Market Fragmentation*

Daniel Chen and Darrell Duffie

Graduate School of Business, Stanford University

February 19, 2020, revised May 26, 2020

Abstract

We model a simple market setting in which fragmentation of trade of the same asset across multiple exchanges improves allocative efficiency. Fragmentation reduces the inhibiting effect of price-impact avoidance on order submission. Although fragmentation reduces market depth on each exchange, it also isolates cross-exchange price impacts, leading to more aggressive overall order submission and better rebalancing of unwanted positions across traders. Fragmentation also has implications for the extent to which prices reveal traders' private information. While a given exchange price is less informative in more fragmented markets, all exchange prices taken together are more informative.

Keywords: market fragmentation, price impact, allocative efficiency, price discovery

JEL codes: G14, D47, D82

*We are grateful for research assistance from David Yang and for conversations with and comments from Mohammad Akbarpour, Bob Anderson, Markus Baldauf, Anirudha Balusubramanian, Jonathan Berk, Eric Budish, Peter DeMarzo, Joe Hall, Elvis Jarnecic, Chafrles-Albert Lehalle, An Qi Liu, Ananth Madhavan, Semyon Malamud, Albert Menkveld, Michael Ostrovsky, Mathieu Rosenbaum, Marzena Rostek, Satchit Sagade, Marcos Salgado, Yuliy Sannikov, Amit Seru, Yazid Sharaiha, Andy Skrzypacz, Bob Wilson, Milena Wittwer, and Haoxiang Zhu. Duffie is a Research Associate of the NBER and a Independent Director of Dimensional's US Mutual Funds Board.

1 Introduction

In modern financial markets, many financial instruments trade simultaneously on multiple exchanges (Budish, Lee, and Shim, 2019; Gresse et al., 2012; Pagnotta and Philippon, 2018). This fragmentation of trade across venues raises concerns over market depth. One might therefore anticipate that fragmentation worsens allocative efficiency through the strategic avoidance of price impact, which inhibits beneficial gains from trade (Vayanos, 1999; Du and Zhu, 2017). Less aggressive trade could in turn impair price informativeness, relative to a centralized market in which all trade flows are consolidated. Perhaps surprisingly, we offer a simple model of how fragmentation of trade across multiple exchanges, despite reducing market depth, actually *improves* allocative efficiency and price informativeness.

In the equilibrium of our market setting, the option to split orders across different exchanges reduces the inhibiting effect of price-impact avoidance on total order submission. Though market depth on each exchange decreases with fragmentation, the common practice of order splitting allows traders to shield orders submitted to a given exchange from the price impact of orders submitted to other exchanges. This effect is sufficiently strong that fragmentation increases overall order aggressiveness. This in turn can result in a more efficient redistribution of unwanted positions across traders and cause prices, collectively across all exchanges, to better reflect traders' private information. Once fragmentation is sufficiently severe, however, any additional fragmentation can cause trade to become *too* aggressive, from a welfare perspective. However, at least in the simple one-period version of our model, any degree of fragmentation is welfare-superior to a centralized market.

Our model abstracts from some important aspects of functioning financial markets. In particular, we do not consider the impact of fragmentation on exchange competition or transaction fees.[1] We also ignore the adverse impact of sniping by fast traders (Budish, Cramton, and Shim, 2015; Malinova and Park, 2019; Pagnotta and Philippon, 2018). Given these and other limitations of our model, we avoid taking a normative or policy stance on fragmentation. Our primary marginal contribution is to identify a potentially important new economic channel for the welfare implications of market fragmentation.

We now briefly summarize our model and the main results. A single asset is traded by N strategic traders participating on E exchanges. Before each round of trade, strategic trader i has a quantity of the asset that is privately observed by trader i. Each trader submits a package of limit orders (forming a demand function) to each of the exchanges, simultaneously. As in common practice (Wittwer, 2020), orders to a given exchange cannot be made contingent on clearing prices at other exchanges. The objective of each strategic

[1] As shown by Budish, Lee, and Shim (2019), transaction fees are economically small.

trader, given the conjectured order submission strategies of the other traders, is to maximize the total expected discounted cash compensation received for executed orders, net of the present value of asset holding costs that are quadratic in the trader's asset position, as in the one-exchange model of Du and Zhu (2017).

At each exchange, "liquidity traders" submit non-discretionary market orders. The aggregate quantities of market orders submitted by liquidity traders to the various exchanges are exogenous random variables, independently and identically distributed across exchanges and periods. In a one-period setting, we also consider a version of the model with no liquidity traders, and a version in which liquidity traders who are local to each exchange are strategic with respect to order quantities. In any version of the model, because agents' preferences are quasilinear in cash and because total cash payments net to zero by market clearing, an unambiguous measure of allocative efficiency is the expected discounted sum of strategic traders' quadratic holding costs.

Price impact is increased by market fragmentation because of cross-exchange price inference, by which traders choose order submissions in light of the positive equilibrium correlation between exchange prices. For example, conditional on a clearing price on a given exchange that is lower than expected, a buyer expects to be assigned higher quantities on all exchanges. This effect dampens the aggressiveness of order submissions, which reduces market depth and heightens market impact, relative to a single-exchange setting. Despite this reduction in market depth, the ability to split orders across exchanges ensures that, in equilibrium, the total order submission of each strategic trader is aggressive enough to achieve the efficient allocation. This natural implication of fragmentation is novel to this paper, as far as we know.

We solve both static and dynamic versions of the model. In the static model, as the number of exchanges increases, the equilibrium allocation becomes more efficient until a point at which trade becomes "too aggressive." We find that the socially optimal number of exchanges depends only on (a) the number of strategic traders and (b) the ratio of the variance of the endowments of strategic traders to the variance of liquidity trade. We show that when there are more exchanges, the price on any individual exchange is less informative of the aggregate asset inventory of strategic traders, the key "state variable" of our model, yet the exchange prices taken together are more informative.

In the dynamic version of the model, we show that market fragmentation still allows efficient trade, despite the associated cross-period cross-exchange price impact and despite within-period price impact that is even higher than in the static model. We do not solve for an equilibrium of the dynamic model for an arbitrary number E of exchanges, given the difficult-to-solve infinite regress of beliefs about beliefs concerning the aggregate asset inventory of strategic traders. Rather than addressing equilibria for general E, we instead construct an equilibrium for a specific number E of exchanges with the property that the associated equilibrium is perfect Bayesian and implements efficient trade. This equilibrium is tractable because efficient trade dramatically simplifies the inference problem of each trader, given that the sum of exchange prices perfectly reveals the aggregate inventory at the end of each trading date. We find that the efficient number of exchanges is invariant to trading frequency, and is the same as that of the static model.

The remainder of the paper is organized as follows. Section 2 provides additional background on exchange market fragmentation and related research. Section 3 gives the setup of the most basic version of our model. Section 4 characterizes properties of the equilibrium. Section 5 presents the implications of fragmentation on price impact, allocative efficiency, and price informativeness. Section 6 studies a formulation of the model in which traders observe the aggregate asset endowment before order submission. Section 8 summarizes the results of various model extensions. Section 7 solves for the efficient number of exchanges in a dynamic formulation of the model with cross-period cross-exchange inference. Section 9 offers some concluding remarks and discusses some potentially important effects that are not captured by our model. Appendices contain proofs and model extensions.

2 Background

We focus in this paper on "visible fragmentation," that is, fragmentation across different lit exchanges (meaning trade venues at which market-clearing prices are set), rather than fragmentation between lit exchanges and size-discovery venues, which cross buy and sell orders at prices that are set on lit exchanges (Körber, Linton, and Vogt, 2013; Zhu, 2014; Degryse, De Jong, and van Kervel, 2015; Duffie and Zhu, 2017; Antill and Duffie, 2019).

In Europe and the U.S., exchange trading is highly fragmented. Budish, Lee, and Shim (2019) document that in the U.S., as of early 2019, annual trade of about one trillion shares is split across 13 U.S. exchanges, and that cross-exchange shares of total exchange-traded volume are stable over time, with 5 exchanges each handling over 10 percent of total exchange volume. Essentially all equities trade on every exchange, with significant volumes of each equity executed on multiple exchanges.[2] Broadly speaking, similar patterns apply to European financial markets (Gresse et al., 2012; Degryse, De Jong, and van Kervel, 2015; Foucault and Menkveld, 2008). This high degree of trade fragmentation is in part a consequence of regulations such as Regulation NMS in the US and MiFid II in Europe, which encourage exchange entry and competition.

[2] Pagnotta and Philippon (2018) and Budish, Lee, and Shim (2019) display the striking facts graphically.

There has been a lonstanding debate (Stoll, 2001) over whether fragmenting trade across exchanges harms market efficiency, in various respects. Empirical findings have been mixed (O'Hara and Ye, 2011; Gomber et al., 2017). Some researchers find that fragmentation has generally been beneficial. For example, O'Hara and Ye (2011), using data from U.S. trade reporting facilities, find that execution speeds are faster, transaction costs are lower, and prices are more efficient when the market is more fragmented. Degryse, De Jong, and van Kervel (2015) analyze a sample of Dutch stocks and measure the degree of visible fragmentation. They find that liquidity, when aggregated over all lit trading venues, improves with fragmentation. Foucault and Menkveld (2008) analyze Dutch stocks and arrive at a similar conclusion. Boehmer and Boehmer (2003) find evidence of improved liquidity when the NYSE began trading ETFs that are also listed on the American Stock Exchange. Gresse (2017), De Fontnouvelle, Fishe, and Harris (2003), Aitken, Chen, and Foley (2017), Hengelbrock and Theissen (2009), Félez-Viñas (2017), and Spankowski, Wagener, and Burghof (2012) generally find that visible fragmentation reduces bid-ask spreads.

Other research, however, suggests less beneficial effects of fragmentation. For example, Bennett and Wei (2006) find that when equity trading migrated from Nasdaq to the NYSE, where trade is more consolidated, there was a decrease in execution costs and an improvement in price efficiency. Chung and Chuwonganant (2012) show that price impact increased following the introduction of Regulation NMS. (In our model, as we have noted, fragmentation indeed reduces market depth, yet increases allocative efficiency and overall price informativeness.) Gentile and Fioravanti (2011) find that MiFID-induced fragmentation "does not have negative effects on liquidity, but it reduces price information efficiency. Moreover, in some cases it leads primary stock exchanges to lose their leadership in the price discovery process." For small-firm equities, Gresse et al. (2012), Gresse (2017), and Degryse, De Jong, and van Kervel (2015) find that market depth declines with sufficient fragmentation, consistent with our theoretical results. Bernales et al. (2018) find that the 2009 consolidation of Euronext's two distinct order books for the same equities was followed by a reduction in bid-offer spreads. Haslag and Ringgenberg (2016) find causal evidence that although fragmentation reduces bid-offer spreads for the equities of large firms, the opposite applies to small firms.

While the empirical evidence regarding the implications of fragmentation are mixed, most of the theoretical literature has shown that visible fragmentation is harmful. For example, Mendelson (1987) shows that fragmentation may isolate individuals for whom there are mutually beneficial trades, because they are located at different venues. Chowdhry and Nanda (1991) show that adverse selection caused by asymmetric information worsens as markets fragment. Baldauf and Mollner (2020) find that welfare is harmed by the ability of fast traders to snipe across fragmented markets.

Of the few theory papers showing that fragmentation may be beneficial, perhaps the closest to ours is Malamud and Rostek (2017). As in our model, they consider a multi-exchange demand submission game in which each exchange operates a double auction. They show that, in certain settings, when agents' risk preferences are sufficiently heterogeneous, fragmented markets can produce outcomes that are welfare superior to centralized markets. Crucially, however, they assume that agents are able to submit demand schedules to each exchange that are contingent on the realization of prices on all exchanges. The channel by which fragmentation is beneficial in our model is not related to that of Malamud and Rostek (2017), and does not rely on heterogeneous risk aversion or cross-exchange contingent order mechanisms, which are extremely rare in practice (Wittwer, 2020).

Of the theoretical papers mentioned, the majority assume that traders are restricted to trade on a strict subset of all trading venues. For example, Pagano (1989) shows that fragmented markets are less stable, in that traders tend to concentrate at a single market venue, at which liquidity is greatest. However, regulations promoting exchange competition may foster fragmentation. If traders are strategic about their price impacts it seems natural to assume they are aware of the option to trade on multiple exchanges simultaneously. The costs of order splitting are economically small (Budish, Lee, and Shim, 2019). So-called Smart Order Routing Technology makes order splitting convenient and practical (Gomber et al., 2016). In our model, strategic traders frictionlessly trade on all exchanges. Empirical research (Malinova and Park, 2019; Menkveld, 2008; Chakravarty et al., 2012; Gomber et al., 2016) finds evidence that some investors strategically split their orders across multiple exchanges, and also between exchanges and size-discovery venues such as dark pools.

Methodologically, our model contributes to the literature on multi-auction demand-function submission games, including work by Wilson (1979), Klemperer and Meyer (1989), and Malamud and Rostek (2017). Within this literature, our paper, like prior work by Wittwer (2020) and a contemporaneous paper by Rostek and Yoon (2020), addresses markets with multiple exchanges. While Wittwer (2020) and Rostek and Yoon (2020) focus on the welfare implications of connecting exchanges through the ability to submit orders contingent on cross-exchange prices, we consider only the common case in practice of "disconnected markets." As opposed to Wittwer (2020) and Rostek and Yoon (2020), we focus on the implications for allocative efficiency and price informativeness of increasing the number of exchanges (fragmentation), and we include a dynamic analysis that captures the implications of cross-time cross-exchange price impact while still showing that enough fragmentation can achieve allocative efficiency.

Since the work of Hamilton (1979), the literature has explored the key tension between the benefit of fragmentation associated with increased competition between exchanges and between specialists, which drives down bid-offer spreads and trading fees, as suggested by the theory of Hall and Rust (2003), versus the cost of fragmentation associated with decreased market depth.[3] Although fragmentation does indeed reduce market depth in our model, consistent with earlier work, we believe that we are the first to point out the benefit of fragmentation associated with increased order aggressiveness, arising from the ability of strategic traders to shield orders on a given exchange from price impacts incurred on other exchanges.

3 Baseline Model

This section presents the setup of our baseline model. All primitive random variables are defined on a complete probability space, $(\Omega, \mathcal{F}, \mathbb{P})$. There is a single asset with a payoff, denoted π, that is a finite-variance random variable with mean μ_π.

We model a market whose agents, called "traders," are of two types: "liquidity" and "strategic." For notational simplicity, we let N denote both the finite set of strategic traders and its cardinality, which is assumed to be at least 3. The only primitive information available to strategic trader i is the trader's own endowment of the asset, $X_i \sim N(0, \sigma_X^2)$. We assume that endowments are i.i.d across traders.

Trade of the asset takes place in a single period on each of a finite number of identical exchanges. For notational simplicity, we let E denote both the set and number of exchanges. Each exchange runs a double auction mechanism. Strategic trader i submits a measurable demand schedule $f_{ie} : \mathbb{R}^2 \to \mathbb{R}$ to exchange e specifying the quantity $f_{ie}(X_i, p)$ of the asset demanded by trader i at any given price $p \in \mathbb{R}$ on exchange e. We emphasize that the demand schedule submitted to a given exchange cannot depend on prices or any other information emanating from the other exchanges. A demand schedule can be viewed as a package of limit orders, each of which is an offer to purchase or sell a given amount of the asset at a given price.[4] Liquidity traders collectively submit an exogenously given quantity of market orders to exchange e denoted $Q_e \sim N(0, \sigma_Q^2/E)$.

We assume that the supply of market orders is i.i.d across exchanges and that $\{X_i \,|\, i \in N\}$, $\{Q_e \,|\, e \in E\}$, and π are independent. We relax these distributional assumptions in Section 6 and in extensions considered in the Appendix H. A useful interpretation of the

[3] For a recent empirical contribution exploring this tradeoff, see Haslag and Ringgenberg (2016).

[4] In this sense, $f(X_i, p)$, if positive, is the aggregate quantity of the limit orders to buy at a price of p or higher, and if negative is the aggregate quantity of the limit orders to sell at price of p or lower. The space of linear combinations of limit orders is dense, in the sense of Brown and Ross (1991), in the space of monotone demand functions.

above assumptions on liquidity trade is that there is a large number of liquidity traders, independent of the number of exchanges in operation, who are spread evenly across exchanges and trade independently of one another.

Given a collection $f = \{f_{ie} \,|\, i \in N, e \in E\}$ of demand schedules, the price on exchange e, if it exists, is a solution[5] p_e^f to the market-clearing condition

$$\sum_{i \in N} f_{ie}(X_i, p_e^f) = Q_e. \tag{1}$$

If there does not exist a unique market clearing price, we assume that no trades are executed. We restrict attention to equilibria consisting of demand schedules with the property that p_e^f is uniquely determined.[6] Based on (1), trader i is able to determine the impact of his or her own demand on the market-clearing price given the conjectured demand schedules of the other traders.

The preferences of the strategic traders are quasi-linear in cash compensation with a quadratic holding cost. Specifically, given a collection $f = \{f_{ie} \,|\, i \in N, e \in E\}$ of demand schedules the associated payoff of trader i is

$$U_i(f) = \left(X_i + \sum_e f_{ie}(X_i, p_e^f) \right) \pi - b \left(X_i + \sum_e f_{ie}(X_i, p_e^f) \right)^2 - \sum_e p_e^f f_{ie},$$

for some $b > 0$. The quadratic term represents a cost for bearing the risk or other costs associated with holding a post-trade position in the asset. Preferences of this form are popular in the market microstructure literature (Vives, 2011; Rostek and Weretka, 2012; Du and Zhu, 2017; Sannikov and Skrzypacz, 2016). Sannikov and Skrzypacz (2016) provide a microfoundation.

An equilibrium is defined as a collection $f = \{f_{ie} \,|\, i \in N, e \in E\}$ of demand schedules with the property that for each strategic trader i the demand schedules $f_i = \{f_{ie} \,|\, e \in E\}$ solve

$$\sup_{\hat{f}} \; \mathbb{E}[U_i(\hat{f}, f_{-i})],$$

where as usual f_{-i} denotes the collection $\{f_j \,|\, j \neq i\}$ of other traders' demand schedules. The model we have specified is a typical demand-function submission game in the sense of Wilson (1979) and Klemperer and Meyer (1989), extended to allow for multiple exchanges. Multi-exchange demand function submission games were earlier analyzed by Malamud and

[5] That is, p_e^f is a random variable such that for each state $\omega \in \Omega$, $\sum_{i \in N} f_{ie}(p_e^f(\omega), X_i(\omega)) = Q_e(\omega)$.

[6] For this, it suffices that, for each $x \in \mathbb{R}^N$, the aggregate demand function $p \mapsto \sum_i f_{ie}(p, x_i)$, which is monotone, is strictly monotone, continuous, and unbounded below and above.

Rostek (2017) and Wittwer (2020).

We conclude this section with an interpretation of the distinction between strategic and liquidity traders. A strategic trader may be viewed as an agent who is sophisticated, internalizes price impact, is able to easily split orders across multiple trading venues, has a relatively low aversion to owning assets, and has a relatively large initial endowment of the asset. A liquidity trader, on the other hand, may be viewed as an agent who is not sophisticated about price impacts, has high aversion to holding assets (thus exercising no discretion in the liquidation of the assets), and has a small initial asset holding, and who therefore submits market orders with no price sensitivity. Liquidity traders are a typical modeling device for settings such as ours in which one wishes to avoid perfect inference of fundamental information from price observations. In our case, the fundamental information to be inferred does not concern asset payoffs but rather the aggregate endowment of strategic traders. Traders have payoff relevant private information about their own endowments but no private information about asset payoffs. We will show that our main results are not driven by the effect of "donations" from liquidity traders to strategic traders.

4 A Symmetric Affine Equilibrium

We can prove the existence and uniqueness of a symmetric affine equilibrium defined by demand schedules of the form

$$f_{ie}(p, X_i) = \Delta_E - \alpha_E X_i - \zeta_E p, \tag{2}$$

for constants Δ_E, α_E, and ζ_E that do not depend on the trader or particular exchange, but do depend on the number E of exchanges.

Using (1) it can be shown that the slope of the inverse residual supply curve facing each agent in each exchange is equal to

$$\Lambda_E \equiv \frac{1}{(N-1)\zeta_E} \tag{3}$$

which we refer to as inverse market depth, or simply as "price impact". Each strategic trader is aware that by deviating from the equilibrium demand schedule and demanding an additional unit on a given exchange, the trader will increase the market-clearing price on that exchange by Λ_E. Price impact is a perceived cost to each strategic trader, but is not a social cost because the payment incurred by any trader is received by another. As emphasized by Vayanos (1999), Rostek and Weretka (2015), and Du and Zhu (2017), the strategic avoidance of price impact through the "shading" of demand schedules is socially costly because it reduces the total gains from the beneficial reallocation of the asset.

By using the form of the demand schedules in (2) we can compute that the final asset position of strategic trader i is

$$(1 - E\alpha_E)X_i + E\alpha_E \frac{\sum_{j\in N} X_j}{N} + \frac{\sum_{e\in E} Q_e}{N}. \tag{4}$$

Generically in the parameters of the model, the equilibrium allocation is inefficient. Given the non-discretionary liquidation $\sum_{e\in E} Q_e$ by liquidity traders, the efficient allocation is one in which each strategic trader receives an equal share of the aggregate supply of the asset, which is

$$\overline{q} = \frac{1}{N}\left(\sum_{e\in E} Q_e + \sum_{i\in N} X_i\right).$$

Inspecting (4), this efficient sharing rule corresponds to the case of $E\alpha_E = 1$. By Jensen's inequality, this produces the efficient allocation because traders have symmetric convex holding costs. Since preferences are quasi-linear in cash compensation, this is also the welfare-maximizing allocation, in that any other allocation would be strictly Pareto dominated by this efficient sharing rule, after allowing voluntary initial side payments.

The equilibrium allocation defined by (4) becomes less efficient the farther is $E\alpha_E$ from 1. This is because replacing $E\alpha_E$ in (4) with a number farther from 1 results in a mean-preserving spread in the cross-sectional distribution of the asset to strategic traders, state by state. Jensen's inequality, applied cross-sectionally in each state $\omega \in \Omega$, then implies an increase in the sum across traders of quadratic holding costs.

The following theorem collects several properties of symmetric affine equilibria. Of primary interest is the property that in the presence of non-trivial liquidity trade, the allocation becomes more efficient as market fragmentation E increases, up to the point at which $E\alpha_E = 1$, and then becomes increasingly less efficient. We will explore this issue in more depth in Section 5. Our proof of the theorem, found in the Appendix B, applies the calculus of variations to verify that a particular set of candidate equilibrium demand coefficients $(\Delta_E, \alpha_E, \zeta_E)$ does in fact uniquely correspond to an equilibrium.

Theorem 1. *For each positive integer number E of exchanges, there exists a unique symmetric affine equilibrium. The associated demand-function coefficients $(\Delta_E, \alpha_E, \zeta_E)$ form the unique solution to appendix equations (34), (35), and (36). Moreover:*

1. *The market-clearing price on exchange e is*

$$p_e^* = \frac{N-1}{N}\Lambda_E\left[N\Delta_E - Q_e - \alpha_E \sum_{i\in N} X_i\right]. \tag{5}$$

2. *The associated price-impact coefficient is*

$$\Lambda_E = \frac{2b(1+\gamma_E(E-1))}{N-2}, \tag{6}$$

where

$$\gamma_E = \frac{E\alpha_E^2\sigma_X^2(N-1)}{E\alpha_E^2\sigma_X^2(N-1)+\sigma_Q^2} \tag{7}$$

is the conditional correlation between prices in any two distinct exchanges e and e' from the perspective of any strategic trader i, given X_i.

3. *The final asset position of strategic trader i is given by (4).*

4. *If there is no liquidity trading, in that $\sigma_Q^2 = 0$, then the equilibrium allocation does not depend on the number E of exchanges.*

5. *If $E = 1$ or $\sigma_Q^2 = 0$, then the final asset position of strategic trader i is*

$$\frac{\Lambda_1}{\Lambda_1+2b}X_i + \frac{2b}{\Lambda_1+2b}\frac{1}{N}\sum_{j\in N} X_j + \frac{\sum_{e\in E} Q_e}{N},$$

where $\Lambda_1 = 2b/(N-2)$.

6. *If $\sigma_Q^2 > 0$, then $E\alpha_E$ is strictly monotone increasing in E and converges to $N/(N-1)$. It follows in this case that a market with only one exchange is strictly dominated, from the viewpoint of allocative efficiency, by a market with any larger number of exchanges.*

Part 5 of Theorem 1 implies that with a single exchange, the fraction of the endowment retained by a trader is increasing in price impact, Λ_1. In a centralized market, price impact avoidance is the only source of allocative inefficiency. As we have described and will later elaborate, the effect of price impact avoidance on allocative efficiency can be mitigated by increasing the degree of market fragmentation. In the next section, we analyze the forces behind this and other effects of market fragmentation. But, as stated in part 6 of Theorem 1, any degree of fragmentation is socially preferred to concentrating all trade on a single exchange.

