit until A3 can be satisfied.

Second, consider the case in which $\lambda_j = 0$, which is to say that the fundamental security value is constant. In that case, the monopoly spread is θ, which exceeds the duopoly spread of $\theta + \alpha - \sqrt{\theta^2 + \alpha^2}$, and the spread decreases still further as X increases beyond 2. The reason is that adverse selection does not increase with the number of exchanges because no adverse selection exists when information never arrives. Yet additional exchanges intensify price competition, resulting in smaller trading fees and hence smaller spreads. Intuitively, the exposure channel is shut down, so the competition channel dominates.

Our empirical application primarily focuses on the comparison between monopoly and duopoly, for which we obtain the following corollary of Propositions 1 and 2:

Corollary 3. $s^*_{\text{monopoly}} \leq s^*_{\text{duopoly}}$ if and only if

$$\theta \left(\frac{\lambda_j^2}{\lambda_i^2} - 1 \right) + 2\alpha \left(\frac{\lambda_j}{\lambda_i} \right) \geq 0.$$

An implication is that a given shock to fragmentation might produce differing effects in the cross section, raising spreads for some securities and reducing them for others.[14] Security-specific policies on fragmentation might therefore be preferable to a one-size-fits-all approach.

C. Comparative Statics

The spread varies monotonically in three of the remaining parameters: α, λ_i, and λ_j.[15] The following result is standard and intuitive:

Proposition 4. The equilibrium spread is i) weakly increasing in α, ii) weakly decreasing in λ_i, and iii) weakly increasing in λ_j.

The parameter α determines the strength of market frictions that distort the exchange choices of investors. If these frictions become stronger (an increase in α), then fee competition is muted, and the spread increases as a result. The arrival rates of investors and information, λ_i and λ_j, affect transaction costs by modulating adverse selection. If investors arrive more frequently (an increase in λ_i), then the liquidity provider faces less adverse selection, since she trades with relatively more investors and fewer snipers. She therefore sets a smaller spread. The reverse is true if information arrives more frequently (an increase in λ_j).

V. Empirical Application

We employ the Australian market for an empirical application of the model. This section lays the groundwork for that exercise by discussing industry background, describing the data sets, and defining the variables constructed from the data for use in estimation.

[14]The corollary also has more nuanced implications. For example, because the condition exhibits single crossing in λ_i, the corollary indicates that, ceteris paribus, fragmentation is more likely to be beneficial when λ_i is higher. Thus, if large-cap stocks attract more investors than their small-cap counterparts, this result might provide a theoretical foundation for the empirical results of Haslag and Ringgenberg (2017).

[15]There is, however, no monotone comparative static with respect to θ, which determines the strength of the private transaction motive of investors.

A. Industry Background

Two exchanges currently operate in Australia: ASX and Chi-X. The baseline make and take fees charged by ASX are each 0.15 basis points (bps) of the value of the trade (ASX (2016)), with larger fees applying to certain advanced order types. For Chi-X, make and take fees are, respectively, 0.06 and 0.12 bps of the traded value (Chi-X (2011)). Over the main sample period, the daily value of trades in the Australian cash market (i.e., equity, warrant, and interest rate market transactions) averaged AUD 4.8 billion (ASX (2014)), or roughly 2% of the U.S. market at the time.

The Australian market is a natural fit for the model because it is particularly simple and self-contained. In particular, i) there are just two exchanges, with independent ownership; ii) there are limited overlaps with foreign markets in terms of trading hours and securities; and iii) off-exchange trading is less relevant than in many other jurisdictions (e.g., the U.S.), largely due to a minimum price improvement rule and a prohibition against payment for order flow. Additional details regarding the Australian market are discussed in Appendix B.B of the Supplementary Material.

Our analysis focuses on a single security: the ETF STW,[16] which is a natural object of focus for three reasons. First, STW is a very significant security, not only because it is Australia's largest ETF but also because it tracks the benchmark index for the Australian market. Second, STW's broad exposure parallels the model's assumption that information is purely public. Although certain traders may be likely to possess private information about individual stocks, such information is relatively less significant for broad composites.[17] Third, STW's relatively large average spread parallels the model's assumption that prices are continuous. For many other thickly traded securities, the bid–ask spread is often one tick, in which case constraints imposed by discrete prices loom large. But discrete prices are less salient if the spread is larger. Fourth, STW is not internationally cross-listed. Because foreign and domestic fragmentation likely differ in their effects, international cross-listings would complicate the interpretation of our findings.

Thus, our framework seems reasonably appropriate for modeling the trading of STW in the Australian market, but its suitability for other applications may depend on the context.

B. Data

We have order-level data from both ASX and Chi-X. In each case, the data are a complete historical record of messages broadcast by the exchange, which market participants can access in real time. Appendix B.C in the Supplementary Material details the data and the required data-processing steps.

[16] As of June 2014, STW had $2.3 billion under management, with 45 million units on offer. The fund consisted of 205 constituents, with a weighted average market capitalization of $54 million (State Street (2014)).

[17] Although private information is a primary driver of volatility in individual stocks (French and Roll (1986), Barclay, Litzenberger, and Warner (1990)), its empirical significance seems to be smaller for ETFs (Tse and Martinez (2007)). Moreover, Subrahmanyam (1991) provides a model explaining this.

The ASX data cover the trading days of February–June 2014. The Chi-X data span only February through May. From the sample, we drop 2 days that were affected by data issues.[18] Our main analysis requires data from both exchanges and therefore uses the 80 remaining days of February–May. We further focus on the hours of 10:30–16:00 during each day.[19] A supplemental out-of-sample analysis in Section VII uses ASX data on the 20 trading days in the month of June. Figure 1 plots the close price and traded volume of STW for February–June of 2014. Aside from a rally early in February, the price remains fairly stable in the neighborhood of $51. Volumes are somewhat more volatile.

FIGURE 1

Price and Traded Volume for STW (2014)

Figure 1 presents the trading statistics for STW for Feb. 3, 2014–June 30, 2014. "Close price" is the close price as announced by ASX. "Volume" includes all trading in the Australian market, in thousands of contracts. Data are from Bloomberg.

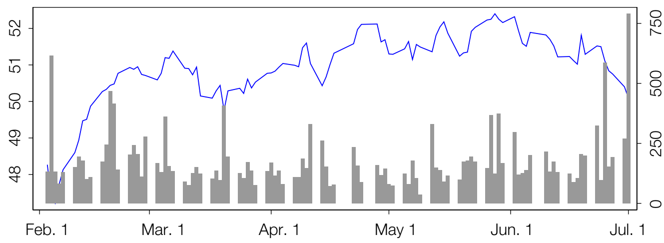


Table 1 presents statistics summarizing trading activity for STW. During our main sample, 169,930 contracts per day were traded on average across the Australian market. Our analysis focuses on a subset of trades that we call *lit book volume*, which consists of on-exchange trades during the continuous session in which the passive order had been visible. (See Appendix B.C in the Supplementary Material for details.) Of the total STW volume, 72.3% is traded in the lit book of ASX and 16.5% in the lit book of Chi-X. The remaining 11.2% includes i) the ASX opening and closing crosses, ii) off-exchange trading (e.g., crossing systems, block trades, and internalization), and iii) on-exchange trading in which the passive order had not been visible.