5 The Effects of Fragmentation

We present several predictions of our model, beginning first with the effects of fragmentation on price impact.

5.1 Price impact

Part 2 of Theorem 1 provides the equilibrium relationship between price impact and the correlation between exchange prices. This relationship reflects the effect on trade demand of cross-exchange inference from prices. The quantity purchased by trader i on exchange e at a given p_e, $f_{ie}(X_i, p_e)$, depends in part on the expectation of the quantities that trader i will execute on the other exchanges, conditional on X_i and p_e.

To illustrate, suppose for example that in state $\omega \in \Omega$ trader i is a buyer of the asset at the equilibrium price in exchange e. If the observed price outcome $p_e(\omega)$ was lowered, trader i would assign a higher conditional likelihood to lower prices on the other exchanges because strategic traders' demands are positively correlated on any two exchanges which implies a positive cross-exchange price correlation, γ_E. But trader i submits demands to the other exchanges before observing p_e. Thus, the lower is $p_e(\omega)$ the higher is the conditional expected quantity executed by trader i on the other exchanges. If $p_e(\omega)$ is lowered, the marginal utility of trader i for purchasing a unit on exchange e would decline. Due to cross-exchange inference, the quantity trader i optimally purchases on exchange e in response to a decrease in price $p_e(\omega)$ is smaller relative to if there was no cross-exchange correlation. Analogous reasoning can be applied to show that due to cross-exchange inference, the quantity trader i optimally purchases on exchange e in response to an increase in price $p_e(\omega)$ decreases relative to if there was no cross-exchange correlation. Overall, the cross-exchange price inference channel reduces the steepness (absolute slope) of the demand schedule of trader on each exchange with respect to price. The result, by (3), is that price impact rises. Since this channel is not present when there is a single exchange, price impact is always higher in a fragmented market than in a centralized market.

We now discuss comparative static results describing the effects of changes in the variance σ_Q^2 of liquidity trade demand and the number E of exchanges on price impact. As σ_Q^2 increases, prices in different exchanges becomes less correlated, so price impact declines, eventually converging to that of a single exchange market as σ_Q^2 tends to infinity. Thus, price impact is lower in markets with noisier liquidity trader supply because the cross-exchange inference channel is weaker. The following proposition characterizes how price impact changes as the number of exchanges increases holding fixed all other model parameters.

Proposition 1. *The price-impact coefficient Λ_E is strictly monotone increasing in the number E of exchanges. If the variance σ_Q^2 of liquidity trade demand is zero, then $\lim_{E\to\infty}\Lambda_E = \infty$. If $\sigma_Q^2 > 0$, then*

$$\lim_{E\to\infty}\Lambda_E = \frac{2b}{N-2}\left(1+\frac{N^2\sigma_X^2}{(N-1)\,\sigma_Q^2}\right),$$

and γ_E declines strictly monotonically to zero as $E \to \infty$.

Proposition 1 states that, with greater market fragmentation, price impact is higher and (in the presence of nontrivial liquidity trade), prices are less correlated. Without liquidity trade ($\sigma_Q^2 = 0$), price impact diverges as the number of exchanges diverges, because γ_E is equal to one. But with liquidity trade ($\sigma_Q^2 > 0$) price impact converges to a finite value. Because price impact depends on $\gamma_E(E-1)$, this follows from the fact that γ_E declines at a rate proportional to $\frac{1}{E}$. The intuition is that as the number of exchanges increases, the expected quantity traded on a given exchange decays at rate $\frac{1}{E}$, which in turn causes the variability in prices due to strategic traders' orders to decay at a rate proportional to $\frac{1}{E^2}$. Since the variability in prices due to exchange-specific liquidity trade is σ_Q^2/E, this implies that γ_E must decline at the rate $\frac{1}{E}$, so that price impact converges.

Figure 1 illustrates the relationship between price impact and the number of exchanges, for different cases of the number N of strategic traders. As illustrated, price impact converges faster when there are more strategic traders. For instance, consider the case of $b = 1/2$, $N = 5$ and $E = 100$. Without liquidity trade, the price impact is $\Lambda_E = 33$. However, with $\sigma_Q^2 > 0$, and strategic traders whose endowments are 10 times more uncertain (in terms of variance) than aggregate liquidity trader supply (in that $\sigma_X^2/\sigma_Q^2 = 10$), price impact drops to approximately 10. As σ_X^2/σ_Q^2 falls below 10, γ_E is reduced and, because of this, price impact is further reduced.

5.2 Allocative Efficiency

We have just shown that price impact is higher in more fragmented markets. However, by Theorem 1, when there is no liquidity trade ($\sigma_Q^2 = 0$), even though price impact diverges as E tends to infinity, total trade aggressiveness is unaffected and the equilibrium allocation remains constant. Moreover, when $\sigma_Q^2 > 0$, even though price impact increases with fragmentation, total trade aggressiveness actually *increases*. One might have expected that the rise in price impact would lead to a reduction in trade aggressiveness and thus lower allocative efficiency, but this is not the case. We turn now to a resolution of this superficial paradox.

As fragmentation rises, price impact increases, but traders can better evade the overall

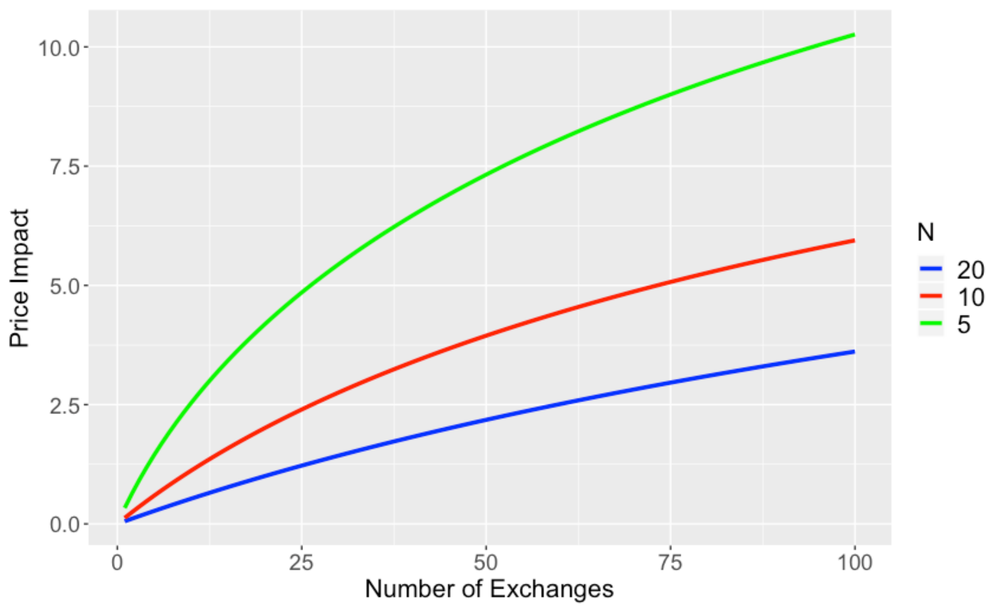

Figure 1: Variation of price impact Λ_E with the number E of exchanges, for various cases of N, the number of strategic traders. In all cases, the variance-aversion coefficient is $b = 1/2$ and the ratio σ_X^2/σ_Q^2 of the variance of strategic-trader asset endowment to total liquidity trade quantity is 10.

cost of price impact by shredding their orders across exchanges. This is because traders bear the cost of price impact on a given exchange only to the extent of the trades executed on that exchange. By order splitting, a trader can shield an order on a given exchange from the price impact of units executed on the other exchanges. When there are more exchanges, the purchase of an additional unit on a given exchange affects a smaller fraction of the total quantity traded. When there is no liquidity trade ($\sigma_Q^2 = 0$) this effect exactly offsets the rise in price impact, leaving the overall agressiveness of a trader's demand invariant to the number of exchanges. When $\sigma_Q^2 > 0$ price impact does not rise quickly enough to offset the effect of increased aggressiveness through order splitting. At low levels of fragmentation, this increase in trade aggressiveness is beneficial for allocative efficiency. But when markets become sufficiently fragmented, the incremental aggressiveness is inefficient, in that $E\alpha_E$ increases past the point of efficiency, at which $E\alpha_E = 1$ (up to $\frac{N}{N-1}$). We emphasize, however, that trade never becomes so aggressive that fragmentation leads to a loss of allocative efficiency relative to that of a market with a single exchange.

By equation (4), the number of exchanges that maximizes allocative efficiency is that for which $E\alpha_E$ is closest to 1.

Proposition 2. *Suppose* $\sigma_Q^2 > 0$. *Let*

$$E^* = 2 + \frac{2}{N-2} + \frac{N-1}{N-2}\frac{N\sigma_X^2}{\sigma_Q^2}.$$

If E^* is an integer, the unique symmetric affine equilibrium for a market with E^* exchanges achieves an efficient allocation of the asset, by allocating an equal amount $\bar{q}$ of the asset to each strategic trader. In general, the number of exchanges that maximizes allocative efficiency is either $\lfloor E^* \rfloor$ or $\lceil E^* \rceil$.

By Proposition 2, the optimal number of exchanges is finite, is at least 2, and depends crucially on the ratio of the variance of the endowment of strategic traders to the variance of the total amount of liquidity trade, σ_X^2/σ_Q^2. This ratio determines γ_E, as seen in equation (7), which in turn determines price impact. As σ_X^2/σ_Q^2 rises, price impact is higher and more fragmentation is needed to offset the adverse effect of price impact with the beneficial effect of increasing the number of exchanges over which strategic traders can split their orders.

It is perhaps surprising that the socially optimal number of exchanges is finite. The intuition associated with order splitting might suggest that inefficiency due to price impact avoidance should only disappear in the limit as the number of exchanges tends to infinity. Only as this limit is approached do agents trade a negligible quantity on any one exchange, so that the marginal unit traded affects the price only negligibly. It turns out, however, that fragmentation introduces a different inefficiency. At the point in time at which traders submit demands to a given exchange, they are unaware of the quantities they will ultimately purchase on other exchanges. Moreover, traders are asymmetrically informed about trading opportunities on the other exchanges because they have different endowments, and equilibrium prices depend on the aggregate endowment. This is a force leading agents to trade more aggressively in fragmented markets that is eventually adverse to efficiency, and that has no counterpart in a centralized market.

Figure 2 illustrates the results of this section. As shown, $E\alpha_E$ is strictly increasing in fragmentation and can exceed the socially efficient level. The socially efficient number of exchanges increases with σ_X^2/σ_Q^2.

5.3 Price Informativeness

Our finding that trade aggressiveness increases with market fragmentation has natural implications for price informativeness. By price informativeness, we mean the degree to which prices reveal information about the average endowment $\overline{X} = \sum_{i\in N} X_i/N$ of strategic traders. This notion is especially relevant when viewing our model as though a snapshot of a dynamic market in which liquidity trade is serially uncorrelated and the aggregate strategic endowment is a persistent markov process. In such a setting, the aggregate endowment of strategic traders is a sufficient statistic for inference regarding future prices and future aggregate endowments.

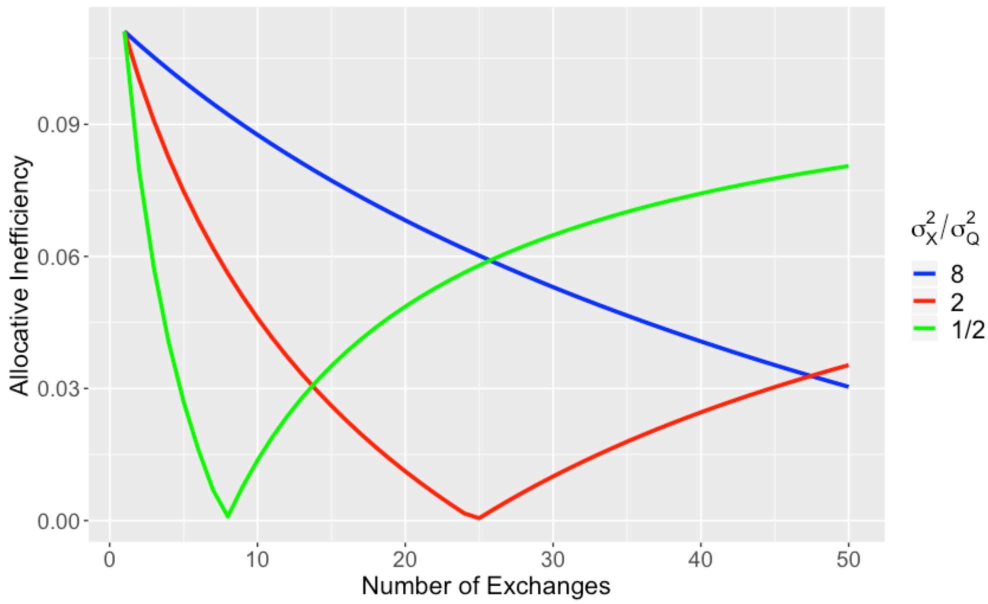

Figure 2: We plot equilibrium allocative inefficiency as measured by $|1 - E\alpha_E|$ against the number of exchanges for different values of the ratio σ_X^2/σ_Q^2 of the variance of the endowment of a strategic trader to the variance of the total amount of liquidity trade. In all cases, the number N of strategic traders is 10. Allocative inefficiency, $|1 - E\alpha_E|$, does not depend on the variance-aversion coefficient b.

Because of the joint normality of prices and endowments in our model, the conditional variance of $\overline{X}$ given exchange prices is an unambiguous metric for price informativeness. Our results are summarized in Proposition 3.

Proposition 3. *Suppose that the variance σ_Q^2 of liquidity trade is not zero. Then:*

1. *For any exchange e, $\text{var}(\bar{X} \mid p_e^*)$ is strictly monotone increasing in the number E of exchanges and converges to $\text{var}(\bar{X})$ as E goes to ∞.*

2. *$\text{var}(\bar{X} \mid \{p_e^* : e \in E\})$ is strictly monotone decreasing in E.*

In words, Proposition 3 shows that the informativeness of the price on any individual exchange worsens with fragmentation but overall price informativeness, taking into consideration information from all exchange prices, improves.

6 The Case of Observable Aggregate Endowment

In this section we present a simplified version of the model in which the aggregate endowment of strategic traders is publicly observable in order to demonstrate the welfare benefits of fragmentation in a setting that does not require liquidity traders or Gaussian X_e and Q_e. Under the assumption of public aggregate endowment the equilibrium price in a given exchange is a linear combination of the aggregate endowment and exchange-specific liquidity trade. As a result, conditional on $\overline{X}$, prices in any two exchanges are uncorrelated so that traders do not need to make cross-exchange price inferences. This allows us to shut down the cross-exchange inference channel and study the welfare benefits of order splitting in isolation.

We retain the model setup of Section 3 with the exceptions that, for any exchange e and any trader i, (a) neither Q_e nor X_i is necessarily normally distributed though Q_e still has mean zero and (b) trader i observes[7] the private endowment X_i and the average endowment $\overline{X}$. The following theorem characterizes the equilibrium of this model.

Theorem 2. *For each number E of exchanges, there exists a symmetric affine equilibrium. If, in addition, for each e, Q_e has full support on $\mathbb{R}$, then the equilibrium is unique in the class of symmetric affine equilibria and has the following properties.*

1. *The price-impact coefficient $\Lambda_E = 2b/(N-2)$ does not depend on the number E of exchanges.*

2. *The price on exchange e is*

$$p_e^* = -2b\left(\bar{X} + \frac{Q_e}{N-2}\frac{N-1}{N}\right) + \mu_\pi.$$

3. *The final asset position of trader i is*

$$\frac{\Lambda_E}{\Lambda_E + 2bE}X_i + \frac{2bE}{\Lambda_E + 2bE}\bar{X} + \frac{\sum_{e\in E} Q_e}{N}.$$

4. *The total expected equilibrium payment $-\mathbb{E}\left[\sum_{e\in E} p_e^* Q_e\right]$ of liquidity traders is invariant to the number E of exchanges and equal to*

$$\frac{\mathrm{var}(\sum_{e\in E} Q_e)}{N-2}\frac{N-1}{N}.$$

5. *Allocative efficiency is increasing in the number E of exchanges. As E diverges, the allocation converges to the efficient allocation, $\bar{q}$ to each strategic trader.*

In this setting, price impact is a constant that does not depend on the level of fragmentation because there is no cross-exchange inference effect. By part 3 of the theorem, more fragmentation is unambiguously beneficial in this setting. In the limit as E tends to infinity, the fully efficient allocation obtains. The benefits of fragmentation arise entirely from the

[7] That is, the demand submitted by trader i on exchange e is a measurable function $f_{ie} : \mathbb{R}^3 \to \mathbb{R}$ that, at any price p, determines the demand $f_{ie}(X_i, \sum_{j\in N} X_j, p)$.

beneficial effects of increased order aggressiveness associated with order splitting. The above equilibrium exists even when there is no liquidity trade, though the presence of liquidity trade is needed for equilibrium uniqueness. Even in the presence of liquidity traders, the expected payment of liquidity traders to strategic traders is invariant to market fragmentation. Thus the beneficial effect of fragmentation is not related to the exploitation of liquidity traders by strategic traders.[8] In the model of Section 3, the liquidity traders were only a convenient modeling device for breaking the perfect correlation in exchange prices. Budish, Cramton, and Shim (2015) note that, at a sufficiently high sampling frequency, the prices of similar assets on different exchanges are virtually uncorrelated, empirically.

7 A Dynamic Model

One might guess that market fragmentation, though capable of alleviating within-period price impact, would not be as effective in a dynamic setting in supporting allocative efficiency, given the strategic avoidance of cross-time cross-exchange price impact. That is, when submitting a trade on exchange e at period t, a trader has no concern about adversely influencing the price on another exchange in the same period, but does internalize the resulting impact on the prices on all exchanges in period $t+1$, given the inference about aggregate inventory that is drawn by other traders from observing p_{et}. Nevertheless, in this section, we show that market fragmentation allows efficient trade even in a dynamic setting, despite the associated cross-period cross-exchange price impact and higher within-period price impact. Moreover, the efficient number of exchanges is invariant to trading frequency, and is the same as that of the static model.

7.1 Setup

Trade occurs at each of a discrete set of times separated by some duration Δ. A positive integer t denotes the t-th trading date. As in the baseline static model, E exchanges operate separate double auctions for a single asset at each trade time. The asset pays π_t at date t, post-trade, where $\pi_1, \pi_2, \ldots$ are independent with common mean $\mu_\pi \Delta$. Liquidity traders supply a Gaussian quantity Q_{et} of the asset to exchange e at trade date t, independent

[8] In the setting of Section 5, our results are not driven by donations from liquidity traders, but liquidity traders do pay more in expectation as fragmentation increases. In the model of Section 3, the total expected payment to strategic traders is

$$\mathbb{E}\left(\sum_{e \in E} p_e^* Q_e\right) = \frac{N-1}{N} \Lambda_E \sigma_Q^2. \tag{8}$$

which is strictly increasing in E since Λ_E is strictly increasing.

across exchanges and dates with common mean zero and variance $\sigma_Q^2\Delta$. At date t, trader i receives a Gaussian inventory shock ϵ_{it}, prior to trade, that has mean zero and variance $\sigma_\epsilon^2\Delta$, independent across trading dates and traders. The inventory shocks, liquidity trader supplies, and the asset payoffs are independent.

The post-trade inventory of trader i at period t is

$$X_{it} = X_{i,t-1} + \sum_{e\in E} q_{ie,t-1} + \epsilon_{it}, \tag{9}$$

where q_{iet} is the quantity purchased by trader i on exchange e at period t. For $t = 0$, we set $X_{i,t-1} + \sum_{e\in E} q_{ie,t-1} = 0$.

During the time interval $[t\Delta, (t+1)\Delta)$, the net payoff to trader i, discounted to the beginning of the interval at the rate $r > 0$, is the total initial payoff from asset holdings, net of asset purchase costs, plus discounted inventory holding costs, given by

$$\begin{aligned} F_{it}(q_{it}) &= \pi_t\left(X_{it} + \sum_{e\in E} q_{iet}\right) - \sum_{e\in E} p_{et}q_{iet} - \int_0^\Delta e^{-rs}\tilde{b}\left(X_{it} + \sum_{e\in E} q_{iet}\right)^2 ds \\ &= \pi_t\left(X_{it} + \sum_{e\in E} q_{iet}\right) - \sum_{e\in E} p_{et}q_{iet} - b\left(X_{it} + \sum_{e\in E} q_{iet}\right)^2, \end{aligned} \tag{10}$$

where $q_{it} = (q_{i1t}, ..., q_{iEt})$ and

$$b = \tilde{b}\frac{1 - e^{-r\Delta}}{r}.$$

Our formulation is in the spirit of Vayanos (1999), differing mainly in that we allow multiple exchanges, introduce liquidity traders, and assume a different inventory preference model. For tractability, a significant part of the analysis in Vayanos (1999) focuses on the case in which σ_ϵ^2 tends to zero. Our analysis applies to arbitrary σ_ϵ^2.

We do not solve for an equilibrium of the model for an arbitrary number E of exchanges because of the problem of infinite regress of beliefs described in the conclusion of Vayanos (1999). In the presence of liquidity traders, strategic traders choose their trades based on their beliefs about the aggregate market asset inventory, as well as beliefs about other traders' beliefs about aggregate inventory, beliefs about the beliefs of other traders about their own beliefs, and so on, causing the state space to explode. To our knowledge, there does not exist a dynamic trading model with double auctions in which traders filter information from prices so as to discern strategic trading from liquidity trading. Rather than addressing equilibria for general E, we instead construct an equilibrium for a specific number E of exchanges with the property that the associated equilibrium is perfect Bayesian and implements efficient trade.

This equilibrium is tractable because efficient trade dramatically simplifies the inference problem of each trader, given that the sum of exchange prices perfectly reveals the aggregate inventory at the end of each trading date.

7.2 An Equilibrium with Efficient Trade

In this section we briefly sketch the derivation of an efficient equilibrium and characterize its key properties, including the associated number E of exchanges. A formal derivation is given in Appendix E.

To start, we conjecture that there exists a number E of exchanges such that in equilibrium each trader i submits the demand schedule to exchange e given by

$$f_{iet}(X_{it}, p_{et}, B_t) = -\frac{1}{E} X_{it} - \zeta p_{et} + \rho B_t + \chi, \tag{11}$$

for some constants ζ, ρ, and χ to be determined, where B_t is defined recursively by $B_0 = 0$ and

$$B_t = NE\rho B_{t-1} + NE\chi - \zeta N \sum_{e \in E} p_{e,t-1}. \tag{12}$$

We later interpret B_t as a variable related to trader beliefs about the aggregate supply of the asset. Given the conjectured form (11) of the demand function, market clearing implies that the equilibrium price on exchange e is

$$p_{et} = \frac{N\rho B_t + N\chi - Q_{et} - \frac{1}{E}\sum_{j \in N} X_{jt}}{\zeta N}. \tag{13}$$

The post-trade aggregate inventory of strategic traders at date t is

$$W_t = \sum_{j \in N} X_{j,t} + \sum_{e \in E} Q_{e,t}.$$

Substituting (13) into (11) and summing across $e \in E$, we verify that the final inventory of trader i at date t is efficient and equal to W_t/N. By substituting (13) into (12) we see that along the equilibrium path, B_t is equal to W_{t-1}. However, if any given trader has deviated from the equilibrium strategy prior to date t, it is possible that $B_t \neq W_{t-1}$. Nonetheless, even if traders had deviated prior to date t, any trader who has not deviated must believe that $W_{t-1} = B_t$ with probability 1 because the Gaussian liquidity trading/inventory shocks ensure that deviations by strategic traders are undetectable. Thus, any given trader i must believe that any other trader j believes that $W_{t-1} = B_t$, and so on with respect to higher-order beliefs. Thus, B_t is a sufficient statistic for higher-order beliefs. This allows for a tractable equilibrium construction.

We now provide intuition for the role of the key state variable B_t in traders' demand schedules. If trader j follows the equilibrium strategy, a substitution of (12) into (11) reveals that

$$X_{j,t-1} + \sum_{e \in E} q_{ie,t-1} = \frac{1}{N} B_t.$$

Summing across $j \neq i$,

$$\sum_{j \neq i} \left(X_{j,t-1} + \sum_{e \in E} q_{je,t-1} \right) = \frac{N-1}{N} B_t.$$

Thus for trader i, B_t is a sufficient statistic for the total post-trade inventory of other traders at date $t-1$. This in turn implies that B_t is sufficient information for trader i to conduct inference about the residual supply that he will face on each of the exchanges at time t. This explains the role of B_t in the demand schedule (11).

In a perfect Bayesian equilibrium, any given trader i, conjecturing that other traders submit demand functions according to (11), solves the stochastic control problem

$$\sup_{\{f_{iet}\}_{e,t}} \mathbb{E} \left[\sum_{t=0}^{\infty} e^{-r\Delta t} F_{it}(q_{it}) \;\middle|\; X_{i0} \right], \tag{14}$$

with demands that are measurable[9] with respect to the history of inventory levels $\{X_{is}\}_{s \leq t}$, trades $\{q_{ies}\}_{e \in E, s<t}$, and prices $\{p_{es}\}_{e \in E, s<t}$, and satisfy[10]

$$\lim_{t \to \infty} e^{-r\Delta t} \mathbb{E} \left(X_{it}^2 \right) = 0, \tag{15}$$

ruling out "Ponzi schemes" that are based on explosive growth in asset positions. An equilibrium is characterized by optimal demands determined by the same function $f_{iet}(\,\cdot\,)$ of (11).

In solving the optimization problem (14), trader i correctly considers the impacts of his or her trades on current and future prices. These impacts occur directly through the formation of the clearing price on the exchange on which an order is submitted and also through the recognition by trader i that other traders draw inference from market prices about the aggregate market supply of the asset, which affects future prices at all exchanges. This impact occurs through the "beliefs" state variable B_t, through the dynamic equation

[9] Although the objective function involves second moments of X_{it}, we allow strategies that do not have finite second moments and show that any such strategy is strictly suboptimal.

[10] This condition is implied by the square-integrability condition $\mathbb{E}\left[\sum_{t=0}^{\infty} e^{-r\Delta t} \sum_{e \in E} q_{iet}^2\right] < \infty$.

(12).

In Appendix E, we use the Bellman principle of optimality to explicitly solve for the required number E of exchanges and for the equilibrum demand coefficients ρ, ζ, and χ. We find that

$$E = 2 + \frac{2}{N-2} + \frac{N(N-1)}{N-2}\frac{\sigma_\epsilon^2}{\sigma_Q^2} \tag{16}$$

$$\rho = -\frac{1}{NE} \tag{17}$$

$$\zeta = \frac{N-2}{2b(N-1)(1+\Gamma(E-1))} - \frac{e^{-r\Delta}}{bE}, \tag{18}$$

$$\chi = \frac{\mu_\pi \Delta}{1-e^{-r\Delta}}\zeta, \tag{19}$$

where

$$\Gamma = \frac{(N-1)\sigma_\epsilon^2}{(N-1)\sigma_\epsilon^2 + E\sigma_Q^2}.$$

From the demand schedule (11), the within-period price impact on any exchange is

$$\frac{1}{\zeta(N-1)} = \frac{2b(1+\Gamma(E-1))}{N-2-e^{-r\Delta}\frac{2N-2}{E}(1+\Gamma(E-1))}, \tag{20}$$

which is higher than in the associated static model. We also compute that cross-period cross-exchange price impact is

$$\frac{dp_{e,t+1}}{dq_{jkt}} = -N\rho\frac{1}{(N-1)\zeta} = \frac{1}{E}\frac{2b(1+\Gamma(E-1))}{N-2-e^{-r\Delta}\frac{2N-2}{E}(1+\Gamma(E-1))}, \tag{21}$$

which is a fraction $1/E$ of the within-period within-exchange price impact. (The differential notation shown for this price sensitivity involves a minor abuse of notation.) The marginal impact of the quantity traded by any trader on any exchange on the sum of exchange prices in the next time period is equal to the within-period within-exchange price impact.

7.3 Summary of Results

The following theorem summarizes the results of our analysis of the dynamic model.

Theorem 3. *If*

$$E = 2 + \frac{2}{N-2} + \frac{N(N-1)}{N-2}\frac{\sigma_\epsilon^2}{\sigma_Q^2}$$

is an integer, then there exists a perfect Bayesian equilibrium in symmetric affine demand schedules for the dynamic market with E exchanges such that:

1. *Trade is allocatively efficient along the equilibrium path.*

2. *Traders submit the demand schedule given by* (11)*, with* ρ*,* ζ*, and* χ *given by* (17)*,* (18)*, and* (19) *respectively.*

3. *Beliefs about the aggregate market inventory evolve according to* (12)*.*

4. *Trades on each exchange have nonzero price impact at each exchange in the next period given by* (21)*.*

5. *The within-period within-exchange price impact* (20) *is higher than that for the associated static model.*

In the equilibrium, by deviating, traders can manipulate other traders' beliefs about the aggregate market asset inventory. Following a one-shot deviation, trade returns to efficiency in the next period, and beliefs become "corrected."

Our analysis shows that in our dynamic model, as for the associated static model, a precise and non-trivial amount of market fragmentation achieves allocative efficiency. Relative to the static model, our dynamic model allows a clearer characterization of how fragmentation improves price discovery. A weakness of our analysis of the dynamic setting is that, because of the need to incorporate the infinite regress of beliefs about beliefs, we are able to characterize equilibrium only for the number of exchanges that is associated with an efficient allocation. A notable implication of Theorem 3 is that this efficient number of exchanges is invariant to the frequency of trade and is identical to that of the static model. Though multiple periods of trade increases price impact relative to the static model, this additional price impact exacerbates the role of traders' asymmetric information, which, as we saw in the static model, leads to more aggressive order submission. This is so because traders rely on their privately observed inventories to conduct inference that is relevant not only to current-period trade prospects, but also to future-period trade prospects. This effect on trade aggressiveness precisely offsets the effects of the rise in price impact on each exchange.

8 Discussion of Model Extensions

In this section we summarize the results of three extensions of the main model that are provided in appendices.

8.1 Endogenous Liquidity Trade, Exchange by Exchange

In our first model extension, found in Appendix F, liquidity traders, who are local to each exchange and conduct no cross-exchange trade, choose the sizes of their trades. Liquidity traders are assumed to have the same preferences as strategic traders, but may have a different quadratic holding cost parameter, c, and may also be endowed with some quantity of the asset prior to trade. Thus, the baseline model is equivalent to the case in which $c = \infty$, in that liquidity traders liquidate their entire endowed positions as though without discretion. Relaxing this baseline extreme assumption to the case of finite c, we find for any positive integer $\overline{E} > 1$, there exists a cutoff $\overline{c}$ such that if $c > \overline{c}$, then a market with $1 < E \leq \overline{E}$ exchanges is welfare superior to a centralized market in that the expected sum of all agents' holding costs is lower.

8.2 Private information about asset payoff

In a second extension, found in Appendix G, agents have differing private information about the asset's final payoff. In this case, allocative efficiency is not necessarily improved by fragmenting a centralized market. This is because fragmentation leads agents to trade more aggressively for two reasons: not only to mitigate holding costs, but also to exploit payoff-relevant private information. While the former motive leads fragmentation to improve allocative efficiency, as we demonstrated in Section 5, the latter effect can cause fragmentation to reduce allocative efficiency. This is because the efficient allocation of the asset does not depend on agents' payoff-relevant private information. Whether fragmentation is beneficial or harmful is shown to depend on the relative magnitudes of these two effects.

8.3 Correlated trade motives

In a third extension, found in Appendix H, we relax the assumption that the underlying random variables $(X_1, \ldots, X_N, Q_1 \ldots, Q_E)$ are jointly independent. We retain the assumption that these random variables are jointly Gaussian, but allow for an essentially arbitrary covariance matrix, subject to the condition that the traders' endowments $X_1, \ldots, X_N$ are symmetrically distributed and that the liquidity-trade quantities $Q_1, \ldots, Q_E$ are symmetrically distributed.

If a strategic trader's endowment X_i does not covary more negatively with aggregate liquidity trader supply $\sum_e Q_e$ than it covaries positively with the aggregate endowment $\sum_j X_j$, there is an interior optimal level of fragmentation which, up to the integer constraint on E, achieves the efficient allocation.[11]

In this setting, however, an arbitrary level of market fragmentation need not, however, coincide with an unambiguous improvement in allocative efficiency over a centralized market. Whether this is so depends on the covariances of endowments. Under certain parameters, agents may trade even more aggressively than they do in the baseline model, which we have shown has the property that trade already becomes "too aggressive" for sufficiently large E. Moreover, if a strategic trader's endowment covaries more negatively with the aggregate liquidity trader supply than it covaries positively with the aggregate endowment, fragmentation is harmful. This is because the inefficiency associated with the inferior trading technology associated with disconnected fragmented markets dominates the beneficial effect of lowering the effect of strategic avoidance of price impact. This follows from the fact that, ex ante, with this correlation structure, traders expect that residual supply on each exchange is on average relatively favorable for offsetting their positions. This, however, leads to less aggressive trade than is socially efficient since agents are less willing to trade large quantities at unfavorable prices on any given exchange because they expect that prices on the other exchanges will be more favorable.

9 Concluding Discussion

We have presented a simple market setting in which fragmentation of trade across multiple exchanges improves allocative efficiency and price informativeness. Our main marginal contributions are (a) a newly identified channel by which cross-exchange price inference exacerbates price impact, and (b) a demonstration of the beneficial effects of cross-exchange order-splitting on allocative efficiency and price informativeness. We find that although fragmentation reduces market depth on any given exchange, this need not be a sign of worsening overall liquidity or market inefficiency. We characterize the number of exchanges that achieves allocative efficiency, and show that this "optimal" degree of fragmentation is invariant to the frequency of trade and indeed the same as that for the static version of the model.

Our stylized model abstracts from many important practical considerations. We do not consider some of the direct frictional costs of trade and order splitting, such as trading fees and subsidies, minimum tick sizes, and bid-offer spreads, which are endogenous to market structure, particularly through the role of competition among exchange operators, specialists, and market makers (Baldauf and Mollner, 2019; Chao, Yao, and Ye, 2018; Colliard and

[11] Positive definiteness of the covariance matrix ensures that, for each i, X_i is positively correlated with $\sum_{j \in N} X_j$.

Foucault, 2012; Malinova and Park, 2019; Foucault and Menkveld, 2008; Chlistalla and Lutat, 2011; Clapham et al., 2019; Hengelbrock and Theissen, 2009; Parlour and Seppi, 2003). For example, Foucault and Menkveld (2008) show that, with non-zero tick sizes, adding an additional limit-order market increases market depth by allowing limit-order submitters to jump the queue of posted orders on one exchange by posting orders on another exchange, due to the absence of cross-exchange time-priority rules. Foley, Jarnecic, and Liu (2020) show that liquidity providers increasingly fragment their activities amongst alternative venues, attempting to jump long queues on larger venues by increasing submissions to venues with short (or empty) queues. This reduces adverse selection costs faced on alternative venues and helps explain the increase in fragmentation for jurisdictions with trade-through prohibitions.

We also do not consider the endogenous entry of exchanges, a common theme in the literature going back to Glosten (1994), as reviewed by Pagnotta and Philippon (2018). Our model does not capture the effect of high-frequency traders who can take advantage of slight discrepancies in order execution times across different exchanges (Budish, Lee, and Shim, 2019; Gresse et al., 2012; Pagnotta and Philippon, 2018). We also ignore the role of trade-through rules such as Regulation NMS, which effectively forces all U.S. lit exchanges to recognize the best bid or offer available across all order books in the market. While Reg NMS has the effect of consolidating markets for small trades, trade-through rules do not play a significant role in price-impact costs, which are only pertinent for large trades. The inefficiencies associated with price-impact cost avoidance through order splitting are the main concern in this paper.

Because we have abstracted from these and other potentially important realistic effects, we make no normative claims or policy recommendations. The mechanisms that we identify do, however, appear to have a natural basis and to be worthy of serious consideration in policy discussions.

Our model also has implications for the welfare impact of innovation of trading technologies. For example, the beneficial welfare effects of order splitting that we have described rely crucially on the realistic assumption that orders submitted to one exchange cannot condition on prices at other exchanges. If, instead, trading technology were to improve so that orders could condition on cross-exchange prices, then trades on a given exchange would have impact on prices at other exchanges, which could eliminate the beneficial effect of order-splitting in fragmented markets, an issue considered by Wittwer (2020) and Rostek and Yoon (2020).

Appendix

A Verification Theorem

Here, we prove a theorem that will be used in later sections to verify a candidate symmetric affine equilibrium. The theorem applies without alteration to the models of Section 4, Section 6, and Appendix H. In what follows, $\mathcal{F}_i$ denotes trader i's information set. In the models of Section 4 and Appendix H, $\mathcal{F}_i = \sigma(X_i)$ while in the model of Section 6, $\mathcal{F}_i = \sigma(X_i, \sum_{j\in N} X_j)$. The set of admissible demand schedules, $\mathcal{M}_i$, is the set of all maps $h : \Omega \times \mathbb{R} \to \mathbb{R}$ that are $\mathcal{F}_i \times \mathcal{B}(\mathbb{R})$-measurable. We denote a candidate symmetric affine equilibrium by the associated triple of demand schedule coefficients, (α, ζ, Δ).

Theorem 4. *Let (α, ζ, Δ) be a candidate symmetric affine equilibrium such that $\zeta > 0$. For each $e \in E$ and $i \in N$ set*

$$r_e^i := \sum_{j\neq i} -\alpha X_j + (N-1)\Delta - Q_e.$$

For each $e \in E$, let f_{ie} be as in (2). A necessary and sufficient condition for (α, ζ, Δ) to be a symmetric affine equilibrium is that

$$\mu_\pi - \mathbb{E}\left[2b\left(X_i + \sum_{e\in E} f_{ie}(\omega, p_e^f)\right) \mid \mathcal{F}_i, r_e^i\right] = p_e^f + \Lambda f_{ie}(\omega, p_e^f) \tag{22}$$

holds almost surely for each $i \in N$ and $e \in E$.

Proof. We prove that trader i optimizes by submitting demand schedules of the form in (2) to each exchange if all other traders do the same. Suppose trader i submits $g_{ie} \in \mathcal{M}$ to exchange e. Then if a market clearing price exists, it satisfies

$$p_e(\omega) = \frac{r_e^i(\omega) + g_{ie}\left(\omega, p_e(\omega)\right)}{\zeta(N-1)}. \tag{23}$$

For any given demand schedule g_{ie} which conditions the quantity purchased on the realization of p_e there is a function $\tilde{g}_{ie}$ which conditions the quantity purchased on the realization of r_e^i such that

$$\tilde{g}_{ie}(\omega, r_e^i(\omega)) = g_{ie}\left(\omega, p_e(\omega)\right)$$

for each $\omega \in \Omega$ for which a unique clearing price exists and

$$\tilde{g}_{ie}(\omega, r_e^i(\omega)) = 0$$

for each $\omega \in \Omega$ such that there is no unique clearing price[12]. To see this, define $\tilde{g}_{ie}$ as follows. For each $r \in \mathbb{R}$ let $p(r,\omega)$ denote the unique solution to

$$p = \frac{r + g_{ie}(\omega, p)}{\zeta(N-1)}$$

if such a solution exists. For all r such that $p(r,\omega)$ is well defined set

$$\tilde{g}_{ie}(\omega, r) = g_{ie}\left(\omega, p(r,\omega)\right).$$

Otherwise, set

$$\tilde{g}_{ie}(\omega, r) = 0.$$

Given $\{f_{ie}\}_{e\in E}$ as in the statement of theorem, define $\{\tilde{f}_{ie}\}_{e\in E}$ in this way. Then

$$\tilde{f}_{ie}(\omega, r) = -\alpha\frac{N-1}{N}X_i - \frac{r}{N} + \frac{N-1}{N}\Delta$$

for each $e \in E$.

It is convenient to relax trader i's optimization problem to

$$\sup_{(\tilde{g}_{i1},...,\tilde{g}_{iE})\in\mathcal{M}_i^E} \mathbb{E}[\pi\sum_{e\in E}\tilde{g}_{ie}(\omega, r_e^i) - b\big(X_i + \sum_{e\in E}\tilde{g}_{ie}(\omega, r_e^i)\big)^2 - \sum_{e\in E}\frac{r_e^i + \tilde{g}_{ie}(\omega, r_e^i)}{\zeta(N-1)}\tilde{g}_{ie}(\omega, r_e^i)]. \quad (24)$$

Above, we have suppressed the dependence of r_e^i on ω. For some $(\tilde{g}_{i1}, ..., \tilde{g}_{iE}) \in \mathcal{M}^E$ the expectation may be infinite. As a result we will first restrict the domain to the set $\hat{\mathcal{M}}^E$ where $\hat{\mathcal{M}}$ is the subset of $h \in \mathcal{M}$ such that $h(\cdot, r_e^i(\cdot))$ is a finite variance random variable. Later we will argue that this is without loss of generality in that any profile of demand schedules outside of $\hat{\mathcal{M}}^E$ leads to $-\infty$ utility.

To derive a first order condition we take the variation of $\tilde{f}_{ie}$ with an arbitrary $h_e \in \hat{\mathcal{M}}$ for each $e \in E$ and substitute into the objective. This gives

$$\begin{aligned}\mathbb{E}\left[\pi\sum_{e\in E}\big(\tilde{f}_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)\big) - b\big(X_i + \sum_{e\in E}\tilde{f}_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)\big)^2\right]\\ - \mathbb{E}\left[\sum_{e\in E}\frac{r_e^i + \tilde{f}_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)}{\zeta(N-1)}\big(\tilde{f}_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)\big)\right] \quad (25)\end{aligned}$$

[12]Recall that if a unique market clearing price does not exist no trades are executed.

where ν is a constant in $\mathbb{R}$. Differentiating with respect to ν and evaluating at $\nu = 0$ gives:

$$\mathbb{E}\left[\pi \sum_{e\in E} h_e(\omega, r_e^i) - 2b\big(X_i + \sum_{e\in E} \tilde{f}_{ie}(\omega, r_e^i)\big) \sum_{e\in E} h_e(\omega, r_e^i)\right] \\ - \mathbb{E}\left[\sum_{e\in E}\big(\frac{\tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)} + \frac{r_e^i + \tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)}\big) h_e(\omega, r_e^i)\right] = 0. \quad (26)$$

It holds if

$$\mathbb{E}\left[-2b\big(X_i + \sum_{k\in E} \tilde{f}_{ik}(\omega, r_k^i)\big) \mid \mathcal{F}_i, r_e^i\right] = \frac{\tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)} + \frac{r_e^i + \tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)} - \mu_\pi. \quad (27)$$

for each $e \in E$. We now show that (27) is a sufficient condition for optimality within $\hat{\mathcal{M}}^E$. Differentiating (25) with respect to ν twice we derive

$$\mathbb{E}\left[-2b\big(\sum_{e\in E} h_{ie}(\omega, r_e)\big)^2 - \frac{2}{\zeta(N-1)} \sum_{e\in E} h_{ie}(\omega, r_e)^2\right], \quad (28)$$

which is less than or equal to 0 for all $(h_1, ..., h_E) \in \hat{\mathcal{M}}^E$. The derivative is negative if one of $h_1, ..., h_N$ is nonzero on a set of positive measure. Suppose for contradiction that $(\tilde{f}_{i1}, ..., \tilde{f}_{iE})$ satisfies (27) but there exists $(h_1^*, ..., h_E^*) \in \hat{\mathcal{M}}^E$ which achieves a strictly higher value of (24). Set $(h_1, ..., h_E) \equiv (h_1^* - \tilde{f}_{i1}, ..., h_E^* - \tilde{f}_{iE}) \in \hat{\mathcal{M}}^E$. Then the function (25) achieves a higher value at $\nu = 1$ than at $\nu = 0$. However (25) is a strictly concave function of ν and thus has a global maximum at $\nu = 0$. This is a contradiction.

To show that it is without loss of generality to restrict attention to optimality within $\hat{\mathcal{M}}^E$ we observe that the coefficient of $\tilde{g}_{ie}(\omega, r_e^i)^2$ in (24) is negative. It is easy to see then that any $(\tilde{g}_{ie}, ..., \tilde{g}_{iE}) \notin \hat{\mathcal{M}}^E$ must result in $-\infty$ for the objective. This can be shown formally using the same method as in step 5 of the proof of Theorem 3.

Using (23) with (27) we see that (27) equivalent to (22) which is therefore a sufficient condition for (α, ζ, Δ) to be a symmetric affine equilibrium. We now show that it is also a necessary condition. Suppose for some $e \in E$, (22) does not hold and set

$$h_e(\omega, r_e^i) = \mu_\pi + \mathbb{E}\left[-2b\big(X_i + \sum_{k\in E} \tilde{f}_{ik}(\omega, r_k^i)\big) \mid \mathcal{F}_i, r_e^i\right] - \frac{\tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)} - \frac{r_e^i + \tilde{f}_{ie}(\omega, r_e^i)}{\zeta(N-1)}.$$

Note that h_e is an affine function of r_e^i. This is because the expectation is affine in r_e^i in each of the models of Section 4, Section 6, and Appendix H. Set $h_k(\omega, r_k^i) = 0$ for $k \neq e$. Then (26) is strictly positive. Thus for all ν sufficiently small $(\tilde{f}_{i1}, ..., \tilde{f}_{ie} + \nu h_e, ..., \tilde{f}_{iE})$ achieves a higher value of the objective (24) than does $(\tilde{f}_{i1}, ..., \tilde{f}_{iE})$. Define the demand schedule d_e such that for any given $p \in \mathbb{R}$ and $\omega \in \Omega$

$$d_e(\omega, p) = (\tilde{f}_{ie} + h_e)(\omega, r(\omega, p))$$

where $r(\omega, p)$ is defined to be the r that solves

$$p = \frac{r + (\tilde{f}_{ie} + \nu h_e)(\omega, r)}{\zeta(N-1)}.$$

If ν was chosen sufficiently small, $r(\omega, p)$ is well defined since the right hand side is an affine function of r with nonzero slope and so $d_e(\omega, p)$ is also well defined. Moreover

$$d_e(\omega, p_e(\omega)) = (\tilde{f}_{ie} + h_e)(\omega, r_e^i(\omega))$$

for each $\omega \in \Omega$. But then $(f_{i1}, ..., d_e, ..., f_{iE})$ gives higher expected utility to trader i than does $(f_{i1}, ..., f_{iE})$ which is a contradiction. Thus (22) is also a necessary condition. $\square$

B Proofs for Section 4

Here, we provide proofs for all results in Section 4 as well as present additional results which were not included in the main text. We first prove Lemma 5 which states that an equilbrium is "more efficient" the closer is $E\alpha_E$ to 1. Lemma 5 will be used in the proof of Theorem 1.

Lemma 5. Let (α, ζ, Δ) be a symmetric affine equilibrium when there are E exchanges and $(\hat{\alpha}, \hat{\zeta}, \hat{\Delta})$ be a symmetric affine equilibrium when there are $\hat{E}$ exchanges. For each $\omega \in \Omega$, the sum of strategic traders' holding costs post trade is strictly lower in the equilibrium corresponding to (α, ζ, Δ) if $|1 - E\alpha| < |1 - \hat{E}\hat{\alpha}|$. If the sum of strategic traders' holding costs post trade are equal across the two equilibria, then $|1 - E\alpha| = |1 - \hat{E}\hat{\alpha}|$

Proof. The sum of holding costs in the equilibrium (α, ζ, Δ) is

$$b \sum_{i \in N} \left((1 - E\alpha) X_i + E\alpha \frac{1}{N} \sum_{j \in N} X_j + \frac{\sum_{e \in E} Q_e}{N} \right)^2.$$

Expanding, rearranging, and combining like terms we obtain

$$b \sum_{i \in N} \left[(1 - E\alpha)^2 X_i^2 - [(1 - E\alpha)^2 - 1] \frac{1}{N} \left(\sum_{j \in N} X_j \right)^2 + \frac{(\sum_{e \in E} Q_e)^2}{N} + 2 \frac{\sum_{e \in E} Q_e}{N} \sum_{j \in N} X_j \right].$$

The result is an implication of the above expression and Jensen's inequality. □

B.1 Proof of Theorem 1

We prove Theorem 1 in 3 steps. In step 1, we derive a system of equations and show that a necessary and sufficient condition for (α_E,ζ_E, Δ_E) to be a symmetric affine equilibrium is that they solve this system. In step 2 we prove that there is exists a unique solution to the system, thus establishing existence and uniqueness of a symmetric affine equilibrium. This proves the preamble in Theorem 1. In step 3, we prove parts 1 through 6.