Table 1 also presents statistics on the messages in the ASX and Chi-X feeds that pertain to STW. The number of total messages is comparable across the two exchanges. But because of differences in the amount of volume traded, the ratio of total messages to trade messages tends to be much higher at Chi-X than at ASX. Finally, Table 1 also displays statistics pertaining to the volatility of STW (based on 1-second returns) and price movement (based on daily returns).

[18]One of those days is Feb. 11, on which there were known issues with the ASX feed (Chi-X (2014a)). The other day is May 2, for which our record of the ASX feed is incomplete.

[19]On ASX, the continuous trading session for STW begins at a random point in the interval [10:08:45, 10:09:15] and ends at 16:00. On Chi-X, the continuous trading session begins at 10:00 and ends at 16:12. We limit attention to continuous trading between 10:30 and 16:00 to ensure a balanced panel and to avoid contamination from the opening and closing auctions at ASX.

TABLE 1
Summary Statistics for Trading of STW

Table 1 presents summary statistics for the trading of STW throughout our sample of 80 trading days. "Lit book" refers to traded volume in which the passive order had been visible; "other" volume is calculated by subtracting that from the total volume as obtained from Bloomberg. "Number of Messages" and "Message-to-Trade Ratio" are calculated from the feeds. "Volatility" is based on 1-second returns (computed using the average of all 4 cum-fee quotes) between 10:30 and 16:00. "Price Movement" is the absolute value of the daily return, computed using Australian Stock Exchange (ASX) open and close prices.

	Mean	St. Dev.	Quartile 1	Median	Quartile 3
Traded Volume (1,000 contracts)					
ASX lit book	122.90	84.92	72.06	96.11	143.22
ASX other	10.52	19.20	3.12	4.81	10.96
Chi-X lit book	28.00	20.85	15.66	22.32	33.19
Chi-X other	8.51	18.33	0.00	0.05	1.77
Total	169.93	100.41	109.26	139.58	178.87
Number of Messages					
ASX	19,350	5,010	16,000	19,053	21,537
Chi-X	17,711	6,175	13,040	16,400	19,962
Message-to-Trade Ratio					
ASX	66.97	27.21	49.12	62.01	77.33
Chi-X	320.77	209.47	181.48	268.42	376.94
Volatility (bps)	0.32	0.06	0.28	0.31	0.35
Price Movement (bps)	32.13	26.84	11.86	26.24	41.90

C. Trade Classification

Our empirical approach requires identifying which of the trades in our sample correspond to the investor-initiated trades of the model. To that end, we recall that in the model, investor trades occur in isolation from other trades. In contrast, sniper trades occur in clusters: taking place on both exchanges simultaneously and in the same direction. Leveraging this distinction, we use isolated trades and clustered trades as empirical proxies for the investor and sniper trades of the model, respectively.[20] Specifically, we classify a lit book trade as *isolated* if no other lit book trade in the same direction occurs within a certain cutoff on either exchange.[21] In the baseline, we set this cutoff to 1 second. Remaining lit book trades are classified as *clustered*.[22] Note that this classification does not rely on trader-level information (which we do not possess), but can be computed from publicly available order-level data.

A potential concern is classification error, which could arise if the distinction between isolated and clustered trades differs from the sharp dichotomy predicted by the model. We address this in two ways. First, Figure 2 illustrates that the predicted dichotomy is, in fact, not far from the truth. For each trade, we compute the length of time to the nearest trade in the same direction on either exchange; the figure plots the empirical distribution of this variable. Many trades are within 50 milliseconds of another trade, beyond which is a long tail. The amount of

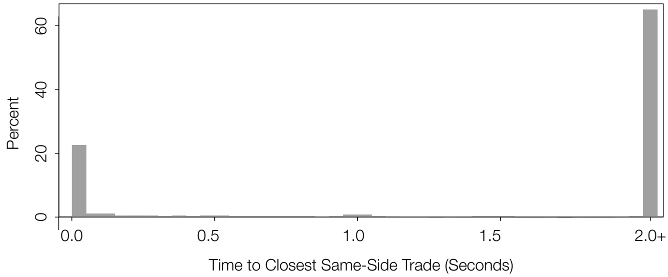
[20] Appendix E.D in the Supplementary Material shows clustered trades to be better predictors of subsequent price movements than isolated trades, which further supports isolation as a proxy for the liquidity-motivated investor trades of the model.

[21] Note that a single marketable limit order might trigger multiple execution messages if it is matched with multiple resting orders. We treat such cases as a single trade.

[22] van Kervel (2015) uses essentially the same classification scheme, proxying for the fraction of fast traders with the percentage of market orders that arrive in clusters. A difference is that he uses a 0.1-second cutoff, whereas we use a 1-second cutoff in our baseline specification. Nevertheless, Appendix D.A in the Supplementary Material shows that a 0.1-second cutoff leads to similar results.

FIGURE 2
Distribution of Time Gaps between Trades

In Figure 2, for each lit book trade of STW that occurred on the Australian Stock Exchange (ASX) or Chi-X Australia (Chi-X) between 10:30 and 16:00 of a day in the main sample, we compute the difference in time stamps to the nearest other such trade in the same direction on either exchange. Figure 2 plots the distribution of these differences in 50-millisecond bins.



classification error must then be relatively limited because only a small fraction of trades lie between any two candidate cutoff points. In particular, for only 12.3% of trades is the nearest trade in the same direction between 50 milliseconds and 2 seconds away. Second, Appendix D.A in the Supplementary Material demonstrates that our results are robust to alternative choices of the cutoff used to classify trades as isolated or clustered.

D. Direct Evidence of Market Frictions

A key feature of the model is that it allows for frictions that might affect the exchange choices of investors. If, in practice, investors often fail to trade at the best price, then it stands to reason that these frictions must be strong. This provides a direct and model-free way to assess the strength of these frictions, which we pursue here before turning to estimation. In this section, we demonstrate that for trades taking place when quoted prices differ at Chi-X and ASX (so that a single "best price" exists), a significant fraction of them occur at the exchange offering the worse price, which indicates that these frictions are indeed strong.