Step 1. Conjecture that there exists a symmetric affine equilibrium $(\alpha_E, \zeta_E, \Delta_E)$. Under this conjecture, each agent $i \in N$ submits a demand schedule of the form in (2) to each $e \in E$ and $i \in N$. Market clearing in exchange e implies that the equilibrium price is

$$p_e^f = \frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_e}{\zeta_E N}. \tag{29}$$

Price impact can also be determined from the market clearing condition. If trader i purchases q units on exchange e when all other traders submit the equilibrium demand schedules then

$$-\alpha_E \left(\sum_{\{j \in N \,|\, j \neq i\}} X_j \right) - \zeta_E(N-1)p_e + \Delta_E(N-1) + q = Q_e.$$

This implies that the inverse residual supply curve trader i faces is

$$p_e(q) = \frac{-\alpha_E \sum_{\{j \in N | j \neq i\}} X_i + q + \Delta_E(N-1) - Q_e}{\zeta_E(N-1)}. \tag{30}$$

Thus the price impact trader i faces in exchange e is $\Lambda := \frac{1}{\zeta_E(N-1)}$, which by symmetry, is the price impact each agent faces in all exchanges. Define $f_{ie}(X_i, p_e^f) := q_{ie}^f$. By Theorem 4 a necessary and sufficient condition for $(\alpha_E, \zeta_E, \Delta_E)$ to be a symmetric affine equilibrium is that

$$-2b\left(X_i + q_{ie}^f + (E-1)\mathbb{E}\left[q_{ik}^f \,|\, p_e^f - \frac{q_{ie}^f}{\zeta_E(N-1)}, X_i \right] \right) = p_e^f + \Lambda_E q_{ie}^f - \mu_\pi \tag{31}$$

holds almost surely for each $e \in E$ and trader $i \in N$. In (31), we have used symmetry of the exchanges. By the projection theorem,

$$\mathbb{E}\left[q_{ik}^f \,|\, p_e^f - \frac{q_{ie}^f}{\zeta_E(N-1)}, X_i\right] = -\alpha_E X_i \frac{N-1}{N} - \left(1 - \frac{N-1}{N}\gamma_E\right)\Delta_E$$
$$- \frac{N-1}{N}\gamma_E \zeta_E p_e^f + \gamma_E \frac{q_{ie}^f}{N} + \Delta_E, \quad (32)$$

where

$$\gamma_E = \mathrm{corr}_{X_i}(p_e^f, p_k^f) = \frac{E\alpha_E^2(N-1)\sigma_X^2}{E\alpha_E^2(N-1)\sigma_X^2 + \sigma_Q^2}. \quad (33)$$

Substituting (32) and (2) into (31) and matching coefficients we derive a system of three equations which characterize the three unknowns, α_E, ζ_E, and Δ_E. These equations are

$$\zeta_E = \frac{1}{2b((E-1)\gamma_E + 1)} \frac{N-2}{N-1} \quad (34)$$

$$\alpha_E = \frac{1}{1 + \frac{\gamma_E(E-1)}{N} + \frac{(E-1)\gamma_E + 1}{N-2} + (E-1)\frac{N-1}{N}}, \quad (35)$$

and

$$\Delta_E = \frac{\mu_\pi}{2b\left(1 + \frac{\gamma_E(E-1)}{N} + \frac{(\gamma_E(E-1)+1)}{N-2} + (E-1)\frac{N-1}{N}\gamma_E\right)}. \quad (36)$$

Equations (34), (35), and (36) are necessary and sufficient conditions for $(\alpha_E, \zeta_E, \Delta_E)$ to be a symmetric affine equilibrium.

Step 2. We now prove existence of a symmetric affine equilibrium. It is straightforward to substitute (33) into (35) and derive a cubic equation that characterizes α_E. Since the equation is cubic, there exists at least one real root. Take this to be the value of α_E and compute ζ_E and Δ_E using equations (33), (34), and (36). Thus a symmetric affine equilibrium exists.

To prove uniqueness, fix $E \in \mathbb{N}$ and define the function g as follows

$$g(a) := a - \frac{1}{E\gamma(\frac{1}{N} + \frac{1}{N-2}) + (1-\gamma)(\frac{1}{N} + \frac{1}{N-2}) + E\frac{N-1}{N}}$$

where γ is a function of a such that $\gamma(a)$ is equal to (33) but with a in place of α_E. Since we have already shown existence there is an $a \in \mathbb{R}$ such that $g(a) = 0$. We observe that the second term in the above expression is strictly monotone decreasing in γ. By (33) we see that γ is strictly monotone increasing in a. Thus $g(a)$ is strictly monotone increasing in a. Hence there can exist at most one value of $a \in \mathbb{R}$ such that $g(a) = 0$.

Step 3. Part 1 follows immediately from (29) and the fact that $\Lambda_E = \frac{1}{(N-1)\zeta_E}$. Part 2 follows immediately from (34). Part 3 follows by substituting (29) into (2). Part 4 can be seen from the fact that when $\sigma_Q^2 = 0$, γ_E is equal to 1 so that inspecting equations (34), (35), and (36) we have closed form solutions for ζ_E, α_E, and Δ_E. Using these closed form solutions we find that $E\alpha_E$, by (35), is equal to $\frac{N-1}{N-2}$ which is independent of E and also equal to $\frac{2b}{2b+\Lambda_1}$. To prove part 5, we combine part 3 with part 4.

Finally, we prove part 6. By Proposition 1, $\gamma \to 0$—note that the proof of Proposition 1 does not rely on Part 6 of Theorem 1 so the logic is not circular. Using (35) with some rearrangement we write

$$E\alpha_E = \frac{1}{\gamma_E(\frac{1}{N} + \frac{1}{N-2}) + (1-\gamma_E)\frac{1}{E}(\frac{1}{N} + \frac{1}{N-2}) + \frac{N-1}{N}}. \tag{37}$$

Taking limits, $E\alpha_E \to \frac{N}{N-1}$. To prove that $E\alpha_E$ is strictly monotone increasing in E, suppose for contradiction that there exists $E \in \mathbb{N}$ such that $(E+1)\alpha_{E+1} < E\alpha_E$. Then by inspection it must be that $\gamma_{E+1} > \gamma_E$. But, inspecting (33), this implies that $(E+1)\alpha_{E+1}^2 > E\alpha_E^2$ which in turn implies that $(E+1)\alpha_{E+1} > E\alpha_E$, a contradiction.

When E is equal to 1, $E\alpha_E$ is equal to $\frac{N-2}{N-1}$ by part 5. When $E \to \infty$, $E\alpha_E$ converges strictly monotonically to $\frac{N}{N-1}$. Thus for any $E > 1$ we have

$$\frac{1}{N-1} = |\,1-\alpha_1\,| > |\,1 - E\alpha_E\,|.$$

That a fragmented market is always more efficient than a centralized market follows from Lemma 5.

C Proofs for Section 5

C.1 Proof of Proposition 1

That Λ_E is strictly monotone increasing and diverges to ∞ when $\sigma_Q^2 = 0$ is immediate from Theorem 1, where we showed that, in this case, $\Lambda_E = \frac{2bE}{N-2}$. For what follows assume $\sigma_Q^2 > 0$.

By Theorem 1 we have $\Lambda_E = \frac{2b(1+\gamma_E(E-1))}{N-2}$. To show Λ_E is strictly monotone increasing it suffices to show that $(E-1)\gamma_E$ is strictly monotone increasing. Fix an arbitrary $E \in \mathbb{N}$. If $\gamma_{E+1} > \gamma_E$, then it must be that $E\gamma_{E+1} > (E-1)\gamma_E$. Suppose $\gamma_{E+1} \leq \gamma_E$. Then to prove that $E\gamma_{E+1} > (E-1)\gamma_E$ it suffices to prove that $(E+1)\gamma_{E+1} > E\gamma_E$. Consider the equation for γ_n derived in the proof of Theorem 1 which holds for arbitrary $n \in \mathbb{N}$:

$$\frac{n\alpha_n^2(N-1)\sigma_X^2}{n\alpha_n^2(N-1)\sigma_X^2 + \sigma_Q^2}.$$

Denote the numerator, num_n so that

$$\gamma_n = \frac{num_n}{num_n + \sigma_Q^2}.$$

By Theorem 1, $(E+1)\alpha_{E+1} > E\alpha_E$ which implies that

$$(E+1)\gamma_{E+1} = \frac{(E+1)num_{E+1}}{num_{E+1} + \sigma_Q^2} > \frac{Enum_E}{num_E + \sigma_Q^2} = E\gamma_E.$$

We next prove that Λ_E converges and give an explicit expression for the limit point. We can, using the expression for γ_E, write Λ_E as

$$\frac{2b}{N-2}\left(1 + \frac{E^2\alpha_E^2(N-1)\sigma_X^2 - E\alpha_E^2(N-1)\sigma_X^2}{E\alpha_E^2(N-1)\sigma_X^2 + \sigma_Q^2}\right).$$

By Theorem 1, $E\alpha_E \to \frac{N}{N-1}$ which implies that $E\alpha_E^2 \to 0$. Taking limits of the right-hand side of the above equation we obtain $\Lambda_E \to \frac{2b}{N-2}(1 + \frac{N^2\sigma_X^2}{(N-1)\sigma_Q^2})$.

To prove that $\gamma_E \to 0$ we inspect (35) to see that

$$\frac{1}{E(\frac{N-1}{N} + \frac{1}{N} + \frac{1}{N-2})} < \alpha_E < \frac{1}{E\frac{N-1}{N}}$$

for all E sufficiently large. Inspecting (33), we see that for large E, the numerator is $O(\frac{1}{E})$ since by Theorem 1 $E\alpha_E$ converges. The denominator is roughly equal to σ_Q^2 for large E so it must be that $\gamma_E \to 0$.

Finally, we prove that γ_E is strictly monotone decreasing in E. Using (37) and substituting into (33) we derive a cubic equation which characterizes γ:

$$\gamma^3 E\sigma_Q^2(1-\frac{1}{E})^2(\frac{1}{N} + \frac{1}{N-2})^2 + \gamma^2 E\sigma_Q^2\left(\frac{N-1}{N} + \frac{1}{E}(\frac{1}{N} + \frac{1}{N-2})\right)^2$$
$$+ 2\gamma^2\sigma_Q^2(1-\frac{1}{E})(\frac{1}{N} + \frac{1}{N-2})(\frac{N-1}{N}E + \frac{1}{N} + \frac{1}{N-2}) + \gamma\sigma_X^2(N-1) = \sigma_X^2(N-1). \quad (38)$$

Each of the coefficients are unambiguously increasing in E except for possibly

$$E\sigma_Q^2\big(\frac{N-1}{N}+\frac{1}{E}(\frac{1}{N}+\frac{1}{N-2})\big)^2.$$

Taking a derivative with respect to E we have

$$\sigma_Q^2\big(\frac{N-1}{N}+\frac{1}{E}(\frac{1}{N}+\frac{1}{N-2})\big)^2-\frac{2}{E}\sigma_Q^2\big(\frac{N-1}{N}+\frac{1}{E}(\frac{1}{N}+\frac{1}{N-2})\big).$$

This derivative is nonegative if

$$E\frac{N-1}{N}+\frac{1}{N}+\frac{1}{N-2}\geq 2.$$

The above holds for $E\geq 2$ since

$$2\frac{N-1}{N}+\frac{1}{N}+\frac{1}{N-2}\geq 2.$$

Since each of the coefficients of the powers of γ in (38) are increasing in E and some are strictly increasing it must be that γ is strictly decreasing in E since the right hand side of (38) is constant.

C.2 Proof of Proposition 2

Substituting (33) into (35) and rearranging we obtain the following cubic equation which defines $E\alpha_E$:

$$(E\alpha_E)^3(\sigma_X^2(N-1)(1+\frac{1}{N-2}))-(E\alpha_E)^2(N-1)\sigma_X^2+E\alpha_E\sigma_Q^2(E-\frac{E}{N}+\frac{1}{N}+\frac{1}{N-2})-E\sigma_Q^2=0.$$

The efficient allocation is acheived at E^* such that $E^*\alpha_{E^*}=1$ provided E^* is in $\mathbb{N}$. Thus

$$(\sigma_X^2(N-1)(1+\frac{1}{N-2}))-(N-1)\sigma_X^2+\sigma_Q^2(E^*-\frac{E^*}{N}+\frac{1}{N}+\frac{1}{N-2})-E^*\sigma_Q^2=0.$$

Solving for E^* yields

$$E^*=2+\frac{2}{N-2}+\frac{N-1}{N-2}\frac{N\sigma_X^2}{\sigma_Q^2}.$$

That the $E\in\mathbb{N}$ whose symmetric affine equilibrium allocation is most efficient is either $\lfloor E^*\rfloor$ or $\lceil E^*\rceil$ follows from the proof of part 6 of Theorem 1 which shows that $E\alpha_E$ is strictly monotone increasing. By inspection, the proof did not rely upon E taking values in $\mathbb{N}$—the same method of proof can be adapted to show that if we increase E continuously, the corresponding α_E which simultaneously solves (33) and (35) is such that $E\alpha_E$ is strictly monotone increasing. Combining this observation with Lemma 5 gives the result.

C.3 Proof of Proposition 3

We first prove part 1. Recall that

$$p_e^* = \frac{N-1}{N}\Lambda_E[\sum_{i\in N} X_i + N\Delta_E - Q_e].$$

By the projection theorem

$$\mathrm{var}[\sum_{i\in N} X_i \,|\, p_e^*] = (1 - \frac{\alpha_E^2 \mathrm{var}[\sum_{i\in N} X_i]}{\alpha_E^2 \mathrm{var}[\sum_{i\in N} X_i] + \frac{\sigma_Q^2}{E}})\mathrm{var}[\sum_{i\in N} X_i].$$

By an argument analogous to the one used to show that γ_E is strictly monotone decreasing to 0 given in Proposition 1, we can show that $\frac{\alpha_E^2 \mathrm{var}[\sum_{i\in N} X_i]}{\alpha_E^2 \mathrm{var}[\sum_{i\in N} X_i] + \frac{\sigma_Q^2}{E}}$ converges to 0 strictly monotonically as E diverges.

We now prove part 2. Since the price in each exchange consists of a common signal component and noise which is i.i.d across exchanges, the sum of prices is a sufficient statistic for inference so that $\mathrm{var}[\sum_{i\in N} X_i \,|\, \sum_{e\in E} p_e^*] = \mathrm{var}[\sum_{i\in N} X_i \,|\, p_1^*, ..., p_E^*]$. We have

$$\sum_{e\in E} p_e^* = \frac{N-1}{N}\Lambda_E[-E\alpha_E \sum_{i\in N} X_i - Q + EN\Delta_E].$$

By the projection theorem,

$$\mathrm{var}[\sum_{i\in N} X_i \,|\, \sum_{e\in E} p_e^*] = \mathrm{var}[\sum_{i\in N} X_i] - \frac{(E\alpha_E)^2 \mathrm{var}[\sum_{i\in N} X_i]}{(E\alpha_E)^2 \mathrm{var}[\sum_{i\in N} X_i] + \sigma_Q^2}\mathrm{var}[\sum_{i\in N} X_i]$$

The result follows since $\frac{(E\alpha_E)^2 \mathrm{var}[\sum_{i\in N} X_i]}{(E\alpha_E)^2 \mathrm{var}[\sum_{i\in N} X_i] + \sigma_Q^2}$ increases strictly monotonically because $E\alpha_E$ increases strictly monotonically as seen from part 6 of Theorem 1.

Proposition 4. *The expected payment of liquidity traders is* $\frac{N-1}{N}\Lambda_E \sigma_Q^2$ *and if* $\sigma_Q^2 > 0$ *is strictly monotone increasing in* E.

Proof.

$$-\mathbb{E}[\sum_{e\in E} p_e^* Q_e] = -\frac{N-1}{N}\Lambda_E \mathbb{E}[(-\sum_{e\in E}(\alpha_E \sum_{i\in N} X_i + N\Delta_E + Q_e)Q_e] = \frac{N-1}{N}\Lambda_E \sigma_Q^2.$$

That the expected payment is strictly monotone increasing follows from the fact that Λ_E is strictly monotone increasing as stated in Proposition 1. $\square$

D Proofs for Section 6

D.1 Proof of Theorem 2

We prove Theorem 2 in 3 steps. In the step 1 we derive a candidate equilibrium. In step 2 we verify that the candidate equilibrium is in fact an equilibrium, and then establish that it is the unique symmetric affine equilibrium if for each $e \in E$, Q_e has full support over the real line. In step 3 we show that the derived equilibrium has properties 1 through 5 given in the statement of the theorem.

Step 1. To begin the first step, we conjecture that there exists a symmetric affine equilibrium, denoted (α, ζ, Δ) in which each trader submits demand schedules of the form in (2). Define q^f_{ie} by $q^f_{ie} := f_{ie}(X_i, p^f_e)$. Under this conjecture, by market clearing, the residual supply curve trader i faces in exchange e is

$$p_e(q) = \frac{(\sum_{j \neq i \in N} -\alpha X_j) + (N-1)\Delta + q - Q_e}{(N-1)\zeta}$$

which implies that $\Lambda = \frac{1}{(N-1)\zeta}$. Also by market clearing we have

$$p^f_e = \frac{(\sum_{j \in N} -\alpha X_j) + N\Delta - Q_e}{N\zeta}. \tag{39}$$

for each $e \in E$. By observing p^f_e trader i can infer the realization of Q_e but this is uninformative of p^f_k for $k \neq e$. Define q^f_{ie} by $q^f_{ie} := f_{ie}(X_i, p^f_e)$. By Theorem 4 a necessary and sufficient condition for (α, ζ, Δ) to be a symmetric affine equilibrium is that

$$-2b(X_i + q^f_{ie} + (E-1)\mathbb{E}[q^f_{ik} \mid \mathcal{F}_i]) = p^f_e + q^f_{ie}\Lambda - \mu_\pi \tag{40}$$

where we have used symmetry. Rearranging, we have

$$q^f_{ie} = \frac{-2bX_i - 2b(E-1)\mathbb{E}[q^f_{ik} \mid \mathcal{F}_i] - p^f_e + \mu_\pi}{\Lambda + 2b}.$$

Substituting p_k^f into the conjectured equilibrium demand schedule, we have

$$q_{ik}^f = -\alpha X_i + \frac{(\sum_{j\in N} \alpha X_j) + Q_k}{N}$$

so that

$$\mathbb{E}[q_{ik}^f \,|\, \mathcal{F}_i] = -\alpha X_i + \frac{(\sum_{j\in N} \alpha X_j)}{N}.$$

We therefore have

$$q_{ie}^f = \frac{(-2b + 2b(E-1)\alpha)X_i - 2b(E-1)\frac{(\sum_{j\in N} \alpha X_j)}{N} - p_e^f + \mu_\pi}{\frac{1}{(N-1)\zeta} + 2b}.$$

We now match coefficients with our conjecture that $q_{ie}^f = -\alpha X_i - \zeta p_e^f + \Delta$ to determine that

$$\zeta = \frac{N-2}{N-1}\frac{1}{2b}, \tag{41}$$

$$\Lambda = \frac{2b}{N-2}, \tag{42}$$

$$\alpha = \frac{2b}{\Lambda + 2bE}, \tag{43}$$

and

$$\Delta = \frac{-2b(E-1)\frac{2b}{\Lambda+2bE}\frac{\sum_{i\in N} X_i}{N} + \mu_\pi}{2b + \Lambda}. \tag{44}$$

Step 2. To complete step two we appeal to Theorem 4 which can be applied since (40) holds. To see that the symmetric affine equilibrium is unique when each Q_e has full support over the real line suppose that there exists a symmetric affine equilibrium such that at least one of the equations (41), (43), and (44) are not satisfied. Then equation (40) is violated for some realization of the price in some exchange $e \in E$ for some agent $i \in N$. Continuity implies that (40) must be violated for realizations of p_e^f in an open neighborhood of positive Lebesgue measure. Since each Q_e has full support over the real line and is independent of $\mathcal{F}_i$ (40) is violated on a set of positive $\mathbb{P}$-measure. This contradicts Theorem 4.

Step 3. Part 1 was shown in equation (42). Part 2. follows from substituting equations (41), (43), and (44) into (39). Part 3 follows from substituting the equation for price in part

2 in to the equilibrium demand schedule. To prove part 4, we have

$$-\mathbb{E}[\sum_{e\in E} p_e^* Q_e] = -\frac{N-1}{N}\frac{2b}{N-2}\mathbb{E}[\sum_{e\in E}(\sum_{i\in N} -\alpha X_i + N\Delta - Q_e)Q_e] = \frac{2b(N-1)}{N(N-2)}\text{var}[\sum_{e\in E} Q_e].$$

Part 5 follows from part 3 and taking the limit as E tends to infinity.

E Proofs for Section 7

This appendix provides a proof of Theorem 3, characterizing an efficient equilibrium for the dynamic version of the model.

E.1 Proof of Theorem 3

The proof proceeds in six steps. In Step 1 we derive the Bellman equation for the dynamic programming problem of trader i. In Step 2 we conjecture a continuation value function V as a solution to the Bellman and we derive a first order condition characterizing the optimal demand schedules of trader i in a restricted domain of demand schedules. In Step 3, we use the first order condition to compute the necessary number E of exchanges and the demand-schedule coefficients ρ, ζ, and χ. In Step 4 we relax the domain restriction on demand schedules. In Step 5, we verify a transversality condition on the value function. In Step 6 we verify that the strategy of submitting demand schedules with coefficients derived in Step 3 from the Bellman equation is in fact optimal.

Step 1.

For a given date t, let $H_t := (\{q_{ie}\}_{e\in E,s<t}, \{p_e\}_{e\in E,s<t}, \{X_{is}\}_{s\le t})$ denote the history of past quantities purchased by trader i, prices on each of the exchanges, and inventory levels. An admissible demand schedule submitted to an exchange e is a function f specifying the quantity $f(H_t, p)$ that trader i will purchase for any given realization $p \in \mathbb{R}$ of the price in the exchange following the history H_t. By inspecting (13) and following a similar argument to that given in the proof of Theorem 4, we see that for any such demand function f there exists a corresponding function $\hat{f}$ that instead specifies the quantity purchased by trader i as a function of H_t and

$$W_{et} := \frac{1}{E}\sum_{j\in N} X_{jt} + Q_{et}.$$

For instance, in the conjectured equilibrium, on exchange e, trader i makes the socially efficient purchase

$$\hat{f}_{iet}\left(H_t, W_{et}\right) = -\frac{1}{E} X_{it} + \frac{W_{et}}{N}, \tag{45}$$

as can be seen by substituting (13) into (11).

We first relax the dynamic programming problem by allowing trader i to select demand functions of the type $\hat{f}$. Let κ_{et} denote (X_{it}, B_t, W_{et}). Under the relaxation, the Bellman equation characterizing trader i's continuation value function $V(\,\cdot\,)$ is

$$V(X_{it}, B_t) = \sup_{\{g_{i1t}, \cdots, g_{iEt}\}} \mathbb{E}_{it}\left[u_{it} + e^{-r\Delta} V(X_{i,t+1}, B_{t+1})\right], \tag{46}$$

where

$$u_{it} = \mu_\pi \Delta \left(X_{it} + \sum_{e \in E} g_{iet}(\kappa_{et}) \right) - b \left(X_{it} + \sum_{e \in E} g_{iet}(\kappa_{et}) \right)^2 - \sum_{e \in E} p_{et} g_{iet}(\kappa_{et}).$$

Above, each $g_{iet} : \mathbb{R}^3 \to \mathbb{R}$ is an arbitrary measurable function. We will assume for now that each g_{iet} is such that $g_{iet}(\kappa_{et})$ is of finite variance conditional on X_{it} and B_t. Call the set of all such measurable functions with this property $\tilde{\mathcal{M}}$. We will show in step 4 that the finite variance assumption is without loss of generality. Note that $\hat{f}_{iet}$ is in $\tilde{\mathcal{M}}$. The operator $\mathbb{E}_{it}$ is the conditional expectation given the state variables, X_{it} and B_t. These are the relevant state variables because, at date t, trader i infers that $\frac{N-1}{N} B_t$ is the total inventory held by the other traders following trade at date $t-1$. Thus X_{it} and B_t are sufficient statistics for trader i to conduct inference on the residual supply curves on each exchange at each future trading date. The law of motion for (X_{it}, B_t) is given by (9) and (12).

A standard verification argument implies that if V satisfies the Bellman equation, and for every feasible strategy, the transversality condition

$$\lim_{t \to \infty} e^{-r\Delta t} \mathbb{E}_{i0}\left[V(X_{it}, B_t)\right] = 0, \tag{47}$$

then V is indeed the value function and the strategy achieving the supremum in (46) determines the optimal policy.

Step 2. We conjecture the value function V defined by

$$V(X_{it}, B_t) = \sum_{s=t}^{\infty} e^{-r\Delta(s-t)} M_s, \tag{48}$$

where

$$M_s = \mathbb{E}_{it}\left[\mu_\pi \Delta\left(X^{\hat{f}}_{is} + \sum_{e\in E} q^{\hat{f}}_{ies}\right) - b\left(X^{\hat{f}}_{is} + \sum_{e\in E} q^{\hat{f}}_{ies}\right)^2 - \sum_{e\in E} p^{\hat{f}}_{es} q^{\hat{f}}_{ies}\right].$$

and where the superscript $\hat{f}$ implies that the inventories, quantities, and prices are those induced by the conjectured equilibrium strategy in which any given trader i selects (45) for any given exchange e. Substituting (48) into the right hand side of the Bellman and using the law of iterated expectations, we can write the objective function in the Bellman equation as

$$\mathbb{E}_{it}\left[\sum_{s=t}^{\infty} e^{-r\Delta(s-t)}\left(\mu_\pi \Delta\left(X^{g}_{is} + \sum_{e\in E} q^{g}_{ies}\right) - b\left(X^{g}_{is} + \sum_{e\in E} q^{g}_{ies}\right)^2 - \sum_{e\in E} p^{g}_{es} q^{g}_{ies}\right)\right],$$

where the superscript g indicates that inventories, quantities, and prices are those generated by a strategy that selects at date t demands according to the functions $g_{i1t}, ..., g_{iEt}$, and then reverts back to the conjectured equilibrium strategy at date $t+1$. We now derive the E, ρ, ζ, and χ such that the optimal choice of $g_{i1t}, ..., g_{iEt}$ coincides with (45), thus verifying the conjecture (48).

To simplify the objective further, we recognize that for any choice of the deviating demands $g_{i1t}, ..., g_{iEt}$, following trade at date $t+1$, the inventory of trader i returns to the efficient level, so all inventories, prices, and quantities at dates $s > t+1$ would be the same as if trader i had never deviated and therefore do not depend on the chosen $g_{i1t}, ..., g_{iEt}$. Thus, it suffices to consider the objective

$$\mathbb{E}_{it}\left[\mu_\pi \Delta \sum_{e\in E} q^{g}_{iet} - b\left(X^{g}_{it} + \sum_{e\in E} q^{g}_{iet}\right)^2 - \sum_{e\in E} p^{g}_{et} q^{g}_{iet} - e^{-r\Delta} \sum_{e\in E} p^{g}_{e,t+1} q^{g}_{ie,t+1}\right]. \tag{49}$$

Let $\eta_{et} \equiv -\frac{1}{E}\sum_{j\neq i} X_{jt} - Q_{et}$. Then

$$\begin{aligned}
\sum_{e\in E} p^{l}_{et} q^{g}_{iet} &= \sum_{e\in E} \frac{\eta_{et} + (N-1)\rho B_t + (N-1)\chi + q^{g}_{iet}}{\zeta(N-1)} q^{g}_{iet} \\
&= \frac{1}{\zeta(N-1)}\left[\sum_{e\in E}\big(\eta_{et} + (N-1)\rho B_t + (N-1)\chi\big) q^{g}_{iet} + \sum_{e\in E} (q^{g}_{iet})^2\right]. \qquad (50)
\end{aligned}$$

From (9), (11), and (13),

$$q^g_{ie,t+1} = -\frac{1}{E}\left(X_{it} + \sum_{e\in E} q^g_{iet} + \epsilon_{i,t+1}\right) + \frac{1}{NE}\sum_{j\in N} X^g_{j,t+1} + \frac{Q_{e,t+1}}{N}$$

and

$$p^g_{e,t+1} = \frac{-\frac{1}{E}\sum_{j\in N} X^g_{j,t+1} + N\rho\left(NE\rho B_t + EN\chi - \zeta N\sum_{e\in E} p^g_{et}\right) - Q_{e,t+1} + N\chi}{\zeta N}.$$

From the above two equations,

$$\begin{aligned} p^g_{e,t+1} q^g_{ie,t+1} = &\left(\frac{1}{\zeta N}\frac{1}{E^2}\sum_{j\in N} X^g_{j,t+1} - \frac{N\rho^2}{\zeta}B_t - \frac{N\rho\chi}{\zeta} - \frac{1}{E}\frac{\chi}{\zeta}\right)\sum_{e\in E} q^g_{ie,t} \\ &- N\rho\left(-\frac{1}{E}X_{it} + \frac{1}{NE}\sum_{j\in N} X^g_{j,t+1}\right)\sum_{e\in E} p^g_{et} + N\rho\frac{1}{E}\sum_{e\in E} p^g_{et}\sum_{e\in E} q^g_{iet} + O_e, \end{aligned} \tag{51}$$

where O_e is a term whose conditional expectation does not depend on the choice of $\{g_{iet}\}_{e\in E}$. Equivalently, we can express (51) as

$$\begin{aligned} p^g_{e,t+1} q^g_{ie,t+1} = &\left(\frac{1}{\zeta N}\frac{1}{E^2}\sum_{j\in N} X^g_{j,t+1} - \frac{N\rho^2}{\zeta}B_t - \frac{N\rho\chi}{\zeta} - \frac{1}{E}\frac{\chi}{\zeta}\right)\sum_{e\in E} q^g_{iet} \\ &- N\rho\left(-\frac{1}{E}X_{it} + \frac{1}{NE}\sum_{j\in N} X^g_{j,t+1}\right)\sum_{e\in E}\frac{\eta_{et} + (N-1)\rho B_t + (N-1)\chi + q^g_{iet}}{\zeta(N-1)} \\ &+ N\rho\frac{1}{E}\sum_{e\in E}\frac{\eta_{et} + (N-1)\rho B_t + (N-1)\chi + q^g_{iet}}{\zeta(N-1)}\sum_{e\in E} q^g_{iet} + O_e. \end{aligned} \tag{52}$$

By substituting (50) and (52) into (49), recalling that by definition $q^g_{iet} = g_{iet}(\kappa_{et})$, and ignoring terms whose conditional expectation does not depend on the choice $\{g_{iet}\}_{e\in E}$ we have transformed the objective function in the Bellman equation into

$$\mathbb{E}_{it}\left[A\left(\sum_{e\in E} g_{iet}(\kappa_{et})\right)^2 + B\sum_{e\in E} g_{iet}(\kappa_{et}) + C\delta_{it},\right], \tag{53}$$

where

$$\delta_{it} = \sum_{e\in E}(\eta_{et} + (N-1)\rho B_t + (N-1)\chi) g_{iet}(\kappa_{et}) + g_{iet}(\kappa_{et})^2,$$

for coefficients

$$
\begin{aligned}
A &= -b - N\rho \frac{1}{\zeta(N-1)} e^{-r\Delta} \\
B &= \mu_\pi \Delta - 2bX_{it} - e^{-r\Delta} \left[\frac{1}{\zeta N} \frac{1}{E} \sum_{j \in N} X^g_{j,t+1} - \frac{NE\rho^2}{\zeta} B_t - \frac{NE\rho\chi}{\zeta} - \frac{\chi}{\zeta} \right] \\
&\quad + \frac{e^{-r\Delta} N\rho}{\zeta(N-1)} \left(-X_{it} + \frac{1}{N} \sum_{j \in N} X^g_{j,t+1} \right) - \frac{e^{-r\Delta} N\rho}{\zeta(N-1)} \sum_{e \in E} \big(\eta_{et} + (N-1)\rho B_t + (N-1)\chi \big) \\
C &= -\frac{1}{\zeta(N-1)}.
\end{aligned}
\tag{54}
$$

Next, for each $e \in E$, we let

$$
g_{iet}(\kappa_{et}) = \hat{f}_{iet}(\kappa_{et}) + \nu h_{iet}(\kappa_{et}),
$$

for an arbitrary measurable deviation h_{iet} in $\tilde{\mathcal{M}}$ from the conjectured optimal $\hat{f}_{iet}$, and for some arbitrary constant ν. Substituting into (53) leaves

$$
\begin{aligned}
\mathbb{E}_{it} \Bigg[& A \left(\sum_{e \in E} \hat{f}_{iet} + \nu \sum_{e \in E} h_{iet} \right)^2 + B \sum_{e \in E} (\hat{f}_{iet} + \nu h_{iet}) \\
& + C \sum_{e \in E} \big(\eta_{et} + (N-1)\rho B_t + (N-1)\chi \big) (\hat{f}_{iet} + \nu h_{iet}) + (\hat{f}_{iet} + \nu h_{iet})^2 \Bigg],
\end{aligned}
\tag{55}
$$

where we have suppressed the argument κ_{et} from the notation, and will continue to do so whenever convenient. Taking a derivative with respect to ν, evaluating the derivative at $\nu = 0$, and setting the derivative equal to 0 gives the necessary optimality condition

$$
\mathbb{E}_{it} \left[2A \sum_{e \in E} \hat{f}_{iet} \sum_{e \in E} h_{iet} + B \sum_{e \in E} h_{iet} + C \sum_{e \in E} (\eta_{et} + (N-1)\rho B_t + (N-1)\chi) h_{iet} + 2\hat{f}_{iet} h_{iet} \right] = 0,
$$

which holds if, for each $k \in E$,

$$
\mathbb{E}_{it} \left[2A \sum_{e \in E} \hat{f}_{iet} + B + C \big(\eta_{et} + (N-1)\rho B_t + (N-1)\chi \big) \,\middle|\, \eta_{kt} \right] = -2C \hat{f}_{ikt}. \tag{56}
$$

The necessary condition (56) is also sufficient for optimality if the second derivative of (55) with respect to ν is negative, that is,

$$\mathbb{E}_{it}\left[A\left(\sum_{e\in E}h_{iet}\right)^2+C\sum_{e\in E}h_{iet}^2\right]<0. \tag{57}$$

To see why, suppose for contradiction that there exists a candidate $(L_{i1t},...,L_{iEt})$ in $\tilde{\mathcal{M}}$ satisfying the first order condition (56) that achieves a strictly higher value of the objective than $(\hat{f}_{i1t},\ldots,\hat{f}_{iEt})$. In that case, let $h_{iet}=L_{iet}-\hat{f}_{iet}$ for each $e\in E$. Then (55) achieves a higher value at $\nu=1$ than at $\nu=0$. This is a contradiction since (57) ensures that (55) is maximized at $\nu=0$.

Step 3. We derive the E, ζ, ρ, and χ such that (56) holds and then show that (57) is satisfied. This implies that we have found a solution to the Bellman equation. We first derive the moments in (56). By (45),

$$\mathbb{E}_{it}\left[\sum_{e\in E}\hat{f}_{iet}\;\middle|\;\eta_{kt}\right]=-X_{it}+\frac{1}{N}\mathbb{E}_{it}\left[\sum_{j\in N}X_{jt}+Q_{kt}\;\middle|\;\eta_{kt}\right]$$

and

$$\begin{aligned}\mathbb{E}_{it}[B\mid\eta_{kt}]&=\\ &-\left(2b+\frac{2e^{-r\Delta}N\rho}{\zeta(N-1)}\right)X_{it}+\left(-\frac{e^{-r\Delta}}{\zeta NE}+\frac{e^{-r\Delta}\rho(N+1)}{\zeta(N-1)}\right)\mathbb{E}_{it}\left[\sum_{j\in N}X_{j,t}+Q_{kt}\;\middle|\;\eta_{kt}\right]\\ &\qquad+e^{-r\Delta}\frac{\chi}{\zeta}+\mu_\pi\Delta.\end{aligned} \tag{58}$$

By the projection theorem,

$$\mathbb{E}_{it}\left[\sum_{j\in N}X_{jt}+Q_{kt}\;\middle|\;\eta_{kt}\right]=\left(\frac{N-1}{N}B_t+X_{it}\right)(1-\Gamma)\frac{E-1}{E}-(1+\Gamma(E-1))\left(\eta_{kt}-\frac{1}{E}X_{it}\right), \tag{59}$$

where

$$\Gamma=\frac{(N-1)\sigma_\epsilon^2}{(N-1)\sigma_\epsilon^2+E\sigma_Q^2}.$$

Finally, we use the fact that

$$\hat{f}_{ikt}=-\frac{1}{E}X_{it}-\frac{1}{N}\left(\eta_{kt}-\frac{1}{E}X_{it}\right)$$

and match coefficients in (56). Matching the coefficient on X_{it} gives

$$\frac{2C}{E} = 2A\frac{1}{N}(1-\Gamma)\frac{E-1}{E} + e^{-r\Delta}\left(\frac{-1}{\zeta NE} + \frac{\rho(N+1)}{\zeta(N-1)}\right)(1-\Gamma)\frac{E-1}{E} + \frac{C}{E}. \tag{60}$$

Matching the coefficient on $\eta_{kt} - \frac{1}{E}X_{it}$ gives

$$\frac{2C}{N} = -2A\frac{1}{N}(1+\Gamma(E-1)) - e^{-r\Delta}\left(\frac{-1}{\zeta NE} + \frac{\rho(N+1)}{\zeta(N-1)}\right)(1+\Gamma(E-1)) + C. \tag{61}$$

Matching the coefficient on B_t gives

$$(1-N)\rho C = 2A\frac{N-1}{N^2}(1-\Gamma)\frac{E-1}{E} + e^{-r\Delta}\left(\frac{-1}{\zeta NE} + \frac{\rho(N+1)}{\zeta(N-1)}\right)\frac{N-1}{N}(1-\Gamma)\frac{E-1}{E}. \tag{62}$$

Matching the constant coefficient gives

$$0 = C(N-1)\chi + e^{-r\Delta}\frac{\chi}{\zeta} + \mu_\pi\Delta. \tag{63}$$

Using (60) and (61), we have

$$\frac{N-2}{N} = \frac{1+\Gamma(E-1)}{(1-\Gamma)(E-1)}.$$

Rearranging gives

$$E = \frac{2N-2}{N-2-N\frac{\Gamma}{1-\Gamma}}. \tag{64}$$

As an aside, this expression is useful in so far as it characterizes the efficient number of exchanges in a partial equilibrium model in which strategic traders percieve the correlation in exchange prices to be Γ. Taking Γ as given, the analysis does not depend on σ_ϵ^2 or σ_Q^2.

We deduce from (64) that

$$E = 2 + \frac{2}{N-2} + \frac{N(N-1)}{N-2}\frac{\sigma_\epsilon^2}{\sigma_Q^2}. \tag{65}$$

Thus the number of exchanges achieving the efficient allocation is precisely that of the static case, as stated by the Theorem.

Next, using (60) and (62), we can solve for

$$\rho = -\frac{1}{NE}.$$

Now, in order to solve for ζ, we use (17) with (61) to get

$$\frac{1}{N} = \frac{N-1}{N} + e^{-r\Delta}\frac{N-1}{N}\left(N\rho - \frac{1}{E}\right)(1+\Gamma(E-1)) - 2b\zeta(N-1)\frac{1}{N}(1+\Gamma(E-1)),$$

which rearranges to

$$2b\zeta(N-1)(1+\Gamma(E-1)) = N-2-e^{-r\Delta}(N-1)\frac{2}{E}(1+\Gamma(E-1)).$$

Thus

$$\zeta = \frac{N-2}{2b(N-1)(1+\Gamma(E-1))} - e^{-r\Delta}\frac{2N-2}{(N-1)2bE}.$$

Using (63) we find

$$\chi = \frac{\mu_\pi \Delta}{1-e^{-r\Delta}}\zeta. \tag{66}$$

The within-period price impact is

$$\frac{1}{\zeta(N-1)} = \frac{2b(1+\Gamma(E-1))}{N-2-e^{-r\Delta}\frac{2N-2}{E}(1+\Gamma(E-1))},$$

as stated in the Theorem. Comparing with the static model, we see that price impact is higher in the dynamic model. We now verify that $\zeta > 0$ by showing that

$$N-2 > e^{-r\Delta}\frac{2N-2}{E}(1+\Gamma(E-1)).$$

The above equality holds since (64) implies that

$$(N-2)E = (2N-2)(1+\Gamma(E-1)).$$

Using (20), the cross-period cross-exchange price impact is

$$\frac{dp_{e,t+1}}{dq_{kt}} = -N\rho\frac{1}{(N-1)\zeta} = \frac{1}{E}\frac{2b(1+\Gamma(E-1))}{N-2-e^{-r\Delta}\frac{2N-2}{E}(1+\Gamma(E-1))},$$

as stipulated by the Theorem. Finally, we verify the sufficient condition for optimality (57) is negative by showing that

$$A\left(\sum_{e\in E} h_{iet}\right)^2 + C\sum_{e\in E} h_{iet}^2 < 0.$$

Using (54) and (17), this is equivalent to

$$\left(-b+\frac{1}{E}\frac{1}{\zeta(N-1)}e^{-r\Delta}\right)\left(\sum_{e\in E}h_{iet}\right)^2-\frac{1}{\zeta(N-1)}\sum_{e\in E}h_{iet}^2<0,$$

which holds by Jensen's inequality because $\zeta>0$. Thus, (48) solves the Bellman equation when the domain of admissible demand functions is restricted to $\tilde{\mathcal{M}}$.

Step 4. In this step, we show that if any measurable $g_{iet}:\mathbb{R}^3\to\mathbb{R}$ outside of $\tilde{\mathcal{M}}$ is chosen, the objective associated with the Bellman equation is $-\infty$. Towards this end, consider the terms in (53) involving $(\sum_{e\in E}g_{iet})^2$ and $\sum_{e\in E}g_{iet}^2$, which sum to

$$\left[-b+\frac{1}{E\zeta(N-1)}e^{-r\Delta}\right]\left(\sum_{e\in E}g_{iet}\right)^2-\frac{1}{\zeta(N-1)}\sum_{e\in E}g_{iet}^2.$$

By Jensen's inequality, the above expression is less than

$$-b\left(\sum_{e\in E}g_{iet}\right)^2-(1-e^{-r\Delta})\frac{1}{\zeta(N-1)}\sum_{e\in E}g_{iet}^2. \tag{67}$$

The other terms in (53) are $B\sum_{e\in E}g_{iet}$, which is only linear in $\sum_{e\in E}g_{iet}$, with B having finite variance, and $C\sum_{e\in E}(\eta_{et}+(N-1)\rho B_t+(N-1)\chi)g_{iet}$, where each η_{et} is of finite variance. We define J_e by

$$Bg_{iet}+C\left(\eta_{et}+(N-1)\rho B_t+(N-1)\chi\right)g_{iet}=J_eg_{iet}.$$

Note that each J_e is of finite variance.

Then

$$\begin{aligned}\mathbb{E}_{it}\left[-\frac{1-e^{-r\Delta}}{\zeta(N-1)}g_{iet}^2+J_eg_{iet}\right]=&\\ &\int_{\left\{\omega\in\Omega:|J_e|>|\frac{1-e^{-r\Delta}}{2\zeta(N-1)}g_{iet}|\right\}}\left(-\frac{1-e^{-r\Delta}}{\zeta(N-1)}g_{iet}^2+J_eg_{iet}\right)d\mathbb{P}(\omega)\\ &+\int_{\left\{\omega\in\Omega:|J_e|\le|\frac{1-e^{-r\Delta}}{2\zeta(N-1)}g_{iet}|\right\}}\left(-\frac{1-e^{-r\Delta}}{\zeta(N-1)}g_{iet}^2+J_eg_{iet}\right)d\mathbb{P}(\omega).\end{aligned}$$

The first integral must be finite since J_e is a finite-variance random variable and the integrand satisfies

$$\left| -\frac{1-e^{-r\Delta}}{\zeta(N-1)} g_{iet}^2 + J_e g_{iet} \right| \leq K J_e^2,$$

for some constant $K \in \mathbb{R}$. The second integral must be $-\infty$ since the integrand satisfies

$$-\frac{1-e^{-r\Delta}}{\zeta(N-1)} g_{iet}^2 + J_e g_{iet} \leq -\frac{1-e^{-r\Delta}}{2\zeta(N-1)} g_{iet}^2.$$

Thus, if g_{iet} is of infinite variance then the second integral must be $-\infty$. Hence, in this case,

$$\mathbb{E}_{it}\left[-\frac{1-e^{-r\Delta}}{\zeta(N-1)} g_{iet}^2 + J_e g_{iet} \right] = -\infty.$$

Inspecting (67) and (53) we see that if a chosen g_{iet} is not in $\tilde{\mathcal{M}}$, the objective function would equal to $-\infty$.

Step 5. We now check that the transversality condition (47) holds. We compute the moments involved in the terms M_s defining the candidate value function V of (48). For $s \geq t$,

$$\mathbb{E}_{it}\left[X_{is} + \sum_{e \in E} q_{es}^{\hat{f}} \right] = \frac{1}{N} X_{it} + \frac{N-1}{N^2} B_t$$

and

$$\mathbb{E}_{it}\left[\left(X_{is} + \sum_{e \in E} q_{es}^{\hat{f}} \right)^2 \right] = \frac{1}{N^2} \left[(X_{it} + \frac{N-1}{N} B_t)^2 + \sigma_\epsilon^2 \big(N(s-t) + N - 1 \big) + \sigma_Q^2 (s-t+1) \right]$$

For $s \geq t+1$ and $e \in E$,

$$\begin{aligned}
\mathbb{E}_{it}[p_{es}^{\hat{f}} q_{ies}^{\hat{f}}] &= \mathbb{E}_{it}\left[\frac{-\frac{1}{E}\sum_{j\in N} X_{js} - \frac{1}{E} W_{s-1} - Q_{es}}{\zeta N} \left(-\frac{1}{E} X_{is} + \frac{1}{NE} \sum_{j \in N} X_{js} + \frac{Q_{es}}{N} \right) \right] \\
&= \mathbb{E}_{it}\left[\frac{-\frac{1}{E}\sum_{j\in N} X_{js} - \frac{1}{E} W_{s-1} - Q_{es}}{\zeta N} \left(\frac{1}{NE} \sum_{j \in N} \epsilon_{js} + \frac{Q_{es}}{N} \right) \right] \\
&= -\frac{\sigma_Q^2}{E\zeta N^2} - \frac{\sigma_\epsilon^2}{N\zeta E^2}.
\end{aligned}$$

Next,

$$\mathbb{E}_{it}[p^{\hat{f}}_{et}q^{\hat{f}}_{iet}] = \mathbb{E}_{it}\left[\frac{-\frac{1}{E}\sum_{j\in N}X_{jt}-\frac{1}{E}B_t-Q_{et}}{\zeta N}\left(-\frac{1}{E}X_{it}+\frac{1}{NE}\sum_{j\in N}X_{jt}+\frac{Q_{et}}{N}\right)\right]$$
$$=\frac{N-1}{\zeta N^2E^2}X_{it}^2+\frac{2}{E^2\zeta N}(\frac{N-1}{N})^2X_{it}B_t-\frac{N-1}{N^2E^2}\frac{2N-1}{N^2\zeta}B_t^2-\frac{\sigma_Q^2}{E\zeta N^2}-\frac{(N-1)\sigma_\epsilon^2}{E^2\zeta N^2}.$$

Substituting these moments into (48) we find that

$$\begin{aligned}V(X_{it},B_t) = {} & X_{it}^2\left[\frac{-b}{N^2(1-e^{-r\Delta})}-\frac{N-1}{\zeta N^2E}\right]+\mu_\pi\Delta\frac{1}{N(1-e^{-r\Delta})}X_{it}\\
&+\mu_\pi\Delta\frac{N-1}{N^2(1-e^{-r\Delta})}B_t-\left[-2b\frac{N-1}{N^3(1-e^{-r\Delta})}-\frac{2}{E\zeta N}(\frac{N-1}{N})^2\right]B_tX_{it}\\
&+[-b(\frac{N-1}{N})^2\frac{1}{N^2(1-e^{-r\Delta})}+\frac{N-1}{N^2E}\frac{2N-1}{N^2\zeta}]B_t^2\\
&+\sigma_Q^2\left[\frac{1}{\zeta N^2(1-e^{-r\Delta})}-\frac{b}{N^2(1-e^{-r\Delta})^2}\right]\\
&+\sigma_\epsilon^2\left[-\frac{be^{-r}}{N(1-e^{-r\Delta})^2}-\frac{b(N-1)}{N^2(1-e^{-r\Delta})}+\frac{1}{N\zeta E(1-e^{-r\Delta})}-\frac{1}{E\zeta N^2}\right].\end{aligned}$$

Recall that an admissible strategy must lead to an inventory process that satisfies the no-Ponzi scheme condition $e^{-r\Delta t}\mathbb{E}_{i0}[X_{it}^2]\to 0$. Thus to show that $e^{-r\Delta t}\mathbb{E}_{i0}[V(X_{it},B_t)]\to 0$ it suffices to show that $e^{-r\Delta t}\mathbb{E}_{i0}[B_tX_{it}]\to 0$ and $e^{-r\Delta t}\mathbb{E}_{i0}[B_t^2]\to 0$.