For this analysis, we focus on isolated trades, which are empirical proxies for the investor trades of the model.[23] We observe the price obtained for every isolated trade (gross of take fees).[24] Similarly, we can infer the price that would have been obtained had one instead traded the same volume at the same time on

[23]Note that a trader accessing liquidity at multiple venues (as with an intermarket sweep order) might rationally trade at an exchange featuring a worse price, but he would also trade at the exchange with the better price at essentially the same time. Thus, our classification algorithm would label such trades as clustered, and they would be omitted from the analysis in this section. In summary, given our focus on isolated trades, such multi-venue trading strategies cannot rationalize the frequency with which trades occur at inferior prices.

[24]Note that a marketable order may execute against multiple resting orders and even "walk the book" by executing against orders resting at different price levels. Following footnote 21, we would consider this as a single trade and compute the price of that trade by aggregating across the individual executions.

the other exchange.[25] We then determine which exchange or exchanges featured the best price for a trade of that size at the time (the lowest price for isolated buys, the highest price for isolated sells), and we compare that to the exchange on which the trade actually occurred. The results are tabulated in Table 2.

TABLE 2
Exchange Choice for Isolated Trades, Number of Trades

Table 2 classifies isolated trades by prevailing quotes at the two exchanges. Columns count the following cases: 1, the trade occurs at the exchange with strictly better price; 2, the trade occurs at the exchange with strictly inferior price; 4, the trade occurs when prices are the same.

	Different Price			Identical Price	Total
	Best Price	Inferior Price	Subtotal		
	1	2	3	4	5
Buy	4,533	441	4,974	2,572	7,546
Sell	4,299	425	4,724	2,497	7,221
Buy or Sell	8,832	866	9,698	5,069	14,767

As Table 2 indicates, many trades occur on exchanges offering inferior prices. Aggregating across buys and sells, 9,698 of the 14,767 isolated trades in our main sample occur when the two exchanges offer different prices. And of those, 8.9% occur on the exchange with the worse price. The magnitude of this frequency suggests that frictions are an empirically important determinant of exchange choice, and it motivates our development of a model that allows for them.[26,27] Moreover, because the strength of these frictions is parameterized in the model by α, this evidence also indicates that we should expect a relatively high value for α when we estimate the model in Section VI. Finally, although this analysis points to considerable frictions, it is silent as to their micro-foundations. As discussed in Section III, several potential sources exist.

E. Variables Used in Estimation

For the estimation, we discretize time into 1-second intervals. For each second during our main sample (10:30–16:00 for 80 trading days in 2014), we construct variables pertaining to the cum-fee prices that prevailed and to the trades that took place.

Prices

The cum-fee bid and ask prevailing at exchange x at the beginning of second t are denoted $b_{x,t}$ and $a_{x,t}$. These are calculated from the quoted bid and ask by

[25] In some cases, it would have been infeasible to trade the same volume at the counterfactual exchange, due to insufficient depth. We then classify the counterfactual exchange as offering an inferior price.

[26] In fact, the foregoing analysis might actually *understate* these frictions by failing to fully consider the potential for order splitting. A more stringent benchmark for evaluating execution quality would be to compare against the price that could have been obtained by splitting orders across exchanges in an optimal way.

[27] One potential concern is that these results might be alternatively explained by misaligned time stamps. However, industry contacts have indicated that this is unlikely to be the case. Additionally, we have conducted a number of unreported robustness checks to alleviate this concern.

adjusting for take fees: 0.15 and 0.12 bps of the value of the trade for ASX and Chi-X, respectively.[28] To denote cum-fee spreads, we use $s_{x,t} = a_{x,t} - b_{x,t}$.

Trades

The indicators $\text{BUY}_{x,t}$ and $\text{SELL}_{x,t}$ evaluate to unity if an isolated trade that is an aggressive buy or, respectively, sell occurs on exchange x in second t.[29]

Summary Statistics

Table 3 presents summary statistics for seven key variables: the cum-fee spreads at ASX and Chi-X, indicators for isolated buys and sells at ASX and Chi-X, and an indicator for clustered trades. For each variable we report the mean and standard deviation over the sample. Spreads at ASX and Chi-X are very similar, albeit slightly smaller and more volatile at ASX. Reflecting ASX's incumbent position, isolated buys and sells are relatively more frequent there. Finally, these isolated trades are somewhat more frequent, in aggregate, than trades we classify as clustered.

TABLE 3
Summary Statistics of Variables Used in Estimation

In Table 3, each observation is a second between 10:30 and 16:00 in one of the 80 trading days in the sample ($N=1,584,000$). For each exchange x, s_x is the cum-fee spread, measured in cents and evaluated at the start of each second. The indicators BUY_x and SELL_x evaluate to unity for seconds in which an isolated buy or, respectively, sell occurs; a lit book trade is isolated if no other lit book trade in the same direction occurs on either exchange within 1 second before or after. Remaining lit book trades are clustered; the indicator CLUSTERED evaluates to unity for seconds in which such a trade occurs.

	s_{ASX}	$s_{\text{Chi-X}}$	BUY_{ASX}	$\text{BUY}_{\text{Chi-X}}$	SELL_{ASX}	$\text{SELL}_{\text{Chi-X}}$	CLUSTERED
Mean	2.80865	2.94904	0.0038	0.00097	0.00356	0.001	0.002
Std. dev.	1.02835	0.7482	0.06151	0.03117	0.05955	0.0316	0.04466

VI. Estimation and Counterfactual Analysis

First, we describe how variation in the data is used to identify and estimate key parameters. We then discuss the estimates and use them to predict the counterfactual monopoly spread.

A. Empirical Strategy

Four parameters require estimation. Three govern the demand system of investors: λ_i, their arrival rate; θ, the strength of their private transaction motive; and α, the strength of market frictions that distort their choice of exchange. The fourth parameter is λ_j, the arrival rate of information, which affects the amount of adverse selection.

[28]Letting $\hat{b}_{x,t}$ and $\hat{a}_{x,t}$ denote the quoted bid and ask: $b_{\text{ASX},t} = 0.999985\hat{b}_{\text{ASX},t}$; $a_{\text{ASX},t} = 1.000015\hat{a}_{\text{ASX},t}$; $b_{\text{Chi-X},t} = 0.999988\hat{b}_{\text{Chi-X},t}$; $a_{\text{Chi-X},t} = 1.000012\hat{a}_{\text{Chi-X},t}$.

[29]Although not used in our main text analysis, we also define the indicator CLUSTERED_t, which evaluates to unity if a clustered trade occurs in second t.

Identification

The identification argument has two parts. First, α, θ, and λ_i are identified by how the arrival rates of isolated buys and sells at ASX and Chi-X fluctuate with variation in prices. Intuitively, α is identified by the cross-price elasticity, θ by the own-price elasticity, and λ_i by the frequency of isolated trades. Second, given values for these three parameters, the duopoly spread is monotone in λ_j (cf. Proposition 4). Thus, the spread identifies λ_j.