We have

$$\begin{aligned}e^{-r\Delta t}\mathbb{E}_{i0}\left[B_tX_{it}\right] &= e^{-r\Delta t}\mathbb{E}_{i0}\left[\frac{N}{N-1}\sum_{j\neq i}\left(X_{j,t-1}+\sum_{e\in E}q_{je,t-1}\right)X_{it}\right]\\
&= e^{-r\Delta t}\mathbb{E}_{i0}\left[\frac{N}{N-1}\left(\sum_{j\in N}X_{j,t-1}+\sum_{e\in E}Q_{e,t-1}-X_{i,t}+\epsilon_{it}\right)X_{it}\right]\\
&= \frac{N}{N-1}e^{-r\Delta t}\mathbb{E}_{i0}\left[\left(\sum_{j\in N}X_{j,t-1}+\sum_{e\in E}Q_{e,t-1}\right)X_{it}-X_{it}^2\right]+e^{-r\Delta t}\frac{N}{N-1}\sigma_\epsilon^2,\end{aligned}$$

where, for the first equality, we have used

$$\frac{N-1}{N}B_t=\sum_{j\neq i}\left(X_{j,t-1}+\sum_{e\in E}q_{je,t-1}\right),$$

and for the second equality we have used $X_{it}=X_{i,t-1}+\sum_{e\in E}q_{ie,t-1}+\epsilon_{it}$. Since, by Cauchy-

Schwarz,

$$e^{-r\Delta t}\mathbb{E}_{i0}\left[\left(\sum_{j\in N} X_{j,t-1}+\sum_{e\in E} Q_{e,t-1}\right) X_{it}\right] \le \sqrt{\mathbb{E}_{i0}\left[\left(\sum_{j\in N} X_{j,t-1}+\sum_{e\in E} Q_{e,t-1}\right)^2\right] e^{-2r\Delta t}\mathbb{E}\left[X_{it}^2\right]},$$

it follows that

$$\lim_{t\to\infty} e^{-r\Delta t}\mathbb{E}_{i0}\left[\left(\sum_{j\in N} X_{j,t-1}+\sum_{e\in E} Q_{e,t-1}\right) X_{it}\right] = 0.$$

Thus, $\lim_{t\to\infty} e^{-r\Delta t}\mathbb{E}_{i0}[B_t X_{it}] = 0$, as desired. That $\lim_{t\to\infty} e^{-r\Delta t}\mathbb{E}_{i0}[B_t^2] = 0$ can be shown using the same method.

Step 6. We now verify that the optimal strategy of trader i is the conjectured equilibrium strategy, coinciding with (45). For an arbitrary admissible strategy, which we denote l, let q_{iet}^l, p_{et}^l, X_{it}^l, and B_t^l denote, respectively, the induced quantity purchased on exchange e, the price on exchange e, the inventory, and the belief at date t. By recursive substitution, using the Bellman equation, for each $t \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}_{i0}\left[V(X_{i0}, B_0)\right] \ge \mathbb{E}_{i0}&\left[\sum_{s=0}^{t} e^{-rs}\left(\mu_\pi \Delta\left(X_{is}^l+\sum_{e\in E} q_{ies}^l\right) - b\left(X_{is}^l+\sum_{e\in E} q_{ies}^l\right)^2 - \sum_{e\in E} p_{es} q_{ies}^l\right)\right] \\ &+ \mathbb{E}_{i0}\left[e^{-r\Delta t} V(X_{i,t+1}^l, B_{t+1}^l)\right].\end{aligned}$$

By taking limits as $t \to \infty$, applying the transversality condition and Fatou's Lemma,

$$V(X_{i0}, B_0) \ge \mathbb{E}_{i0}\left[\sum_{s=0}^{\infty} e^{-r\Delta s}\left(\mu_\pi \Delta\left(X_{is}^l+\sum_{e\in E} q_{ies}^l\right) - b\left(X_{is}^l+\sum_{e\in E} q_{ies}^l\right)^2 - \sum_{e\in E} p_{es} q_{ies}^l\right)\right].$$

The right-hand side is the utility of the arbitrary strategy l, whereas the left-hand side is the utility of the conjectured equilibrium strategy. This completes the proof of the Theorem.

F Extension: Endogeneous Liquidity Trade

This appendix offers an extension in which liquidity traders, who are local to each exchange and conduct no cross-exchange trade, choose the sizes of their trades.

F.1 Setup

In this section we extend the baseline model by allowing liquidity traders to endogenously choose the quantity of market orders that they supply. There are M liquidity traders who are each restricted to trade on a single exchange. We assume that M is divisible by E and that a fraction $1/E$ of them trade on any given exchange. Liquidity trader j has endowment

$$H_j \sim N(0, \frac{1}{M}\sigma_H^2)$$

where the $\{H_j\}$ are mutually independent. Suppose further that each liquidity trader j has preferences of the same form that we have assumed for the strategic traders. If liquidity trader j is restricted to trade on exchange e, his or her ex-ante expected utility of purchasing h_j units via a market order is

$$\mathbb{E}[\pi h_j - c(H_j + h_j)^2 - h_j p_e \,|\, H_j, h_j].$$

Above $c \in \mathbb{R}_+$ is the holding cost parameter of the liquidity traders. It is useful to think of c being high relative to b, the holding cost parameter of strategic agents. Finally, for simplicity, for this section only, we assume that $\mu_X = 0$ and $\mu_\pi = 0$.

F.2 Analysis

Theorem 6. *There exists a symmetric affine equilibrium. In any symmetric affine equilibrium the following are true.*

1. *The quantity of market orders submitted by agent j is*

$$h_j = \frac{-cH_j}{c + \Lambda_E \frac{N-1}{N}}.$$

2. *For each $e, e' \in E$ distinct, the correlation between prices in the two exchanges from the perspective of a strategic trader is*

$$\gamma_E = \frac{(E\alpha_E)^2\sigma_X^2(N-1)}{(E\alpha_E)^2\sigma_X^2(N-1) + (\frac{c}{c+\Lambda_E\frac{N-1}{N}})^2\sigma_H^2 E} \tag{68}$$

3. *A strategic trader's price impact satisfies*

$$\Lambda_E = \frac{2b((E-1)\gamma_E + 1)}{N-2} \tag{69}$$

while the price impact of a liquidity trader is

$$\frac{N-1}{N}\Lambda_E. \tag{70}$$

4. $E\alpha_E$ *satisfies*

$$E\alpha_E = \frac{1}{\gamma_E(\frac{1}{N}+\frac{1}{N-2})+(1-\gamma_E)\frac{1}{E}(\frac{1}{N}+\frac{1}{N-2})+\frac{N-1}{N}}. \tag{71}$$

Proof. We conjecture that there exists a symmetric affine equilibrium in which each strategic trader $i \in N$ submits a demand schedule of the form $-\alpha_E X_i - \zeta_E p$ and each liquidity trader j submits a market order of the form $-\tilde{\alpha}_E H_j$. We study the best response problem of trader $j \in M$. Via market clearing, we can compute the market clearing price in exchange e is

$$p_e = \frac{\sum_{i\in N} -\alpha_E X_i - \sum_{\{k\in M\,|\,k\neq j\}} \tilde{\alpha}_E H_k + h_j}{N\zeta_E}$$

if all agents $i \in N$ and $k \in M$ such that $k \neq j$ behave as conjectured and agent j purchases h_j units on the exchange. Retaining the notation that $\Lambda_E = \frac{1}{(N-1)\zeta_E}$ the price impact of liquidity trader j is $\Lambda_E\frac{N-1}{N}$. He seeks to maximize

$$\mathbb{E}[\pi h_j - c(H_j+h_j)^2 - h_j p_e \,|\, H_j, h_j] = -c(H_j+h_j)^2 - \Lambda_E\frac{N-1}{N}h_j^2$$

by choosing $h_j \in \mathbb{R}$. Taking a first order condition with respect to h_j we have

$$-2c(H_j+h_j) - 2h_j\Lambda_E\frac{N-1}{N} = 0,$$

which implies that

$$h_j = \frac{-cH_j}{c+\Lambda_E\frac{N-1}{N}}.$$

Thus

$$\tilde{\alpha}_E = \frac{c}{c+\Lambda_E\frac{N-1}{N}}.$$

If strategic traders take the variance of aggregate liquidity trade to be

$$\sigma_Q^2 = \left(\frac{c}{c+\Lambda_E\frac{N-1}{N}}\right)^2\sigma_H^2,$$

we see that the analysis of the baseline model applies. That is, strategic traders maximize by submitting affine demand schedules such that equations (68), (69) and (71) are satisfied. Then the analysis of the baseline model therefore ensures that provided there exists α_E and γ_E which satisfies (68), (69), and (71), there exists a symmetric affine equilibrium with the four properties given in the statement of the theorem. To show existence it suffices to recognize that substituting expressions (69) and (71) into (68) and re-arranging yields a cubic equation in γ_E. Since the equation is cubic there always exists at least one real root. Thus there always exists a solution to the system of equations. □

The above theorem has characterized a symmetric affine equilibrium of the model with endogenous liquidity traders. The following proposition states some results relevant for assessing the allocative efficiency of the symmetric affine eqiulibrium.

Proposition 5. *The following are true of any symmetric affine equilibrium.*

1. $E\alpha_E \in [\frac{N-2}{N-1}, \frac{N}{N-1}]$ *is always higher in fragmented markets than in centralized markets.*
2. *Fixing arbitrary E, in the limit as c tends to infinity, the expected sum of liquidity traders' holding costs tends to zero.*
3. *Fixing arbitrary $E > 1$, for all c sufficiently large, a market with E exchanges is more efficient than a market with a single exchange in the sense that the expected sum of all traders' holding costs is lower.*
4. *For any $\overline{E} > 1$, there exists an $\bar{c}$ such that if $c > \bar{c}$ then a market with $1 < E \leq \overline{E}$ exchanges is more efficient than a market with a single exchange in the sense that the expected sum of all traders' holding costs is lower.*

Proof. Centralized markets correspond to the case when E is 1. To prove Part 1, it is clear by inspecting (71) that $E\alpha_E \in [\frac{N-2}{N-1}, \frac{N}{N-1}]$. Next recognize that in fragmented markets $E > 1$ and $\gamma_E < 1$ so that again by inspection, $E\alpha_E$ is always higher in fragmented markets.

To prove part 2 recognize that, using part 1 of Theorem 6, the expected sum of liquidity agents' holding costs is

$$c\left(\frac{\Lambda_E \frac{N-1}{N}}{c + \Lambda_E \frac{N-1}{N}}\right)^2 \sigma_H^2,$$

which decays to 0 as c diverges.

To prove part 3, fix $E > 1$ and inspect equation (68). Since $E\alpha_E \in [\frac{N-2}{N-1}, \frac{N}{N-1}]$ there exists $a, b \in \mathbb{R}$ such that $1 > b > a > 0$ and $\gamma_E \in [a, b]$ for all c sufficiently large. This implies that $|1 - E\alpha_E|$ is bounded above by a constant strictly less than $\frac{1}{N-1}$ whenever c is sufficiently large. In the limit as $c \to \infty$ the aggregate quantity of liquidity trader supply absorbed by strategic traders when there is a single exchange as well as when there are E exchanges becomes arbitrarily close to $\sum_{j\in M} H_j$. Therefore, by Proposition 5, in the limit as $c \to \infty$, the expected sum of holding costs is strictly lower when there are E exchanges than when there is a single exchange since $|\,1-\alpha_E\,| < |\,1-\alpha_1\,| = \frac{1}{N-1}$. However, the sum of liquidity traders' holding costs converges to 0 as $c \to \infty$. This implies the claim asserted in Part 3 of the theorem.

Part 4 is an immediate implication of part 3. □

We now prove the following proposition which implies that $E\alpha_E$ must be strictly monotone increasing in E at least until a certain cutoff point. As c increases the range that we can prove that $E\alpha_E$ is strictly monotone increasing in is larger.

Proposition 6. *Fix $E^* \in \mathbb{N}$. If c is sufficiently large such that*

$$\left(\frac{c}{c+\frac{2bE^*}{N-2}\frac{N-1}{N}}\right)^2 > \frac{E^*}{E^*+1},$$

then $E\alpha_E$ is strictly monotone increasing for all $E < E^*$.

Proof. We begin by proving that

$$\left(\frac{c}{c+\Lambda_E\frac{N-1}{N}}\right)^2 E$$

is strictly monotone increasing in E for all $E < E^*$. Since Λ_E is bounded above by $\frac{2bE^*}{N-2}$ we have that

$$\left(\frac{c}{c+\Lambda_E\frac{N-1}{N}}\right)^2 E > \frac{E^*}{E^*+1}E$$

for each $E < E^*$. Thus we have

$$\left(\frac{c}{c+\Lambda_{E+1}\frac{N-1}{N}}\right)^2 (E+1) - \left(\frac{c}{c+\Lambda_E\frac{N-1}{N}}\right)^2 E > \frac{E^*}{E^*+1}(E+1) - E$$

for each $E < E^*$. But the right hand side is equal to

$$\left(\frac{E^*}{E^*+1} - 1\right)E + \frac{E^*}{E^*+1} > \left(\frac{E^*}{E^*+1} - 1\right)E^* + \frac{E^*}{E^*+1} = 0.$$

Now we prove that $E\alpha_E$ is strictly monotone increasing at each $E < E^*$. Inspect the equation (71). Suppose $E\alpha_E$ is decreasing in E then it must be that γ_E is increasing. Consider now

(68). Since $(\frac{c}{c+\Lambda_E \frac{N-1}{N}})^2 E$ is strictly monotone increasing and $E\alpha_E$ is decreasing it must be that γ_E is decreasing, a contradiction. $\square$

G Extension: Private Information about Asset Payoff

This appendix addresses an extension of the model in which strategic traders are asymmetrically informed about the asset payoff.

G.1 Setup

We alter the baseline model so that each agent has private information about the asset's final payoff, $\pi \sim N(\mu_\pi, \sigma_\pi^2)$. We assume the aggregate endowment of strategic traders, $Z \equiv \sum_i X_i$, is public information. As before, liquidity traders supply a quantity $Q_e \sim N(0, \frac{\sigma_Q^2}{E})$ to each exchange, independent across exchanges. Strategic traders receive private signals of π:

$$S_i = \pi + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ is i.i.d across individuals.

G.2 Analysis

Theorem 7. *In any symmetric affine equilibrium with demand schedules which are each monotone decreasing in price,*

1. *Each strategic trader i submits a demand schedule to each exchange e of the form*

$$f_{ie}(X_i, p) = -\alpha X_i - \zeta p + w S_i + \Delta.$$

where α, ζ, w, and Δ are defined by the system of equations (73)—(80).

2. *Price impact is*

$$\Lambda_E = \frac{(2b[(E-1)\tilde{\gamma}_1 + 1] + N\frac{\tilde{\gamma}_3}{w})}{N-2}$$

where $\tilde{\gamma}_1$ and $\tilde{\gamma}_3$ are defined by equations (73) and (75).

3. *The final inventory of strategic trader i is*

$$X_i + \sum_{e\in E} f_{ie}(X_i, p_e^f) = (1-E\alpha)X_i + E\alpha\frac{1}{N}\sum_{j\in N} X_j + Ew\left(S_i - \frac{1}{N}\sum_{j\in N} S_j\right) + \frac{\sum_{e\in E} Q_e}{N}.$$

Proof. Conjecture a symmetric affine equilibrium in which agent i submits demand schedule

$$f_{ie}(X_i, p) = -\alpha X_i - \zeta p + w S_i + \Delta$$

to exchange $e \in E$ for each $i \in N$ and $e \in E$. By market clearing the residual supply curve trader i faces in exchange e is

$$p_e(q) = \frac{1}{(N-1)\zeta}\left[\sum_{j \neq i}(-\alpha X_j + w S_j + \Delta) - Q_e + q\right].$$

Thus price impact is $\Lambda = \frac{1}{(N-1)\zeta}$. Also by market clearing, the equilibrium price is

$$p_e^f = \frac{1}{N\zeta}\left[\sum_{j \in N}(-\alpha X_j + w S_j + \Delta) - Q_e\right].$$

Going forward, let us define $q_{ie}^f := f_{ie}(X_i, p_e^f)$ for each $e \in E$ for ease of notation. In any equilibrium, trader i must equate marginal utility with marginal cost for every realization of the price:

$$-2b(X_i + q_{i1}^f + (E-1)\mathbb{E}[q_{i2}^f \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, X_i, S_i]) = p_1^f - \mathbb{E}[\pi \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, X_i, S_i] + \frac{1}{(N-1)\zeta}q_{i1}^f. \quad (72)$$

Above we have used symmetry. We now compute the two conditional moments $\mathbb{E}[q_{i2}^f \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, S_i, X_i]$ and $\mathbb{E}[\pi \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, S_i, X_i]$ by using the projection theorem. We begin with the former. We can, using the projection theorem, express

$$\mathbb{E}\left[\sum_{\{j \in N \,|\, j \neq\}} S_i \,|\, p_1^f - \frac{q_{i1}^f}{(N-1)\zeta}, S_i\right] = \mu_\pi(N-1) + \gamma_1(p_1^f - \frac{q_{i1}^f}{(N-1)\zeta} - \frac{w\mu_\pi}{\zeta} + \alpha\frac{Z - X_i}{(N-1)\zeta} - \frac{\Delta}{\zeta}) + \gamma_2(S_i - \mu_\pi).$$

Here, γ_1 and γ_2 are derived as follows. The variables, $\sum_{j \neq i} S_j$,S_i, $p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f$ are jointly

Gaussian with variance matrix

$$\Sigma = \begin{bmatrix} (N-1)^2\sigma_\pi^2 + \sigma_\epsilon^2(N-1) & (N-1)\sigma_\pi^2 & \frac{w}{\zeta}(\sigma_\pi^2(N-1)+\sigma_\epsilon^2) \\ (N-1)\sigma_\pi^2 & \sigma_\pi^2+\sigma_\epsilon^2 & \frac{w}{\zeta}\sigma_\pi^2 \\ \frac{w}{\zeta}(\sigma_\pi^2(N-1)+\sigma_\epsilon^2) & \frac{w}{\zeta}\sigma_\pi^2 & \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] \end{bmatrix}.$$

Define

$$\overline{\Sigma} \equiv \begin{bmatrix} \sigma_\pi^2+\sigma_\epsilon^2 & \frac{w}{\zeta}\sigma_\pi^2 \\ \frac{w}{\zeta}\sigma_\pi^2 & \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}], \end{bmatrix}$$

with

$$\overline{\Sigma}^{-1} = [(\sigma_\pi^2+\sigma_\epsilon^2)\frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - \frac{w^2}{\zeta^2}\sigma_\pi^4]^{-1}$$
$$\times$$
$$\begin{bmatrix} \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) + \frac{\sigma_Q^2}{E(N-1)^2}] & -\frac{w}{\zeta}\sigma_\pi^2 \\ -\frac{w}{\zeta}\sigma_\pi^2 & \sigma_\pi^2+\sigma_\epsilon^2 \end{bmatrix}.$$

Define

$$\Sigma_{12} \equiv \begin{bmatrix} (N-1)\sigma_\pi^2 & \frac{w}{\zeta}(\sigma_\pi^2(N-1)+\sigma_\epsilon^2) \end{bmatrix}$$

By the rules of conditional normals

$$\begin{bmatrix} \gamma_2 & \gamma_1 \end{bmatrix} = \Sigma_{12}\overline{\Sigma}^{-1}.$$

This yields,

$$\gamma_2 = \frac{\frac{\sigma_\pi^2\sigma_Q^2}{E(N-1)}}{(\sigma_\pi^2+\sigma_\epsilon^2)[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - w^2\sigma_\pi^4}.$$

Note that $\frac{1}{N-1}\gamma_2 \in [0,1]$. Next, we have

$$\gamma_1 = \zeta\frac{w\sigma_\pi^2\sigma_\epsilon^2(N-1) + w\sigma_\epsilon^2(\sigma_\pi^2+\sigma_\epsilon^2)}{(\sigma_\pi^2+\sigma_\epsilon^2)[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - w^2\sigma_\pi^4}$$

Note that $\frac{w}{\zeta(N-1)}\gamma_1 \in [0,1]$. We have

$$\mathbb{E}[q_{i2}^f \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, S_i] = -\alpha X_i + wS_i + \Delta + \frac{\alpha Z}{N} - \frac{wS_i}{N} - \Delta$$
$$- \frac{w}{N}[\mu_\pi(N-1) + \gamma_1(p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f - \frac{w\mu_\pi}{y} + \alpha\frac{Z-X_i}{(N-1)\zeta} - \frac{m}{\zeta}) + \gamma_2(S_i-\mu_\pi)].$$

Next, we move on to compute, $\mathbb{E}[\pi \,|\, p_1^f - \frac{1}{(N-1)\zeta}q_{i1}^f, S_i, X_i]$. We can, using the rules of conditional normals, express

$$\mathbb{E}[\pi \mid p_1^f - \frac{q_{i1}^f}{(N-1)\zeta}, S_i] = \mu_\pi + \gamma_3 (p_1 - \frac{q_{i1}^f}{(N-1)\zeta} - \frac{w\mu_\pi}{\zeta} + \alpha \frac{Z - X_i}{(N-1)\zeta} - \frac{\Delta}{\zeta}) + \gamma_4 (S_i - \mu_\pi).$$

The variables, π,S_i, $p_1^f - \frac{q_{i1}^f}{(N-1)\zeta}$ are jointly Gaussian with variance matrix

$$\Sigma = \begin{bmatrix} \sigma_\pi^2 & \sigma_\pi^2 & \frac{w}{\zeta}\sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\epsilon^2 & \frac{w}{\zeta}\sigma_\pi^2 \\ \frac{w}{\zeta}\sigma_\pi^2 & \frac{w}{\zeta}\sigma_\pi^2 & \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] \end{bmatrix}.$$

Define

$$\overline{\Sigma} \equiv \begin{bmatrix} \sigma_\pi^2 + \sigma_\epsilon^2 & \frac{w}{\zeta}\sigma_\pi^2 \\ \frac{w}{\zeta}\sigma_\pi^2 & \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] \end{bmatrix}$$

and

$$\Sigma_{12} \equiv \begin{bmatrix} \sigma_\pi^2 & \frac{w}{\zeta}\sigma_\pi^2 \end{bmatrix}$$

Then

$$\begin{bmatrix} \gamma_4 & \gamma_3 \end{bmatrix} = \Sigma_{12}\overline{\Sigma}^{-1}.$$

We obtain,

$$\gamma_4 = \frac{\sigma_\pi^2 \frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - \frac{w^2}{\zeta^2}\sigma_\pi^4}{(\sigma_\pi^2 + \sigma_\epsilon^2)\frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - \frac{w^2}{\zeta^2}\sigma_\pi^4},$$

and

$$\gamma_3 = \frac{\frac{w}{\zeta}\sigma_\pi^2\sigma_\epsilon^2}{(\sigma_\pi^2 + \sigma_\epsilon^2)\frac{1}{\zeta^2}[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - \frac{w^2}{\zeta^2}\sigma_\pi^4}.$$

Note that $\gamma_3 \frac{w}{\zeta(N-1)} \in [0,1]$ and $\gamma_4 \in [0,1]$. It is useful, for the analysis to follow, to redefine the inference coefficients so that they all lie in the interval $[0,1]$. Specifically, define $\tilde{\gamma}_1 = \frac{w}{\zeta(N-1)}\gamma_1$, $\tilde{\gamma}_2 = \frac{1}{N-1}\gamma_2$, $\tilde{\gamma}_3 \equiv \frac{w}{\zeta(N-1)}\gamma_3$, and $\tilde{\gamma}_4 = \gamma_4$. Then

$$\tilde{\gamma}_1 = \frac{w^2\sigma_\pi^2\sigma_\epsilon^2 + \frac{w^2\sigma_\epsilon^2(\sigma_\pi^2+\sigma_\epsilon^2)}{N-1}}{w^2[(\sigma_\pi^2 + \sigma_\epsilon^2)(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) - \sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2 + \sigma_\epsilon^2)} \tag{73}$$

$$\tilde{\gamma}_2 = \frac{\frac{\sigma_\pi^2\sigma_Q^2}{E(N-1)^2}}{w^2[(\sigma_\pi^2 + \sigma_\epsilon^2)(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) - \sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2 + \sigma_\epsilon^2)}. \tag{74}$$

$$\tilde{\gamma}_3 = \frac{w^2\sigma_\pi^2\frac{\sigma_\epsilon^2}{N-1}}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4]+\frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)}. \tag{75}$$

$$\tilde{\gamma}_4 = \frac{\sigma_\pi^2[w^2(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{(N-1)})+\frac{\sigma_Q^2}{(E(N-1)^2)}]-w^2\sigma_\pi^4}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4]+\frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)}. \tag{76}$$

We can now use the equation (72) together with the conditional moments we just computed, to match coefficients and pin down α, ζ, w, and Δ. The coefficient of q_{i1} gathered on to the LHS is

$$-2b - \frac{1}{(N-1)\zeta} - 2b(E-1)\frac{1}{N}\tilde{\gamma}_1 - \frac{\tilde{\gamma}_3}{w}.$$

The coefficient of p_1 gathered on to the RHS is

$$1 - 2b(E-1)\frac{1}{N}\tilde{\gamma}_1(N-1)\zeta - \zeta\frac{(N-1)\tilde{\gamma}_3}{w}.$$

The coefficient of S_i gathered on to the RHS is

$$2b(E-1)w(\frac{N-1}{N})(1-\tilde{\gamma}_2) - \gamma_4.$$

The coefficient of X_i gathered on to the RHS is

$$2b + 2b(E-1)[-\alpha + \alpha\frac{\tilde{\gamma}_1}{N}] + \frac{\tilde{\gamma}_3}{w}\alpha.$$