Estimating Equations

Equation (7) in Section IV.A states the arrival rate of investor buys as a function of the prevailing asks. In the duopoly case, its empirical analogue for each exchange x and second t is

$$(8)\qquad \text{BUY}_{x,t} = \frac{\lambda_i}{2}\left[\frac{1}{2}+\frac{a_{-x,t}-a_{x,t}}{\alpha}\right]_0^1\left[1-\frac{a_{x,t}-v_t}{\theta}\right]_0^1+\varepsilon_{x,t}^{\text{buy}},$$

where $\varepsilon_{x,t}^{\text{buy}}$ is an error term whose mean, conditional on the quotes, is assumed to be 0. In the model, investor arrivals follow a Poisson process. The error term captures deviations of this random process from its mean, as well as any unmodeled determinants of trade. Likewise, we obtain the following for investor sells:

$$(9)\qquad \text{SELL}_{x,t} = \frac{\lambda_i}{2}\left[\frac{1}{2}+\frac{b_{x,t}-b_{-x,t}}{\alpha}\right]_0^1\left[1-\frac{v_t-b_{x,t}}{\theta}\right]_0^1+\varepsilon_{x,t}^{\text{sell}},$$

where $\varepsilon_{x,t}^{\text{sell}}$ is an error term with a conditional mean of 0. Because we do not observe v_t, we proxy it with the average midprice $(b_{\text{ASX},t}+b_{\text{Chi-X},t}+a_{\text{ASX},t}+a_{\text{Chi-X},t})/4$ in both equations (8) and (9).

Finally, Proposition 2 provides an expression for the duopoly spread. Its empirical analogue for each exchange x and second t is

$$(10)\qquad s_{x,t} = \theta+\alpha-\sqrt{\theta^2+\alpha^2-\frac{4\alpha\theta\lambda_j}{\lambda_i}}+\varepsilon_{x,t}^{\text{spread}},$$

where $\varepsilon_{x,t}^{\text{spread}}$ is an error term with an expectation of 0. Note that the model does not suggest a reason for why the spread should vary around its mean. Nevertheless, spreads do display a limited amount of variation in the data. There are a number of potential explanations for short-term deviations from this long-run pricing equation. First, we have assumed that the parameters of the model are constant over time. Although this seems accurate as a first approximation, small deviations may arise in practice, which would induce variation in the spread. Second, we have assumed that all agents are risk neutral. In practice, liquidity providers are likely to be either risk averse or constrained in their ability to take on inventory, in which case they may vary their quotes based on their net positions. We assume further that the data-generation process is stationary and weakly dependent.

Estimation Procedure

We estimate the parameters using systems nonlinear least squares on equations (8)–(10) by minimizing the objective

$$Q_T(\alpha,\theta,\lambda_i,\lambda_j) \quad = \quad \frac{1}{T}\sum_{t=1}^{T}\sum_{x\in\{\text{ASX,Chi-X}\}} \left(\varepsilon_{x,t}^{\text{buy}}\right)^2 + \left(\varepsilon_{x,t}^{\text{sell}}\right)^2 + \left(\varepsilon_{x,t}^{\text{spread}}\right)^2. \tag{11}$$

The consistency of this procedure is formally proven in Appendix B.D of the Supplementary Material. Standard errors are computed using a nonoverlapping block bootstrap procedure. See Appendix B.E in the Supplementary Material for further details on the implementation of estimation and the computation of standard errors.

B. Parameter Estimates

Table 4 reports the estimation results. The point estimate of α of 7.34¢ implies that an exchange attracts every investor in the market only if it offers prices at least 3.67¢ better than its competitor. This considerably exceeds the average half spread of 1.44¢, which indicates that market frictions significantly distort the exchange choices of investors. Moreover, the null hypothesis of frictionless routing (i.e., $\alpha = 0$) is rejected at all common significance levels. The estimate of θ of 1.53¢ means that the average magnitude of private transaction motives among investors (i.e., $\theta/2$) is 53% of the average half spread, which indicates that transaction costs crowd out a substantial number of potential investor trades. Finally, the estimates of λ_i and λ_j indicate that investors arrive 2.2 times more frequently than information.[30]

TABLE 4
Parameter Estimates

In Table 4, point estimates are computed to minimize the objective given by equation (11). Standard errors (in parentheses) are based on 200 block bootstrap replications. *$p<0.1$, **$p<0.05$, ***$p<0.01$.

	α	θ	λ_i	λ_j
Point estimate	7.33504***	1.52817***	0.00172***	0.00078***
	(1.06129)	(0.15127)	(0.00022)	(0.00012)

To illustrate the elasticities implied by the estimated model, note that if quotes are symmetric about the value of the security, then investors arrive at an exchange x at the rate

$$\lambda_i\left[\frac{s_{-x}-s_x}{\alpha}+\frac{1}{2}\right]_0^1\left[1-\left(\frac{1}{\theta}\right)\left(\frac{s_x}{2}\right)\right]_0^1.$$

Evaluated at the estimates, the elasticity with respect to s_x is -17.1, whereas the elasticity with respect to s_{-x} is 0.8. The relatively high own-price elasticity is due

[30]In the model, adverse selection has one source: public information about fundamentals. In reality, adverse selection has other sources as well, including public information about order flow (which would be relevant to a risk-averse liquidity provider) and private information. The presence of these additional sources might drive up the spread and, given our estimation procedure, also be contributing to the estimate of λ_j.

to the spread falling on the highly elastic portion of the linear demand schedule, and the relatively low cross-price elasticity reflects the presence of considerable market frictions.

C. The Effect of Fragmentation

In our model, it is theoretically ambiguous whether fragmentation benefits investors. This is because the competition channel (i.e., additional exchanges intensify price competition) and the exposure channel (i.e., additional exchanges intensify adverse selection) act in opposite directions. Nevertheless, the estimates from the previous section can be used to resolve this theoretical ambiguity for the case of Australia and STW.

We first consider the effect of a reduction in the number of exchanges. Under the prevailing duopoly regime, the average spread observed in the data is 2.88¢. Using Proposition 1, our estimates imply that in the counterfactual of a monopoly, the spread would be 2.22¢, or 22.9% lower. Moreover, the monopoly spread is lower than the duopoly spread in each of the 200 bootstrap replications. Two features of the Australian market might contribute to why the exposure channel appears to outweigh the competition channel: i) the lack of an order-protection rule because the presence of one might counteract market frictions and strengthen the competition channel, and ii) the relatively small natural base of investors, which suggests high adverse selection and a strong exposure channel.

In principle, the estimates could also be used to consider the effects of counterfactual increases in the number of exchanges. According to the estimated model, three or more exchanges could not operate simultaneously without market breakdown because A2 and A3 cannot simultaneously hold with $X \geq 3$. However, one reason for interpreting this prediction with caution is that exchanges, in practice, have revenue sources beyond trading fees. Although such additional revenue streams may be strategically orthogonal to the trading-fee decision, and hence to our previous conclusions, they would imply that A3 might be more than is needed to ensure the simultaneous profitability of all exchanges.

VII. Natural Experiment: Chi-X Shutdown

Our estimates imply that the spread of STW would be substantially lower if instead of two exchanges, the Australian market were to have only one. To provide additional support for this conclusion, we study a natural experiment in which Chi-X shut down for a day, leaving ASX as the only operating exchange. Because the shutdown was short and unanticipated, ASX did not strategically alter its fees in response. Thus, this episode constitutes an isolated test of the exposure channel of the model and offers an opportunity to quantify its magnitude. Consistent with the predictions of the model, the STW spread is substantially smaller on that day than on the surrounding days, as well as relative to an unaffected control group. Combining these results with a back-of-the-envelope quantification of the competition channel, we find a net effect on par with the estimates of the previous section.

The Chi-X shutdown has several advantages as a natural experiment with which to test the predictions of the model. First, the shutdown was unanticipated

https://doi.org/10.1017/S0022109019000814 Published online by Cambridge University Press

and exogenous. Second, its occurrence within weeks of the end of our estimation sample means that the underlying parameters are more likely to resemble those prevailing during the estimation sample than if we had used an episode several years prior or hence (e.g., Chi-X's 2011 entry).

A. Event Description

On June 16, 2014, a technical issue caused Chi-X to halt its trading at 11:08 (Chi-X (2014b)), leaving ASX as the only exchange in operation until Chi-X resumed trading the next morning. The issue arose due to a small change Chi-X had made in its handling of certain orders. That change was implemented with an error, which, when noticed, prompted the shutdown. Hence, the shutdown was exogenous to trading activity. Indeed, June 16 (the "monopoly day") appears similar to other trading days in June 2014 (the "duopoly days") in terms of volume, message flow, volatility, and price movement. As Table 5 illustrates, the monopoly day is within 1 standard deviation of the duopoly mean for each statistic.[31]

TABLE 5

Summary Statistics for Trading of STW (June 2014)

Table 5 presents summary statistics for the trading of STW for all 20 trading days in June 2014. "Lit book" refers to traded volume in which the passive order had been visible. The total volume for Australia is obtained from Bloomberg. "Number of Messages" and "Message-to-Trade Ratio" are calculated from the Australian Stock Exchange (ASX) feed. "Volatility" is based on 1-second returns (computed using the average of the ASX cum-fee quotes) between 10:30 and 16:00. "Price Movement" is the absolute value of the daily return, computed using ASX open and close prices.

	Monopoly Day	Duopoly Days (June 2014, $N=19$ days)				
		Mean	St. Dev.	Quartile 1	Median	Quartile 3
Volume (1,000 contracts)						
ASX lit book	121.65	160.18	97.10	90.33	145.32	198.16
Total	123.84	197.81	115.36	126.62	173.73	215.93
Number of Messages						
ASX	20,941	17,222	4,531	13,179	16,192	20,537
Message-to-Trade Ratio						
ASX	85.82	76.19	36.89	48.49	68.10	88.83
Volatility (bps)	0.298	0.354	0.162	0.278	0.332	0.35
Price Movement (bps)	17.60	41.99	31.26	13.69	37.30	69.28

B. Analysis

To corroborate the conclusions of our structural estimation, we investigate how the spread of STW is affected by this shock to fragmentation. To that end, we compute the ASX cum-fee spread prevailing at the beginning of every second between 11:08 and 16:00 of every trading day in June 2014. We use 11:08 because that was the time at which Chi-X shut down.

We then regress the STW spread on an indicator for June 16, 2014. In our baseline specification, the sample consists of all seconds between 11:08 and 16:00

[31]Although not an outlier, the monopoly day's volume is relatively low. But this does not drive our findings: Appendix E.B in the Supplementary Material shows that our results survive even if we control for volume or estimate only on low-volume days.

in all 20 trading days in June 2014. The results, reported in column 1 of Table 6, indicate that the monopoly day is associated with a statistically significant 28.0% reduction in the average spread. (See Appendix E.A in the Supplementary Material for more details on the shift in the spread distribution.)

TABLE 6

ASX Spreads, Australian Equity Exchange-Traded Funds (June 2014)

In Table 6, the dependent variable is the cum-fee spread in cents prevailing on the Australian Stock Exchange (ASX) at the beginning of the second. An observation is a pair: a second between 11:08 and 16:00 in June 2014 and a security traded on ASX. STW × MONOPOLY is an indicator for June 16, 2014 and STW. Samples for the respective columns are as follows: column 1, all trading days in June 2014 and STW; column 2, all Mondays in June 2014 and STW; column 3, all trading days from June 6 until June 23 and STW; column 4, all trading days from June 11 until June 19 and STW; and column 5, all trading days in June 2014 and the securities STW, IOZ, ISO, MVW, QOZ, SSO, VAS, VLC, and VSO. Coefficients are estimated by ordinary least squares. Standard errors (in parentheses) are clustered by 120-second blocks on each trading day. STW duopoly mean is the average of the dependent variable for STW, nonmonopoly observations. "Change (%)" is the estimate relative to the STW duopoly mean. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

	All	Mondays	± 5 Days	± 3 Days	All
	1	2	3	4	5
STW × MONOPOLY	−1.127***	−0.981***	−1.069***	−0.882***	−1.036***
	(0.0567)	(0.0901)	(0.0673)	(0.0818)	(0.0643)
STW duopoly mean	4.023	3.877	3.965	3.777	4.023
Change (%)	−28.02	−25.31	−26.97	−23.34	−25.76
Day × hour fixed effects	No	No	No	No	Yes
Security fixed effects	Yes	Yes	Yes	Yes	Yes
Control group	No	No	No	No	Yes
No. of obs.	350,400	70,080	192,720	122,640	3,153,600

A potential concern is that this reduction in the spread is attributable to forces other than the shock to fragmentation. For instance, a similar effect might obtain from diminished information arrival or heightened retail trader activity. To the extent that such forces are correlated with the day of the week or with the time of the year, we can control for them by repeating the previous analysis on different subsamples. In column 2 of Table 6, we focus only on Mondays. In columns 3 and 4, we shorten the event window around the monopoly day to 3 and 5 trading days, respectively. In each case, we obtain qualitatively similar results, which suggests that the finding is indeed driven by fragmentation.

As a second means of ruling out alternative explanations, we introduce a control group of eight securities. Like STW, these are ETFs with exposure to Australian equities that were traded on ASX in June 2014, but unlike STW, they were not also traded on Chi-X. (See Appendix B.F in the Supplementary Material for more details on the selection of this group.) Using this control group, we pursue a difference-in-differences approach to isolate the effects of the change in the number of venues on which STW was traded from those of any market-wide shocks that may have influenced trading conditions on the day of the shutdown. The results are reported in column 5 of Table 6 and are again qualitatively similar.