The constant coefficient gathered on to the RHS is

$$2b(E-1)[\frac{\alpha Z}{N} - \frac{w}{N}(\mu_\pi(N-1) + \frac{\tilde{\gamma}_1\alpha Z}{w} - \frac{m\tilde{\gamma}_1(N-1)}{w} - \tilde{\gamma}_2(N-1)\mu_\pi - \tilde{\gamma}_1(N-1)\mu_\pi)]$$

$$-\mu_\pi + \tilde{\gamma}_3\mu_\pi(N-1) - \frac{\tilde{\gamma}_3\alpha Z}{w} + \tilde{\gamma}_3\frac{(N-1)m}{w} + \tilde{\gamma}_4\mu_\pi$$

We now match coefficients to compute y as a function of $\tilde{\gamma}_1$ and $\tilde{\gamma}_3$:

$$\zeta = \frac{N-2}{N-1}\frac{1}{(2b[(E-1)\tilde{\gamma}_1+1]+N\frac{\tilde{\gamma}_3}{w})}.$$

Price impact is therefore

$$\frac{1}{(N-1)\zeta} = \frac{(2b[(E-1)\tilde{\gamma}_1+1]+N\frac{\tilde{\gamma}_3}{w})}{N-2}. \tag{77}$$

Notice that compared with the model without private information about asset payoffs, there is now a $\frac{N\tilde{\gamma}_3}{w}$ term which is a result of using the price in an exchange to do inference on the asset's payoff, π. We now match coefficients to derive a cubic equation which characterizes w:

$$-2b(E-1)w\frac{N-1}{N}(1-\tilde{\gamma}_2)+\gamma_4= w[2b+\frac{(2b[(E-1)\tilde{\gamma}_1+1]+\frac{N}{w}\tilde{\gamma}_3)}{N-2}+2b(E-1)\frac{1}{N}\tilde{\gamma}_1+\frac{\tilde{\gamma}_3}{w}] \quad (78)$$

We now match coefficients to compute α as a function of the inference coefficients:

$$\alpha=\frac{2b}{2b+\frac{(2b[(E-1)\tilde{\gamma}_1+1]+N\frac{\tilde{\gamma}_3}{w})}{N-2}+2b(E-1)\frac{1}{N}\tilde{\gamma}_1+2b(E-1)(1-\frac{\tilde{\gamma}_1}{N})}. \quad (79)$$

We now match coefficients to compute Δ as a function of the inference coefficients:

$$\Delta=-[\frac{2b(E-1)[\frac{\alpha Z}{N}-\frac{w}{N}(\mu_\pi(N-1)+\frac{\tilde{\gamma}_1\alpha Z}{w}-\tilde{\gamma}_2(N-1)\mu_\pi-\tilde{\gamma}_1(N-1)\mu_\pi)]}{2b+\frac{1}{(N-1)y}+\frac{2b(E-1)\tilde{\gamma}_1}{N}+\frac{\tilde{\gamma}_3}{w}+\frac{2b(E-1)\tilde{\gamma}_1(N-1)}{N}+\frac{\tilde{\gamma}_3(N-1)}{w}} + \frac{-\mu_\pi+\tilde{\gamma}_3\mu_\pi(N-1)-\frac{\tilde{\gamma}_3\alpha Z}{w}+\tilde{\gamma}_4\mu_\pi}{2b+\frac{1}{(N-1)y}+\frac{2b(E-1)\tilde{\gamma}_1}{N}+\frac{\tilde{\gamma}_3}{w}+\frac{2b(E-1)\tilde{\gamma}_1(N-1)}{N}+\frac{\tilde{\gamma}_3(N-1)}{w}}]. \quad (80)$$

Thus equations (79), (77), (78), (80), (73), (74), (75), and (76) are necessary conditions that any symmetric affine equilibrium must satisfy. An argument analogous to that of Theorem 4 can be used to show that a solution to these equations constitute a symmetric affine equilibrium provided that y is positive. Part 2 follows from equation (77). This completes the proof of parts 1 and 2. We omit the proof of part 3 since it is a straightforward computation. □

Proposition 7. *For any value of E, if there exists a symmetric affine equilibrium $w>0$ if $\zeta>0$.*

Proof. Recall that a requirement of a symmetric affine equilibrium is that y is positive. The cubic equation characterizing w is

$$\tilde{\gamma}_4-\tilde{\gamma}_3=w[2b+\frac{1}{\zeta(N-1)}+2b(E-1)\frac{1}{N}\tilde{\gamma}_1+2b(E-1)(\frac{N-1}{N})(1-\tilde{\gamma}_2)].$$

The left hand side is positive as seen by inspecting the equations defining the inference coefficients. The bracketed term on the right hand side is also always positive if the demand schedules are downward sloping since the inference coefficients are in the unit interval. Thus the only way for the cubic equation to be satisfied is if w is positive. $\square$

We now focus on characterizing how Ew and $E\alpha$ change as E varies. In this model, the efficient allocation is the same as that of the baseline model. Thus by Part 3 of Theorem 7, perfect allocative efficiency is acheived if $Ew = 0$ and $E\alpha = 1$.

Proposition 8. *The following are true.*

1. *There exists a unique symmetric affine equilibrium when $E = 1$.*

2. *When there is just a single exchange,*

$$0 < w_1 < \frac{1}{2b}\frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_\epsilon^2}$$

   where w_1 corresponds to the unique symmetric affine equilibrium.

3. *There exist at least one and at most three symmetric affine equilibria for all E sufficiently large.*

4. *For any sequence $\{Ew_E\}$ corresponding to symmetric affine equilibria,*

$$Ew_E \to \frac{1}{2b}\frac{N}{N-1}\frac{\sigma_\pi^2}{\sigma_\epsilon^2} > \frac{1}{2b}\frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_\epsilon^2}$$

   as $E \to \infty$.

5. *For any sequence $\{E\alpha_E\}$ corresponding to symmetric affine equilibria $E\alpha_E \to 1$ which is strictly greater than α_1.*

Proof. Part 1. When there is a single exchange,

$$w_1 = \frac{\tilde{\gamma}_4 - (1 + \frac{N}{N-2})\tilde{\gamma}_3}{2b(1 + \frac{1}{N-2})}. \tag{81}$$

Rearranging (81), we derive

$$2b(1 + \frac{1}{N-2})w^3[(\sigma_\pi^2 + \sigma_\epsilon^2)(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) - \sigma_\pi^4] + w2b(1 + \frac{1}{N-2})\frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2 + \sigma_\epsilon^2)$$
$$= \sigma_\pi^2[w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{E(N-1)^2}] - w^2\sigma_\pi^4 - (1 + \frac{N}{N-2})w^2\sigma_\pi^2\frac{\sigma_\epsilon^2}{N-1}$$

Thus, when E is 1, w_1 satisfies a cubic equation with coefficients:

$$[w_1^3] : 2b(1 + \frac{1}{N-2})[(\sigma_\pi^2 + \sigma_\epsilon^2)(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) - \sigma_\pi^4]$$

$$[w_1^2] : \frac{N}{N-2}\sigma_\pi^2 \frac{\sigma_\epsilon^2}{N-1}$$

$$[w_1] : 2b(1 + \frac{1}{N-2})\frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2 + \sigma_\epsilon^2)$$

$$[constant] : -\frac{\sigma_\pi^2 \sigma_Q^2}{E(N-1)^2}.$$

Since the coefficient of w_1^2 is positive, the coefficient of w_1^3 is positive, and the constant is negative, there always exists exactly one positive real root. Let p, q, and r denote the roots of the cubic equation. Then $pqr = -\frac{constant\ coefficient}{coefficient\ of\ w_1^3} > 0$. Thus if there is one real root and 2 complex roots, the real root must be positive. If there are are three real roots, at least one must be positive. Next, $p + q + r = -\frac{coefficient\ of\ w_1^2}{coefficient\ of\ w_1^3} < 0$ so if there are three real roots, two must be negative and one must be positive. There always exists a unique positive real root. Take this positive real root. For this value of w_1, by (77), ζ_1 is positive. An approach analogous to that of Theorem 4 (which we omit) can then be used to verify that there is a symmetric affine equilibrium corresponding to this value of w_1. Since it is the unique positive real root, the equilibrium must be unique since (81) is a necessary condition which must be satisfied in any symmetric affine equilibrium.

Part 2. We rearrange (76) to derive

$$\tilde{\gamma}_4 = \frac{\sigma_\pi^2(w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) + \frac{\sigma_Q^2}{N-1}) - w^2\sigma_\pi^2}{(\sigma_\pi^2 + \sigma_\epsilon^2)(w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) + \frac{\sigma_Q^2}{N-1}) - w^2\sigma_\pi^2}$$

$$< \frac{\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}[(\sigma_\pi^2 + \sigma_\epsilon^2)(w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) + \frac{\sigma_Q^2}{N-1}) - w^2\sigma_\pi^2]}{(\sigma_\pi^2 + \sigma_\epsilon^2)(w^2(\sigma_\pi^2 + \frac{\sigma_\epsilon^2}{N-1}) + \frac{\sigma_Q^2}{N-1}) - w^2\sigma_\pi^2}$$

$$= \frac{\sigma^2\pi}{\sigma_\pi^2 + \sigma_\epsilon^2}.$$

Inspecting (81) together with the above inequality gives the result.

Parts 3 and 4. Rearranging equation (78), we derive

$$w_E = \frac{\tilde{\gamma}_4 - (1+\frac{N}{N-2})\tilde{\gamma}_3}{2b + \frac{2b}{N-2} + 2b(E-1)(\frac{1}{N}+\frac{1}{N-2})\tilde{\gamma}_1 + 2b(E-1)(\frac{N-1}{N})(1-\tilde{\gamma}_2)}.$$

we observe that $|\, w_E \,|$ is less than $\frac{C}{E}$ for large E for some constant C since $\tilde{\gamma}_2$ is by inspection bounded away from 1 (we can derive a bound which holds for all E) and the numerator is bounded above by $2 + \frac{N}{N-2}$. Thus, it must be the case that $\tilde{\gamma}_4 \to \frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}$ in the limit as $E \to \infty$. By inspection $\tilde{\gamma}_1$ and $\tilde{\gamma}_3$ converges to 0 while $\tilde{\gamma}_2 \to \frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}$. We can express

$$Ew_E = \frac{\tilde{\gamma}_4 - (1+\frac{N}{N-2})\tilde{\gamma}_3}{\frac{2b+\frac{2b}{N-2}}{E} + 2b\frac{(E-1)}{E}(\frac{1}{N}+\frac{1}{N-2})\tilde{\gamma}_1 + 2b\frac{(E-1)}{E}(\frac{N-1}{N})(1-\tilde{\gamma}_2)}.$$

Thus in the limit as $E \to \infty$,

$$Ew_E \to \frac{1}{2b}\frac{1}{\frac{N-1}{N}\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}}\frac{\sigma_\pi^2}{\sigma_\epsilon^2+\sigma_\pi^2} = \frac{1}{2b}\frac{N}{N-1}\frac{\sigma_\pi^2}{\sigma_\epsilon^2}.$$

Note that this implies that for large enough E, any real root of the cubic equation for w_E must be positive, which by (77) implies that ζ_E is positive for any real root. An argument analogous to Theorem 4 can then be used to verify that there is a symmetric affine equilibrium corresponding to any positive root of the cubic equation for w_E. Since a cubic equation always has at least one real root and at most three, there always exists at least one and at most three symmetric affine equilibrium for E sufficiently large.

Part 5. Using earlier results we can write

$$E\alpha_E = \frac{2bE}{2b + \frac{(2b[(E-1)\tilde{\gamma}_1+1]+\frac{N}{N-1}\tilde{\gamma}_3)}{N-2} + 2b(E-1)\frac{1}{N}\tilde{\gamma}_1 + \tilde{\gamma}_3\frac{1}{(N-1)} + 2b(E-1)(1-\frac{\tilde{\gamma}_1}{N})}.$$

Thus, if $\sigma_Q^2 > 0$, as $E \to \infty$,

$$E\alpha_E \to 1.$$

When $E = 1$,

$$\alpha_1 = \frac{2b}{2b + \frac{2b+\frac{\tilde{\gamma}_3}{w_1}}{N-2} + \frac{\tilde{\gamma}_3}{w_1(N-1)}} < 1.$$

Thus, an increase in fragmentation means a more efficient redistribution of endowments, at least in the limit. $\square$

Next, we give a coarse analysis of welfare which compares the expected holding costs of strategic agents as E tends infinity with the case of centralized exchange when $E = 1$.

Proposition 9. *If $\frac{\sigma_\pi^4}{\sigma_\epsilon^2}$ is sufficiently small, then for all E sufficiently large the allocation of any symmetric affine equilibrium is more efficient than the allocation of the unique symmetric affine equilibrium when E is 1.*

Proof. By symmetry it suffices to study the expected holding cost of an individual agent. Recall, in what follows, that we have assumed for simplicity that the mean of the liuqidity trader supply is zero. The expected holding cost of an agent is

$$\mathbb{E}\left[b\left((1 - E\alpha_E)X_i + E\alpha_E\frac{Z}{N} + Ew_E(S_i - \frac{1}{N}\sum_{j\in N} S_j) + \frac{\sum_{e\in E} Q_e}{N}\right)^2\right] =$$
$$b\big((1 - E\alpha_E)X_i + E\alpha_E\frac{Z}{N})^2 + (Ew_E)^2\big((\frac{N-1}{N})^2 + \frac{N-1}{N^2})\sigma_\epsilon^2 + \frac{\sigma_Q^2}{N^2}$$

Consider taking a limit as $E \to \infty$ of the above expression. Then we obtain

$$b\frac{Z^2}{N^2} + \frac{\sigma_Q^2}{N^2} + (\frac{1}{2b}\frac{N}{N-1})^2\frac{\sigma_\pi^4}{\sigma_\epsilon^2}((\frac{N-1}{N})^2 + \frac{N-1}{N^2})$$

The only difference between this expected holding cost and the expected holding cost at the efficient allocation is the last term. Thus when $\frac{\sigma_\pi^4}{\sigma_\epsilon^2}$ is small, a large level of fragmentation is preferred to centralized exchange. □

H Extension: Arbitrary Covariance Matrix

In this appendix, we extend the baseline model to allow for correlation among the primitive asset quantities $\{X_1, \ldots, X_N, Q_1, \ldots, Q_E\}$ setting the sizes of trading interests. This model variant nests the baseline model. Consequently, many of the proofs are quite similar.

H.1 Setup

We retain the same model setup as in the baseline but alter the assumptions about the joint distribution of $(X_1, \ldots, X_N, Q_1, \ldots, Q_E)$. We assume that $Q = C + \sum_{e\in E}\xi_e$ and $Q_e = \frac{C}{E} + \xi_e$ for each $e \in E$, where C and $\{\xi_e\}_{e\in E}$ are random variables in $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Here, C is the component of liquidity trader supply which is common across exchanges and ξ_e is the component idiosyncratic to exchange e. We assume that the distribution of C does not depend on E and that $\{\xi_e\}_{e\in E}$ is a collection of i.i.d, Gaussian distributed random variables with a mean of 0 and variance of $\frac{\sigma_\xi^2}{E}$ that are independent of $X_1, \ldots, X_N$, and C. Under these assumptions, the distribution of Q does not depend on E. Next, we assume that $X_1, \ldots, X_N, C$ are jointly Gaussian with $\mathbb{E}[C] = \mu_Q$, $\text{var}[C] = \rho$, $\text{cov}(X_i, X_j) = \Sigma$ for all $i, j \in N$ such that $i \neq j$, and $\text{cov}(X_i, C) = \eta$, $\mathbb{E}[X_i] = \mu_X$, and $\text{var}[X_i] = \sigma_X^2$ for all $i \in N$. For the distribution to be well defined, ρ, Σ, η, and σ_X^2 are such that the covariance matrix of $X_1, \ldots, X_N, C$ is positive definite.

H.2 Analysis

Lemma 8. The condition, $\sigma_X^2 + (N-1)\Sigma > 0$, holds.

Proof of Lemma 8. The covariance matrix of $(X_1, \ldots, X_N)$ is positive definite. Denote the covariance matrix V_X. Each element of the diagonal of V_X is σ_X^2 while all other elements are Σ. This implies that $\mathbf{1}^T V_X \mathbf{1} = N[\sigma_X^2 + (N-1)\Sigma] > 0$ where $\mathbf{1}$ is an $N \times 1$ vector of ones. □

Theorem 9. *For each $E \in \mathbb{N}$, there exists at least one and up to three symmetric affine equilibria. If either $\eta \geq 0$ or $\sigma_\xi^2 = 0$, there is a unique symmetric affine equilibrium. Given an arbitrary $E \in \mathbb{N}$ let $(\alpha_E, \zeta_E, \Delta_E)$ be an arbitrary corresponding symmetric affine equilibrium. Then α_E, ζ_E, and Δ_E satisfy equations (96), (97), and (98). Moreover:*

1. *For each $e \in E$,*
$$\Lambda_E = \frac{2b(1 + \gamma_E(E-1))}{N-2}$$
where
$$\gamma_E \equiv \text{corr}_{X_i}(p_e^*, p_k^*)$$
for $k \neq e$ such that $k \in E$.

2. *Price in exchange $e \in E$ is*
$$p_e^* = \frac{N-1}{N}\Lambda_E[\sum_{i \in N} -\alpha_E X_i - Q_e + N\Delta_E].$$

3. *The final asset position of trader $i \in N$ is*
$$(1 - E\alpha_E)X_i + E\alpha_E \frac{\sum_{i \in N} X_j}{N} + \frac{Q}{N}.$$

4. *If $\sigma_\xi^2 = 0$ or $E = 1$, for each $E \in \mathbb{N}$, the equilibrium allocation corresponds with that of the centralized benchmark.*

5. *If $\sigma_\xi^2 > 0$, given an arbitrary sequence of symmetric affine equilibria, $\{(\alpha_E, \zeta_E, \Delta_E)\}_{E\in\mathbb{N}}$, we have*

$$E\alpha_E \to \frac{N}{N-1}\frac{1+\frac{\eta}{N\sigma_X^2}}{1-\frac{\Sigma}{\sigma_X^2}}.$$

Proof of Theorem 9. The proof proceeds in 3 steps. In the step 1 we compute some relevant moments corresponding to a symmetric affine equilibrium, $(\alpha_E, \zeta_E, \Delta_E)$. In step 2, we substitute the derived moments from step 1 into the optimality condition for a traders' demand submission problem and match coefficients to derive a system of three equations for α_E, ζ_E, and Δ_E. In step 3 we prove existence and uniqueness of a symmetric affine equilibrium and parts 1 through 5.

Step 1: To begin we conjecture an arbitrary symmetric affine equilibrium $(\alpha_E, \zeta_E, \Delta_E)$ in which each trader submits a demand schedule of the form in (2) to each exchange e. For ease of notation define

$$q_{ie}^f := f_{ie}(X_i, p_e^f).$$

We compute the following unconditional moments.

$$\mathbb{E}[\frac{-\alpha_E(\sum_i X_i) + mN - Q_{e'}}{yN}] = \frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E N} \tag{82}$$

$$\mathbb{E}[\frac{\sum_{j\neq i} -\alpha_E X_j}{\zeta_E(N-1)} - \frac{Q_e}{\zeta_E(N-1)} + \frac{\Delta_E}{\zeta_E}] = \frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E(N-1)} \tag{83}$$

$$\text{var}[\sum_i X_i] = N\sigma_X^2 + 2\Sigma\sum_{i=1}^{N}(i-1) = N\sigma_X^2 + \Sigma(N-1)N \tag{84}$$

Using the above moments we can then compute the following moments, conditional on X_i, using the projection theorem.

$$\mathbb{E}[\frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_{e'}}{\zeta_E N} \mid X_i] =$$
$$\frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E N} + \frac{\frac{1}{\zeta_E N}(-\alpha_E(N-1)\Sigma - \alpha_E\sigma_X^2 - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) \tag{85}$$

$$\mathbb{E}[\frac{(\sum_{j\neq i} -\alpha_E X_j) - Q_e + \Delta_E(N-1)}{\zeta_E(N-1)} \,|\, X_i] =$$
$$\frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E(N-1)} + \frac{\frac{1}{\zeta_E(N-1)}(-\alpha_E\Sigma(N-1) - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) \quad (86)$$

$$\mathrm{var}[-\alpha_E(\sum_{j\neq i} X_j) + \Delta_E(N-1) - Q_{e'} \,|\, X_i] =$$
$$\alpha_E^2(N-1)\sigma_X^2 + \alpha_E^2\Sigma(N-2)(N-1) + \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} + \frac{2\eta\alpha_E(N-1)}{E} - \frac{[(-\alpha_E\Sigma(N-1) - \frac{\eta}{E})]^2}{\sigma_X^2} \quad (87)$$

$$\mathrm{cov}_{X_i}(\sum_j -\alpha_E X_j - Q_{e'}, \sum_{j\neq i} -\alpha_E X_j - Q_e) =$$
$$\mathrm{var}[\sum_{j\neq i} -\alpha_E X_j \,|\, X_i] - 2\mathrm{cov}_{X_i}(Q_{e'}, \sum_{j\neq i} -\alpha_E X_j) + \mathrm{cov}_{X_i}(Q_{e'}, Q_e) \quad (88)$$

Using the above moments, we compute the following moments, conditional on X_i and $p_e^f - \Lambda q_{ie}^f$, (the portion of price in exchange e which is unknown to agent i—see equation (94)) by using the projection theorem. We have,

$$\mathbb{E}[p_{ie'}^f \,|\, p_e^f - \frac{q_{ie}^f}{y(N-1)}, X_i] =$$
$$(1 - \frac{N-1}{N}\gamma_E)\frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - (1-\gamma_E)\frac{\mu_Q}{E\zeta_E N} + (1-\gamma_E)\frac{\frac{1}{\zeta_E N}(-\alpha_E(N-1)\Sigma - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X)$$
$$+ \frac{-\alpha_E X_i}{\zeta_E N} + \frac{N-1}{N}\gamma_E p_e^f - \gamma_E\frac{q_{ie}^f}{\zeta_E N} \quad (89)$$

$$\mathbb{E}[q_{ie'}^f \,|\, p_e^f - \frac{q_{ie}^f}{\zeta_E(N-1)}, X_i] =$$
$$-\alpha_E X_i\frac{N-1}{N} - (1 - \frac{N-1}{N}\gamma_E)(-\alpha_E\mu_X + \Delta_E) + (1-\gamma_E)\frac{\mu_Q}{EN}$$
$$-(1-\gamma_E)\frac{\frac{1}{N}(-\alpha_E(N-1)\Sigma - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) - \frac{N-1}{N}\gamma_E\zeta_E p_e^f + \gamma_E\frac{q_{ie}^f}{N} + \Delta_E \quad (90)$$

Above, γ_E denotes

$$\frac{\mathrm{cov}_{X_i}(\sum_i -\alpha_E X_i - Q_e, \sum_{j\neq i} -\alpha_E X_j - Q_{e'})}{\mathrm{var}[\sum_{j\neq i} -\alpha_E X_j - Q_e \,|\, X_i]}. \tag{91}$$

Of course, $\mathbb{E}[q^f_{ie'} \,|\, p^f_e - \frac{q^f_{ie}}{\zeta(N-1)}, X_i]$ could have been computed in one step by just a single application of the projection theorem, but we found it less algebraicly taxing to apply the projection theorem twice. To finish deriving $\mathbb{E}[q^f_{ie'} \,|\, p^f_e - \frac{q^f_{ie}}{\zeta(N-1)}, X_i]$, we must compute an expression for γ_E. The denominator was computed earlier in equation (6). To compute the numerator, we make use of the decomposition in equation (88). The terms $\sum_{j\neq i} X_j$, Q'_e, Q_e, and X_i are jointly normally distributed with covariance matrix

$$\overline{\Sigma} = \begin{bmatrix} (N-1)\sigma_X^2 + \Sigma(N-2)(N-1) & \frac{\eta(N-1)}{E} & \frac{\eta(N-1)}{E} & \Sigma(N-1) \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} & \frac{\rho}{E^2} & \frac{\eta}{E} \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} & \frac{\eta}{E} \\ \Sigma(N-1) & \frac{\eta}{E} & \frac{\eta}{E} & \sigma_X^2 \end{bmatrix}.$$

The goal is to derive the covariance matrix of $\sum_{j\neq i} X_j$, $Q_{e'}$, Q_e conditional on X_i, which we denote $\tilde{\Sigma}$. To do this we can apply the projection theorem. Then

$$\tilde{\Sigma} =$$

$$\begin{bmatrix} (N-1)\sigma_X^2 + \Sigma(N-2)(N-1) & \frac{\eta(N-1)}{E} & \frac{\eta(N-1)}{E} \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} & \frac{\rho}{E^2} \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} \end{bmatrix} - \begin{bmatrix} \Sigma(N-1) \\ \frac{\eta}{E} \\ \frac{\eta}{E} \end{bmatrix} \frac{1}{\sigma_X^2} \begin{bmatrix} \Sigma(N-1) & \frac{\eta}{E} & \frac{\eta}{E} \end{bmatrix}$$

$$\Leftrightarrow$$

$$\tilde{\Sigma} =$$

$$\begin{bmatrix} (N-1)\sigma_X^2 + \Sigma(N-2)(N-1) & \frac{\eta(N-1)}{E} & \frac{\eta(N-1)}{E} \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} & \frac{\rho}{E^2} \\ \frac{\eta(N-1)}{E} & \frac{\rho}{E^2} & \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} \end{bmatrix} - \frac{1}{\sigma_X^2} \begin{bmatrix} \Sigma^2(N-1)^2 & \frac{\Sigma\eta(N-1)}{E} & \frac{\Sigma\eta(N-1)}{E} \\ \frac{\Sigma\eta(N-1)}{E} & \frac{\eta^2}{E^2} & \frac{\eta^2}{E^2} \\ \frac{\Sigma\eta(N-1)}{E} & \frac{\eta^2}{E^2} & \frac{\eta^2}{E^2} \end{bmatrix}$$

From above, we have

$$\mathrm{cov}_{X_i}(-\alpha_E X_i + \sum_{j\neq i} -\alpha_E X_j - Q_{e'}, \sum_{j\neq i} -\alpha_E X_j - Q_e)$$

$$= \alpha_E^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1) - \frac{\Sigma^2(N-1)^2}{\sigma_X^2}) + \frac{2\alpha_E\eta(N-1)}{E}(1 - \frac{\Sigma}{\sigma_X^2}) + \frac{\rho}{E^2} - \frac{\eta^2}{E^2\sigma_X^2}.$$

We finally derive that

$$\gamma_E = \frac{\alpha_E^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1) - \frac{\Sigma^2(N-1)^2}{\sigma_X^2}) + 2\alpha_E \frac{\eta}{E}(N-1)(1-\frac{\Sigma}{\sigma_X^2}) + \frac{\rho}{E^2} - \frac{\eta^2}{E^2\sigma_X^2}}{\alpha_E^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1)) + \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} + 2\frac{\eta}{E}\alpha_E(N-1) - \frac{\left(-\alpha_E\Sigma(N-1)-\frac{\eta}{E}\right)^2}{\sigma_X^2}}. \tag{92}$$

This concludes step one.