Our model not only predicts a reduction in the spread on the monopoly day but also specifies the channel for that effect: short-lived adverse selection. In the model, eliminating an exchange concentrates uninformed order flow on the remaining venue and reduces adverse selection there. To test this prediction, we compute the adverse selection of a trade as the signed return over the subsequent

10 seconds (or 5 minutes).[32] As expected, adverse selection is on average lower on the monopoly day. Averaging across all trades, the 10-second (5-minute) adverse selection is 0.785 bps (1.398 bps), compared with an average of 1.757 bps (2.240 bps) on duopoly days. This evidence supports the mechanism underlying the exposure channel, and in so doing, it provides additional backing for our modeling approach.

C. Discussion

In the model, the full equilibrium effect of subtracting an exchange is determined by the magnitudes of both the competition and the exposure channels. That spreads are on average 28.0% smaller on the monopoly day is consistent with the large exposure channel predicted by the model. However, our analysis does not speak to the competition channel: Because the shutdown was unexpected and short-lived, ASX did not alter its fees in response. Therefore, this figure can be interpreted only as an upper bound on the net effect.

A rough estimate of the magnitude of the competition channel might be derived from the July 2010 fee change made by ASX in response to the announcement that Chi-X would enter the following year: a reduction from 0.56 to 0.30 bps. Multiplying this change of 0.26 bps by the closing price of STW on the monopoly day ($51.23) yields 0.133¢. Under the assumption that fee changes are passed through on a one-for-one basis into the spread, we can interpret this figure as the magnitude of the competition channel. Adding it to the −1.127¢ estimate from column 1 of Table 6, which we interpret as the magnitude of the exposure channel, suggests a net effect of −0.994¢. This would equate to a 24.7% decrease from the average spread on duopoly days, very much in line with the 22.9% reduction predicted by the model.

Although the identification strategy underlying our analysis is strong, a potential concern is that trader behavior on the day of the shutdown might not reflect long-run behavior in a monopoly environment. For instance, certain traders, knowing the shutdown to be temporary, might have withdrawn from trading on that day, whereas they would adapt over the long run to a permanent change in the number of venues. Nevertheless, several factors mitigate this worry. First, that the spread narrows on the monopoly day suggests that liquidity providers, at least, did not withdraw. Second, private conversations with industry participants in Australia and the United States indicate that traders are well equipped to deal with such contingencies and have no need to withdraw.[33] Third, the total volume for STW on the day after the shutdown was only 88,953 contracts, the lowest value for the entire month of June. This is at odds with the theory that traders withdrew on the monopoly day: In that case, one would expect larger volumes on the following day as the displaced trades materialize.

[32]The adverse selection of a trade is computed as $q(m_{t+\Delta} - m_t)/m_t$, where $q=1$ for a buy and $q=-1$ for a sell, m_t is the midprice just before the trade, and $m_{t+\Delta}$ is the midprice Δ time units thereafter.

[33]Consistent with this, a similar episode that occurred in the United States 1 year later had almost no impact on trading (the *Wall Street Journal* (2015)), despite a primary venue, NYSE, shutting down in that case.

In spite of the aforementioned caveats, this natural experiment supports the model's prediction that STW spreads would be lower in a monopoly. Moreover, these results might be interpreted as a proof of concept for the model, suggesting that our approach could be a valid way of ascertaining the effects of fragmentation in other contexts.

VIII. Potential Concerns

The model omits several important features of financial markets. Some omissions prove to be without consequence: Appendix F in the Supplementary Material shows that our conclusions persist in a variety of extensions.[34] Nevertheless, other omissions could have import.

One potential concern is our focus on symmetric equilibria, which could limit the extent to which our model is a good fit for our chosen empirical application as well as for other potential applications. Although ASX and Chi-X are reasonably symmetric with respect to some key variables (e.g., spreads, message traffic), symmetry fails to hold in other respects. For example, ASX handles more volume and also charges higher fees.[35]

A second limitation is that our model and estimation focus on only a single security. We emphasize that the effects of fragmentation depend on many factors (e.g., those corresponding to the parameters of our model). Thus, although our analysis predicts that fragmentation leads to larger spreads in the case of STW and Australia, we would not wish to imply that it always has this effect. For instance, other studies of the Australian market (He et al. (2015), Aitken et al. (2017)) have linked fragmentation to smaller spreads. Those results may differ from ours because they apply to an earlier point in time (around the 2011 entry of Chi-X) and a different set of securities (S&P/ASX 200 constituents). Related is that our single-security model is incapable of capturing cross-security substitution patterns. An interesting direction for future work would be to expand this framework to a multi-security setting.

Third, the model makes quite stark parametric assumptions. In particular, we assume a linear demand framework, and we assume that both investors and information arrive at constant and exogenous rates. Although restrictive, these assumptions can be thought of as first-order approximations, which facilitate tractability and closed-form derivations.

Fourth, although the model allows for some forms of heterogeneity among investors (i.e., in the strength of their private transaction motives and in their susceptibility to market frictions), it does not allow for heterogeneity in other

[34]We consider extensions involving i) inventory constraints, ii) operation costs for liquidity providers and exchanges, iii) richer evolution processes for the security value, iv) across-exchange order splitting, v) short-lived private information, vi) successful order cancellation, and vii) stochastic horizon times.

[35]Thus, alternative (although perhaps less parsimonious) models that permit asymmetries might yield better fits. One possibility would be to relax the assumption that investors are uniformly distributed on the circle. Another would be to allow for a mass of investors without access to the smaller exchange, as in Foucault and Menkveld (2008). Although this latter modeling choice may have been appropriate for their 2004–2005 data, it seems less suitable for modern trading: Essentially all traders now have easy access to all exchanges.

dimensions. For example, the model does not permit differences in either the volume that investors seek to trade or in their patience for spreading trades over time. With such heterogeneity, the spread would cease to be a sufficient welfare statistic as it is in our model because the spread quantifies transaction costs only for small trades. Moreover, such heterogeneity might also extend to concerns with the empirical classification of clustered trades as information motivated because a large but uninformed investor might initiate simultaneous trades on multiple exchanges.

Finally, the model necessarily falls short of incorporating every potential channel through which fragmentation might affect market quality. Some such omissions seem without much loss. For instance, although network externalities were the focus of many early analyses of fragmentation (e.g., Pagano (1989)), they are less significant in the electronically linked trading environments of today. But other channels absent from our model do remain important. For example, the competitive benefits of fragmentation might not be limited to the price dimension: Fragmentation could induce some exchanges to improve their speed (as in Pagnotta and Philippon (2018)) or technological services (as in Cespa and Vives (2019)). Further, fragmentation, together with maker-taker pricing, allows liquidity providers to compete on a finer price grid, which could reduce the spread of the aggregate book (as in Chao et al. (2019)). Conversely, fragmentation increases the costs of communication (as in Mendelson (1987)), and fixed costs are also required to establish new venues. Although our focus on two channels has the advantage of yielding a parsimonious and tractable model, it would also be valuable to develop a richer model that incorporates some of these other elements.

IX. Conclusion

This article provides a tractable and estimable model that captures several first-order aspects of the interaction between fragmentation and liquidity. The model features two countervailing forces: i) the competition channel, whereby adding more exchanges induces lower fees and therefore smaller spreads, and ii) the exposure channel, whereby adding more exchanges induces more adverse selection against liquidity providers and therefore larger spreads.

The parameters of the model are identified from the data by the average spread, together with how the incidence of certain trades depends on prevailing prices. We also demonstrate a procedure for estimating these parameters. For an empirical application, we use data pertaining to the trading of an Australian ETF. Our estimates imply the existence of significant market frictions, suggesting the competition channel to be limited in magnitude and outweighed by the exposure channel. Indeed, according to the estimates, traders of the ETF would fare better under a monopoly than under Australia's prevailing duopoly. To corroborate the approach, we show this prediction to align with that of a separate analysis conducted in the same setting but based on exogenous variation in the number of venues. We submit that our approach can be useful when rule changes affecting fragmentation have to be decided, especially in the absence of sufficient identifying variation in market structure.

Supplementary Material

Supplementary Material for this article is available at https://doi.org/10.1017/S0022109019000814.

References

Aït-Sahalia, Y., and M. Sağlam. "High Frequency Market Making: Implications for Liquidity." Working Paper, available at http://ssrn.com/abstract=2908438 (2017a).

Aït-Sahalia, Y., and M. Sağlam. "High Frequency Market Making: Optimal Quoting." Working Paper, available at http://ssrn.com/abstract=2331613 (2017b).

Aitken, M.; H. Chen; and S. Foley. "The Impact of Fragmentation, Exchange Fees and Liquidity Provision on Market Quality." *Journal of Empirical Finance*, 41 (2017), 140–160.

Amihud, Y.; B. Lauterbach; and H. Mendelson. "The Value of Trading Consolidation: Evidence from the Exercise of Warrants." *Journal of Financial and Quantitative Analysis*, 38 (2003), 829–846.

Anand, A.; M. Samadi; J. Sokobin; and K. Venkataraman. "Institutional Order Handling and Broker-Affiliated Trading Venues." Working Paper, available at https://www.finra.org/sites/default/files/OCE_WP_jan2019.pdf (2019).

Angel, J. J.; L. E. Harris; and C. S. Spatt. "Equity Trading in the 21st Century: An Update." *Quarterly Journal of Finance*, 5 (2015), 1–39.

Arnold, T.; P. Hersch; J. H. Mulherin; and J. Netter. "Merging Markets." *Journal of Finance*, 54 (1999), 1083–1107.

ASX Group. "Past Editions of Australian Cash Market Reports." Available at http://www.asx.com.au/services/trading-services/past-editions-australian-cash-market-reports.htm (2014).

ASX Group. "ASX Trade: Markets Participant and Trading Schedule of Fees." Available at https://www.asxonline.com/content/dam/asxonline/public/documents/schedule-of-fees/asx-trade-markets-participant-and-trading-schedule-of-fees-v01122016.pdf (2016).

Australian Securities and Investments Commission. "Report 331: Dark Liquidity and High-Frequency Trading." Available at http://download.asic.gov.au/media/1344182/rep331-published-18-March-2013.pdf (2013).

Barclay, M. J.; R. H. Litzenberger; and J. B. Warner. "Private Information, Trading Volume, and Stock-Return Variances." *Review of Financial Studies*, 3 (1990), 233–253.

Battalio, R. H. "Third Market Broker-Dealers: Cost Competitors or Cream Skimmers?" *Journal of Finance*, 52 (1997), 341–352.

Battalio, R. H.; S. A. Corwin; and R. Jennings. "Can Brokers Have It All? On the Relation between Make-Take Fees and Limit Order Execution Quality." *Journal of Finance*, 71 (2016), 2193–2238.

Battalio, R. H.; B. C. Hatch; and M. Sağlam. "The Cost of Routing Orders to High Frequency Traders." Working Paper, available at http://ssrn.com/abstract=3281324 (2018).

Bennett, P., and L. Wei. "Market Structure, Fragmentation, and Market Quality." *Journal of Financial Markets*, 9 (2006), 49–78.

Bernales, A.; I. Riarte; S. Sagade; M. Valenzuela; and C. Westheide. "A Tale of One Exchange and Two Order Books: Effects of Fragmentation in the Absence of Competition." Working Paper, available at http://ssrn.com/abstract=3276548 (2018).

Bernhardt, D., and E. Hughson. "Splitting Orders." *Review of Financial Studies*, 10 (1997), 69–101.

Bessembinder, H., and H. M. Kaufman. "A Cross-Exchange Comparison of Execution Costs and Information Flow for NYSE-Listed Stocks." *Journal of Financial Economics*, 46 (1997), 293–319.

Biais, B.; D. Martimort; and J.-C. Rochet. "Competing Mechanisms in a Common Value Environment." *Econometrica*, 68 (2000), 799–837.

Boehmer, B., and E. Boehmer. "Trading Your Neighbor's ETFs: Competition or Fragmentation?" *Journal of Banking and Finance*, 27 (2003), 1667–1703.

Boneva, L.; O. Linton; and M. Vogt. "The Effect of Fragmentation in Trading on Market Quality in the UK Equity Market." *Journal of Applied Econometrics*, 31 (2016), 192–213.

Branch, B., and W. Freed. "Bid–Asked Spreads on the AMEX and the Big Board." *Journal of Finance*, 32 (1977), 159–163.

Budish, E.; P. Cramton; and J. Shim. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *Quarterly Journal of Economics*, 130 (2015), 1547–1621.

Budish, E.; R. Lee; and J. Shim. "A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?" Working Paper, available at http://ssrn.com/abstract=3391461 (2019).

CBOE Global Markets, Inc. "Historical Market Volume Data." Available at http://markets.cboe.com/us/equities/market_statistics/historical_market_volume/ (2018).

Cespa, G., and X. Vives. "Exchange Competition, Entry, and Welfare." Working Paper, available at http://ssrn.com/abstract=3316784 (2019).

Chao, Y.; C. Yao; and M. Ye. "Why Discrete Price Fragments U.S. Stock Exchanges and Disperses Their Fee Structures." *Review of Financial Studies*, 32 (2019), 1068–1101.