Step 2. By market clearing, we have

$$p_e^f = \frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_e}{\zeta_E N}. \tag{93}$$

Also by market clearing, the residual supply curve trader i faces in exchange e is

$$p_e(q) = \frac{-\alpha_E(\sum_{j\neq i} X_i) + q + \Delta_E(N-1) - Q_e}{\zeta_E(N-1)}. \tag{94}$$

This implies that the price impact agent i faces in exchange e is $\Lambda := \frac{1}{\zeta_E(N-1)}$, which by symmetry, is the price impact each agent i faces in all exchanges. In equilibrium trader i equates his expected marginal utility conditional on $p_e^f - \frac{q_{ie}^f}{\zeta_E(N-1)}$ and X_i, with his marginal cost. That is

$$\mu_\pi - 2b(X_i + q_{ie}^f + (E-1)\mathbb{E}[q_{i2}^f \,|\, p_e^f - \frac{q_{ie}^f}{\zeta_E(N-1)}, X_i]) = p_e^f + \Lambda q_{ie}^f. \tag{95}$$

Substiting equation (90) into (95) and matching coefficients we obtain a system of three equations which characterize the three unknowns, α_E, ζ_E, and Δ_E. We do not explicitly list the algebraic steps here. Matching the coefficients on price, we obtain

$$\zeta_E = \frac{1}{2b((E-1)\gamma_E + 1)}\frac{N-2}{N-1}. \tag{96}$$

Matching the coefficients on X_i we obtain

$$\alpha_E = \frac{1 + (E-1)\frac{(1-\gamma_E)\frac{\eta}{E}}{N\sigma_X^2}}{1 + \frac{\gamma_E(E-1)}{N} + \frac{(E-1)\gamma_E+1}{N-2} + (E-1)\frac{N-1}{N} - (1-\gamma_E)(E-1)\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}}. \tag{97}$$

Matching the constant coefficients, we obtain

$$\Delta_E = \frac{N-2}{N-1} \frac{\mu_\pi - 2b(E-1)\mu_X\left(\frac{(1-\gamma_E)\mu_Q}{EN} - \frac{(1-\gamma_E)\frac{1}{N}(\alpha_E(N-1)\Sigma + \frac{\eta}{E})}{\sigma_X^2} + \alpha_E(1 - \frac{N-1}{N}\gamma_E)\right)}{2b(1+\gamma_E(E-1))} \tag{98}$$

Above, γ_E, as we saw in equation (92) is dependent on α_E. By inspecting (97) and (92) we see that α_E satisfies a cubic equation. It is clear that a neccessary condition for $(\alpha_E, \zeta_E, \Delta_E)$ to be a symmetric affine equilibrium is that α_E, ζ_E, and Δ_E satisfy the above equations (since otherwise the distributional assumptions ensure that the condition (95) is violated on a set of strictly positive $\mathbb{P}$-measure). This concludes step 2.

Step 3. By Theorem 4, equations (97), (96), and (98) are necessary and sufficient conditions for $(\alpha_E, \zeta_E, \Delta_E)$ to be a symmetric affine equilibrium. To prove existence of at least one and up to three such symmetric affine equilibria, it suffices to observe from (97) and (92) that α_E satisfies a cubic equation which must have at least one real root and up to three real roots. We now prove uniqueness of the equilibrium when $\eta \geq 0$. Fix $E \geq 1$, denote $y \equiv \alpha_E$, and define

$$g(y) \equiv y - \frac{1 + \frac{E-1}{E}\frac{(1-\gamma_E)\eta}{N\sigma_X^2}}{E\gamma_E(\frac{1}{N} + \frac{1}{N-2} + \frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}) + (1-\gamma_E)(\frac{1}{N} + \frac{1}{N-2} + \frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}) + E\frac{N-1}{N}(1 - \frac{\Sigma}{\sigma_X^2})}.$$

There exists a symmetric affine equilibrium for each y positive such that $g(y) = 0$. Using the assumption that $\eta \geq 0$, the second term in the above expression is strictly monotone decreasing in γ_E. By multiplying the numerator and denominator in equation (92) by E^2 we see that γ_E is strictly monotone increasing in y. Thus $g(y)$ is strictly monotone increasing in y. Hence there can exist at most one value of $y \in \mathbb{R}$ such that $g(y) = 0$.

We now prove the remaining parts of the theorem. Part 1 follows immediately from (96). Part 2 follows immediately from (94). Part 3 of the theorem is true of any symmetric affine equilibrium independent of the joint distribution of the random variables and the proof is analogous to that of Theorem 1. Part 4 follows from part 3 and (97) when substituting in $\gamma_E = 1$ which is the value γ_E takes on when $\sigma_Q^2 = 0$. To prove part 5, observe that using Proposition 11, $\gamma_E \to 0$. By equation (97),

$$E\alpha_E = \frac{1 + \frac{(E-1)}{E}\frac{(1-\gamma_E)\eta}{N\sigma_X^2}}{\frac{1}{E} + \frac{\gamma_E(E-1)}{EN} + \frac{(E-1)\gamma_E+1}{E(N-2)} + (E-1)\frac{N-1}{EN} - (1-\gamma_E)(E-1)\frac{N-1}{EN}\frac{\Sigma}{\sigma_X^2}}.$$

Since $\gamma_E \to 0$, $E\alpha_E \to \frac{1+\frac{\eta}{N\sigma_X^2}}{\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})}$. □

Corollary 9.1. *Let $\{E\alpha_E\}_{E\in\mathbb{N}}$ be defined as in Theorem 9. Then $-El_E$ converges to a constant that exceeds 1 if and only if $\sigma_\xi^2 > 0$ and $\eta > -[\sigma_X^2+(N-1)\Sigma]$, where, by the positive definiteness of the covariance matrix of $X_1,\ldots,X_N$, we have $\sigma_X^2+(N-1)\Sigma \geq 0$. Further, $E\alpha_E$ converges to a constant that exceeds $\frac{N-1}{N-2}$ if and only if $\sigma_\xi^2 > 0$ and $\eta > -[\sigma_X^2+(N-1)\Sigma]$.*

Proof of corollary 9.1. Theorem 9 supplies a closed form expression for the limiting value of $E\alpha_E$ as $E \to \infty$. The rest of the proof is a simple computation.

□

Proposition 10. *Let*

$$E^* \equiv \frac{-\frac{(N-1)\sigma_X^2+\Sigma(N-2)(N-1)-\frac{\Sigma^2(N-1)^2}{\sigma_X^2}+2\eta(N-1)(1-\frac{\Sigma}{\sigma_X^2})+\rho-\frac{\eta^2}{\sigma_X^2}}{N-2} - \sigma_\xi^2(\frac{1}{N}+\frac{1}{N-2}+\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}) - \sigma_\xi^2\frac{\eta}{N\sigma_X^2}}{\sigma_\xi^2\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2}) - \sigma_\xi^2(1+\frac{\eta}{N\sigma_X^2})}.$$

If E^* is in $\mathbb{N}$, there is a unique symmetric affine equilibrium when $E = E^*$ whose allocation is the efficient allocation. If $\eta \geq 0$, by Theorem 9, there is a unique symmetric affine equilibrium allocation associated with each $E \in \mathbb{N}$. The $E \in \mathbb{N}$ whose symmetric affine equilibrium is most efficient (more efficient than that of any $E' \in \mathbb{N}$ with $E' \neq E$) is either $\lfloor E^* \rfloor$ or $\lceil E^* \rceil$.

Proof of proposition 10. Let $(\alpha_E, \zeta_E, \Delta_E)$ denote an arbitrary symmetric affine equilibrium. Define $g_E \equiv E\alpha_E$. Substituting equation (92) into (97) and rearranging yields a cubic equation in g_E with coefficients

$$[g_E^3] : A(1+\frac{1}{N-2})$$

$$[g_E^2] : B(1+\frac{1}{N-2}) - A$$

$$[g_E] : F(1+\frac{1}{N-2}) + \sigma_\xi^2(\frac{1}{N}+\frac{1}{N-2}+\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}) + \sigma_\xi^2 E\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2}) - B$$

$$[constant] : -F - E\sigma_\xi^2(1+\frac{\eta}{N\sigma_X^2}) + \sigma_\xi^2\frac{\eta}{N\sigma_X^2}$$

where

$$A \equiv ((N-1)\sigma_X^2 + \Sigma(N-2)(N-1) - \frac{\Sigma^2(N-1)^2}{\sigma_X^2}),$$

$$B \equiv 2\eta(N-1)(1-\frac{\Sigma}{\sigma_X^2})$$

and

$$F \equiv \rho - \frac{\eta^2}{\sigma_X^2}.$$

By definition, at E^*, $g_{E^*} = 1$. Therefore, we have

$$A(1+\frac{1}{N-2}) + B(1+\frac{1}{N-2}) - A + F(1+\frac{1}{N-2})$$

$$+\sigma_\xi^2(\frac{1}{N}+\frac{1}{N-2}+\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2})+\sigma_\xi^2 E^*\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})-B-F-E^*\sigma_\xi^2(1+\frac{\eta}{N\sigma_X^2})+\sigma_\xi^2\frac{\eta}{N\sigma_X^2}=0.$$

Solving for E^* we obtain,

$$E^* = \frac{-\frac{A+B+F}{N-2} - \sigma_\xi^2(\frac{1}{N}+\frac{1}{N-2}\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}) - \sigma_\xi^2\frac{\eta}{N\sigma_X^2}}{\sigma_\xi^2\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2}) - \sigma_\xi^2(1+\frac{\eta}{N\sigma_X^2})}.$$

That the $E \in \mathbb{N}$ whose symmetric affine equilibrium allocation is most efficient is either $\lfloor E^* \rfloor$ or $\lceil E^* \rceil$ when $\eta \geq 0$ follows from proposition 14. $\square$

Proposition 11. *For each $E \in \mathbb{N}$ denote an arbitrary corresponding symmetric affine equilibria, $\{(l_E, \zeta_E, \Delta_E)\}_{e\in E}$. Let Λ_E be the corresponding equilibrium price impact and γ_E the equilibrium inference coefficient. Then, if $\sigma_\xi^2 = 0$, $\{\Lambda_E\}_{E\in\mathbb{N}}$ diverges to ∞ and $\{\gamma_E\}_{E\in\mathbb{N}}$ is the constant sequence of ones. If $\sigma_\xi^2 > 0$, $\{\Lambda_E\}_{E\in\mathbb{N}}$ converges to*

$$\frac{2b+\frac{1}{\sigma_\xi^2}[(\frac{1+\frac{\eta}{N\sigma_X^2}}{\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})})^2((N-1)\sigma_X^2+\Sigma(N-2)(N-1)-\frac{\Sigma^2(N-1)^2}{\sigma_X^2})+2N(1+\frac{\eta}{N\sigma_X^2})\eta+\rho-\frac{\eta^2}{\sigma_X^2}]}{N-2}$$

while $\{\gamma_E\}_{E\in\mathbb{N}}$ converges to 0.

Proof of propostion 11. The claims when $\sigma_\xi^2 = 0$ are obvious in light of Theorem 9. We prove the claims when $\sigma_\xi^2 > 0$. By inspecting equation (97), and recognizing that Lemma 8 implies that $\frac{1}{N}+\frac{1}{N-2}+\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2} > 0$, we see that

$$\frac{|\,1+\frac{E-1}{E}\frac{(1-\gamma_E)\eta}{N\sigma_X^2}\,|}{E(\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})+\frac{1}{N}+\frac{1}{N-2}+\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2})} < |\,\alpha_E\,| < \frac{1+\frac{E-1}{E}\frac{(1-\gamma_E)\,|\,\eta\,|}{N\sigma_X^2}}{E\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})}.$$

Inspecting the equation (92), we see that for large E, the numerator of γ_E is $O(\frac{1}{E^2})$ while the denominator, because of the $\frac{\sigma_\xi^2}{E}$ term, is $\omega(\frac{1}{E^2})$ so that $\gamma_E \to 0$. To prove that Λ_E converges to a positive constant, we can express $E\gamma_E$ as

$$E\frac{E^2\alpha_E^2((N-2)\sigma_X^2+\Sigma(N-2)(N-1)-\frac{\Sigma^2(N-1)^2}{\sigma_X^2})+2E\alpha_E\eta(N-1)(1-\frac{\Sigma}{\sigma_X^2})+\rho-\frac{\eta^2}{\sigma_X^2}}{E^2\alpha_E^2+\Sigma(N-2)(N-1)-\frac{\Sigma^2(N-1)^2}{\sigma_X^2})+2E\alpha_E\eta(N-1)(1-\frac{\Sigma}{\sigma_X^2})+\rho-\frac{\eta^2}{\sigma_X^2}+E\sigma_\xi^2}.$$

By Theorem 9, $-El_E$ converges so by inspection it is clear that $E\gamma_E$ must converge. Since both $E-1$ and γ_E are always weakly positive, and $\Lambda_E = \frac{2b(1+\gamma_E(E-1))}{N-2}$, it must converge to a strictly positive constant. We can directly compute this constant using part 5 of Theorem 9 to be:

$$\frac{2b + \frac{1}{\sigma_\xi^2}[(\frac{1+\frac{\eta}{N\sigma_X^2}}{\frac{N-1}{N}(1-\frac{\Sigma}{\sigma_X^2})})^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1) - \frac{\Sigma^2(N-1)^2}{\sigma_X^2}) + 2N(1+\frac{\eta}{N\sigma_X^2})\eta + \rho - \frac{\eta^2}{\sigma_X^2}]}{N-2}.$$

□

Proposition 12. *Suppose $\eta \geq 0$. For each $E \in \mathbb{N}$, let Λ_E denote the equilibrium price impact in the unique symmetric affine equilibrium. The sequence, $\{-\Lambda_E\}_{E\in\mathbb{N}}$, is strictly monotone increasing.*

Proof of proposition 12. The proof is analogous to that of Proposition 1 so we omit it. □

Proposition 13. *The total expected payment of liquidity traders is*

$$\frac{N-1}{N}\Lambda_E(-\mu_Q N\Delta_E + \sigma_\xi^2 + \frac{\rho+\mu_Q^2}{E} - \alpha_E N(\eta + \mu_X\mu_Q)).$$

Proof of propostion 13. We compute

$$-\mathbb{E}[\sum_{e\in E} p_e^* Q_e] = -\frac{N-1}{N}\Lambda_E \mathbb{E}[\sum_{e\in E}(\sum_{i\in N} -\alpha_E X_i + N\Delta_E - Q_e)Q_e]$$

$$= \frac{N-1}{N}\Lambda_E(-\mu_Q N\Delta_E + \sigma_\xi^2 + \frac{\rho+\mu_Q^2}{E} + \alpha_E N(\eta + \mu_X\mu_Q)).$$

□

Proposition 14. *Suppose $\sigma_\xi^2 > 0$ and $\eta \geq 0$. For each, $E \in \mathbb{N}$, denote the unique symmetric affine equilibrium, $(\alpha_E, \zeta_E, \Delta_E)$. The sequence, $\{E\alpha_E\}_{E\in\mathbb{N}}$, is strictly monotone increasing.*

Proof of proposition 14. The proof is analogous to that of part 6 of Theorem 1. □

References

Aitken, M., H. Chen, and S. Foley. 2017. The impact of fragmentation, exchange fees and liquidity provision on market quality. *Journal of Empirical Finance* 41:140–60.

Antill, S., and D. Duffie. 2019. Augmenting markets with mechanisms. Working Paper, National Bureau of Economic Research.

Baldauf, M., and J. Mollner. 2019. High-frequency trading and market performance. *Journal of Finance, forthcoming. Available at SSRN 2674767* .

———. 2020. Trading in fragmented markets. *Journal of Financial and Quantitative Analysis* forthcoming.

Bennett, P., and L. Wei. 2006. Market structure, fragmentation, and market quality. *Journal of Financial Markets* 9:49–78.

Bernales, A., N. Garrido, S. Sagade, M. Valenzuela, and C. Westheide. 2018. A tale of one exchange and two order books: Effects of fragmentation in the absence of competition. SAFE Working paper 234.

Boehmer, B., and E. Boehmer. 2003. Trading your neighbor's ETFs: Competition or fragmentation? *Journal of Banking & Finance* 27:1667–703.

Brown, D., and S. Ross. 1991. Spanning, valuation and options. *Economic Theory* 1:3–12.

Budish, E., P. Cramton, and J. Shim. 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics* 130:1547–621.

Budish, E., R. Lee, and J. Shim. 2019. Will the market fix the market? A theory of stock exchange competition and innovation. Working paper, University of Chicago.

Chakravarty, S., P. Jain, J. Upson, and R. Wood. 2012. Clean sweep: Informed trading through intermarket sweep orders. *Journal of Financial and Quantitative Analysis* 47:415–35.

Chao, Y., C. Yao, and M. Ye. 2018. Why discrete price fragments u.s. stock exchanges and disperses their fee structures. *Review of Financial Studies* 32:1068–101.

Chlistalla, M., and M. Lutat. 2011. Competition in securities markets: The impact on liquidity. *Financial Markets and Portfolio Management* 25:149–72.

Chowdhry, B., and V. Nanda. 1991. Multimarket trading and market liquidity. *The Review of Financial Studies* 4:483–511.

Chung, K. H., and C. Chuwonganant. 2012. Regulation NMS and market quality. *Financial Management* 41:285–317.

Clapham, B., P. Gomber, J. Lausen, and S. Panz. 2019. Liquidity provider incentives in fragmented securities markets. SAFE Working Paper No. 231.

Colliard, J.-E., and T. Foucault. 2012. Trading fees and efficiency in limit order markets. *The Review of Financial Studies* 25:3389–421.

De Fontnouvelle, P., R. P. Fishe, and J. H. Harris. 2003. The behavior of bid-ask spreads and volume in options markets during the competition for listings in 1999. *The Journal of Finance* 58:2437–63.

Degryse, H., F. De Jong, and V. van Kervel. 2015. The impact of dark trading and visible fragmentation on market quality. *Review of Finance* 19:1587–622.

Du, S., and H. Zhu. 2017. What is the optimal trading frequency in financial markets? *Review of Economic Studies* 84:1606–51.

Duffie, D., and H. Zhu. 2017. Size discovery. *The Review of Financial Studies* 30:1095–150.

Félez-Viñas, E. 2017. Market fragmentation, mini flash crashes and liquidity. Working paper, Stockholm Business School.

Foley, S., E. Jarnecic, and A. Liu. 2020. Forming an orderly line — how queue-jumping drives fragmentation. Working paper, Macquarie University.

Foucault, T., and A. J. Menkveld. 2008. Competition for order flow and smart order routing systems. *The Journal of Finance* 63:119–58.

Gentile, M., and S. Fioravanti. 2011. The impact of market fragmentation on European stock exchanges. Working paper, Consob, Research Division.

Glosten, L. 1994. Is the electronic open-limit order book inevitable? *Journal of Finance* 49:1127–61.

Gomber, P., S. Sagade, E. Theissen, M. C. Weber, and C. Westheide. 2016. Spoilt for choice: Order routing decisions in fragmented equity markets. Working paper, Goethe University, SAFE.

———. 2017. Competition between equity markets: A review of the consolidation versus fragmentation debate. *Journal of Economic Surveys* 31:792–814. doi:10.1111/joes.12176.

Gresse, C. 2017. Effects of lit and dark market fragmentation on liquidity. *Journal of Financial Markets* 35:1–20.

Gresse, C., et al. 2012. Market fragmentation in europe: Assessment and prospects for market quality. Working Paper, The Future of Computer Trading in Financial Markets: Foresight Driver Review 19, Government Office for Science, United Kingdom.

Hall, G., and J. Rust. 2003. Middlemen versus market makers: A theory of competitive exchange. *Journal of Political Economy* 111:353–403.

Hamilton, J. L. 1979. Marketplace fragmentation, competition, and the efficiency of the stock exchange. *The Journal of Finance* 34:171–87.

Haslag, P. H., and M. Ringgenberg. 2016. The causal impact of market fragmentation on liquidity. In *Fourth Annual Conference on Financial Market Regulation*.

Hengelbrock, J., and E. Theissen. 2009. Fourteen at one blow: The market entry of turquoise. Working paper, University of Bonn.

Klemperer, P. D., and M. A. Meyer. 1989. Supply function equilibria in oligopoly under uncertainty. *Econometrica* 57:1243–77.

Körber, L., O. B. Linton, and M. Vogt. 2013. The effect of fragmentation in trading on market quality in the uk equity market. *Available at SSRN 2315484* .

Malamud, S., and M. Rostek. 2017. Decentralized exchange. *American Economic Review* 107:3320–62.

Malinova, K., and A. Park. 2019. Sniping in fragmented markets. Working paper, University of Toronto.

Mendelson, H. 1987. Consolidation, fragmentation, and market performance. *Journal of Financial and Quantitative Analysis* 22:189–207.

Menkveld, A. 2008. Splitting orders in overlapping markets: A study of cross-listed stocks. *Journal of Financial Intermediation* 17:145–74.

O'Hara, M., and M. Ye. 2011. Is market fragmentation harming market quality? *Journal of Financial Economics* 100:459–74.

Pagano, M. 1989. Endogenous market thinness and stock price volatility. *The Review of Economic Studies* 56:269–87.

Pagnotta, E., and T. Philippon. 2018. Competing on speed. *Econometrica* 86:1067–115.

Parlour, C. A., and D. J. Seppi. 2003. Liquidity-based competition for order flow. *The Review of Financial Studies* 16:301–43.

Rostek, M., and M. Weretka. 2012. Price inference in small markets. *Econometrica* 80:687–711.

———. 2015. Dynamic thin markets. *Review of Financial Studies* 28:2946–92.

Rostek, M., and J. H. Yoon. 2020. Exchange design and efficiency. Working paper, University of Wisconsin.

Sannikov, Y., and A. Skrzypacz. 2016. Dynamic trading: Price inertia and front-running. Working paper, Graduate School of Business, Stanford University. https://economics.uchicago.edu/sites/economics.uchicago.edu/files/uploads/PDF/sannikov_dynamic_trading.pdf.

Spankowski, U., M. Wagener, and H.-P. Burghof. 2012. The role of traditional exchanges in fragmented markets. Working paper, University of Hohenheim.

Stoll, H. R. 2001. Market fragmentation. *Financial Analysts Journal* 57:16–20.

Vayanos, D. 1999. Strategic trading and welfare in a dynamic market. *The Review of Economic Studies* 66:219–54.

Vives, X. 2011. Strategic supply function competition with private information. *Econometrica* 79:1919–66.

Wilson, R. 1979. Auctions of shares. *Quarterly Journal of Economics* 93:675–89.

Wittwer, M. 2020. Connecting disconnected financial markets? *American Economic Journal: Microeconomics* forthcoming.

Zhu, H. 2014. Do dark pools harm price discovery? *Review of Financial Studies* 27:747–89.