Chi-X Australia. "Market Operations Notice (Reference Number: 0019/11)." Available at https://www.chi-x.com.au/wp-content/uploads/history/Market%20Operations%20Notice%200019-11.pdf (2011).

Chi-X Australia. "Market Operations Notice (Reference Number: 0002/14)." Available at http://www.chi-x.com.au/wp-content/uploads/history/Market%20Operations%20Notice%200002-14.pdf (2014a).

Chi-X Australia. "Technical Notice Regarding Trading Halt on 16th June 2014 (Reference Number: 0012/14)." Available at https://www.chi-x.com.au/wp-content/uploads/history/Chi-X%20Technical%20Notice%200012-14.pdf (2014b).

Chlistalla, M., and M. Lutat. "Competition in Securities Markets: The Impact on Liquidity." *Financial Markets and Portfolio Management*, 25 (2011), 149–172.

Chowdhry, B., and V. Nanda. "Multimarket Trading and Market Liquidity." *Review of Financial Studies*, 4 (1991), 483–511.

Cohen, K. J., and R. M. Conroy. "An Empirical Study of the Effect of Rule 19c-3." *Journal of Law and Economics*, 33 (1990), 277–305.

Colliard, J.-E., and T. Foucault. "Trading Fees and Efficiency in Limit Order Markets." *Review of Financial Studies*, 25 (2012), 3389–3421.

Copeland, T. E., and D. Galai. "Information Effects on the Bid–Ask Spread." *Journal of Finance*, 38 (1983), 1457–1469.

De Fontnouvelle, P.; R. P. H. Fishe; and J. H. Harris. "The Behavior of Bid–Ask Spreads and Volume in Options Markets during the Competition for Listings in 1999." *Journal of Finance*, 58 (2003), 2437–2464.

Degryse, H.; F. de Jong; and V. van Kervel. "The Impact of Dark Trading and Visible Fragmentation on Market Quality." *Review of Finance*, 19 (2015), 1587–1622.

Dennert, J. "Price Competition between Market Makers." *Review of Economic Studies*, 60 (1993), 735–751.

Fink, J.; K. E. Fink; and J. P. Weston. "Competition on the NASDAQ and the Growth of Electronic Communication Networks." *Journal of Banking and Finance*, 30 (2006), 2537–2559.

Foucault, T. "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market." *Journal of Financial Markets*, 2 (1999), 99–134.

Foucault, T., and A. J. Menkveld. "Competition for Order Flow and Smart Order Routing Systems." *Journal of Finance*, 63 (2008), 119–158.

French, K. R., and R. Roll. "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics*, 17 (1986), 5–26.

Gajewski, J.-F., and C. Gresse. "Centralised Order Books versus Hybrid Order Books: A Paired Comparison of Trading Costs on NSC (Euronext Paris) and SETS (London Stock Exchange)." *Journal of Banking and Finance*, 31 (2007), 2906–2924.

Glosten, L. R. "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49 (1994), 1127–1161.

Glosten, L. R., and P. R. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, 14 (1985), 71–100.

Hamilton, J. L. "Marketplace Fragmentation, Competition, and the Efficiency of the Stock Exchange." *Journal of Finance*, 34 (1979), 171–187.

Haslag, P. H., and M. Ringgenberg. "The Demise of the NYSE and NASDAQ: Market Quality in the Age of Market Fragmentation." Working Paper, available at http://ssrn.com/abstract=2591715 (2017).

He, P. W.; E. Jarnecic; and Y. Liu. "The Determinants of Alternative Trading Venue Market Share: Global Evidence from the Introduction of Chi-X." *Journal of Financial Markets*, 22 (2015), 27–49.

Hendershott, T., and C. M. Jones. "Island Goes Dark: Transparency, Fragmentation, and Regulation." *Review of Financial Studies*, 18 (2005), 743–793.

van Kervel, V. "Competition for Order Flow with Fast and Slow Traders." *Review of Financial Studies*, 28 (2015), 2094–2127.

Kyle, A. S. "Continuous Auctions and Insider Trading." *Econometrica*, 53 (1985), 1315–1335.

Mayhew, S. "Competition, Market Structure, and Bid–Ask Spreads in Stock Option Markets." *Journal of Finance*, 57 (2002), 931–958.

Mendelson, H. "Consolidation, Fragmentation, and Market Performance." *Journal of Financial and Quantitative Analysis*, 22 (1987), 189–207.

Menkveld, A. J. "High Frequency Trading and the New Market Makers." *Journal of Financial Markets*, 16 (2013), 712–740.

Neal, R. "Potential Competition and Actual Competition in Equity Options." *Journal of Finance*, 42 (1987), 511–531.

Nguyen, V.; B. F. Van Ness; and R. A. Van Ness. "Short- and Long-Term Effects of Multimarket Trading." *Financial Review*, 42 (2007), 349–372.

Nielsson, U. "Stock Exchange Merger and Liquidity: The Case of Euronext." *Journal of Financial Markets*, 12 (2009), 229–267.

O'Hara, M., and M. Ye. "Is Market Fragmentation Harming Market Quality?" *Journal of Financial Economics*, 100 (2011), 459–474.

Pagano, M. "Trading Volume and Asset Liquidity." *Quarterly Journal of Economics*, 104 (1989), 255–274.

Pagnotta, E., and T. Philippon. "Competing on Speed." *Econometrica*, 86 (2018), 1067–1115.

Salop, S. C. "Monopolistic Competition with Outside Goods." *Bell Journal of Economics*, 10 (1979), 141–156.

State Street Global Advisors. "SPDR S&P/ASX 200 Fund." Available at http://spdrs.com.au/etf/fund/fund_detail_STW.html (2014).

Subrahmanyam, A. "A Theory of Trading in Stock Index Futures." *Review of Financial Studies*, 4 (1991), 17–51.

Tse, Y., and V. Martinez. "Price Discovery and Informational Efficiency of International iShares Funds." *Global Finance Journal*, 18 (2007), 1–15.

U.S. Department of the Treasury. *A Financial System That Creates Economic Opportunities: Capital Markets*. Washington, DC: U.S. Department of the Treasury (2017).

U.S. Securities and Exchange Commission. "Regulation NMS: Final Rules and Amendments to Joint Industry Plans." *Federal Register*, 70 (2005).

U.S. Securities and Exchange Commission. Equity Market Structure 2019: Looking Back and Moving Forward." Available at https://www.sec.gov/news/speech/clayton-redfearn-equity-market-structure-2019 (2019).

The Wall Street Journal. "Ho-Hum, Just Another Day for Traders." Available at https://www.wsj.com/articles/why-the-nyse-outage-may-not-matter-to-investors-1436374396 (2015).

Weston, J. P. "Electronic Communication Networks and Liquidity on the NASDAQ." *Journal of Financial Services Research*, 22 (2002), 125–139.