Market Fragmentation*

Daniel Chen† and Darrell Duffie‡
Graduate School of Business, Stanford University

February 19, 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 Mohammad Akbarpour, Bob Anderson, Anirudha Balusubramanian, Eric Budish, Ananth Madhavan, Albert Menkveld, Michael Ostrovsky, Marcos Salgado, Yuliy Sannikov, Yazid Sharaiha, Andy Skrzypacz, Bob Wilson, Milena Wittwer, and Haoxiang Zhu.

†Economic Analysis and Policy, Stanford Graduate School of Business

‡Stanford Graduate School of Business. Duffie is a Research Associate of the National Bureau of Economic Research and a director of Dimensional Funds.

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). Market fragmentation 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 causes trade to become *too* aggressive, from a welfare perspective. However, in our model setting, any degree of fragmentation is welfare-superior to a centralized market.

Our simple model abstracts from some important aspects of functioning financial markets. We do not consider the impact of fragmentation on exchange competition or transaction fees.[1] We also abstract from trader inferences related to cross-exchange cross-time order submission and the associated 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 in a single period by N strategic traders participating on E exchanges. Prior to trade, strategic trader i is endowed with a quantity X_i 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

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

exchange cannot be made contingent on clearing prices at other exchanges. The objective of each strategic trader, given the conjectured order submission strategies of the other traders, is to maximize the total expected cash compensation received for executed orders, net of a holding cost that is quadratic in the trader's final 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. We also consider a version of the model with no liquidity traders, and a version in which liquidity traders are replaced by a "competitive fringe" of traders that 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 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 more aggressive. This natural implication of fragmentation is novel to this paper, as far as we know. 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 endowment of strategic traders, the key "state variable" of our model, yet the exchange prices taken together are more informative.

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 7 summarizes the results of various model extensions. Section 8 offers some concluding remarks, including some 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.

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.

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

Other research, however, suggests less beneficial effects of fragmentation. For example, Bennett and Wei (2006) find that when equity trading migratesd 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 split orders between exchanges and size-discovery venues such as dark pools.

Methodologically, our model relates to the literature on multi-auction demand submission games (Wilson, 1979; Klemperer and Meyer, 1989; Malamud and Rostek, 2017; Wittwer, 2020). Within this literature, our model is closest to that of Wittwer (2020), which studies a demand-function submission game that is based on two exchanges, and which examines the welfare implications of connecting the two 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), we focus on properties of the equilibrium as the number of exchanges is increased.

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 Setup

This section presents the setup of the most basic version of our 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.

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

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

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, which is a finite-variance random variable.

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 given by a finite variance random variable Q_e.

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

[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.

[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.

associated with holding a post-trade position in the asset. Preferences of this form, although they have not been micro-founded, are popular in the market microstructure literature, including Vives (2011), Rostek and Weretka (2012), Du and Zhu (2017), and Sannikov and Skrzypacz (2016).

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 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 Simple Equilibrium

We assume that the $\{X_i \,|\, i \in N\}$ are *iid* and normally distributed with mean μ_X and variance σ_X^2. We also assume that the $\{Q_e \,|\, e \in E\}$ are *iid* normal with mean μ_Q/E and variance σ_Q^2/E. Finally, we assume that $\{X_i \,|\, i \in N\}$, $\{Q_e \,|\, e \in E\}$, and Z are independent. We relax these distributional assumptions in Section 6 and in extensions considered in the Appendix. A useful interpretation of the 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.

Under the above assumptions, following the approach of Du and Zhu (2017), 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

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

Generically in the parameters of the model, the equilibrium allocation is inefficient. Given the non-discretionary liquidation 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

$$\bar{q} = \frac{1}{N}\left(Q_e + \sum_i X_i\right).$$

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, 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 equilibrium in symmetric affine demand functions. The associated demand-function coefficients $(\Delta_E, \alpha_E, \zeta_E)$ form the unique solution to appendix equations (21), (22), and (23). 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*

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

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 $b = 1/2$ and $E = 100$ exchanges. Without liquidity trade, the price impact is $\Lambda_E = 33$. However, with $\sigma_Q^2 > 0$, and just $N = 5$ 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.

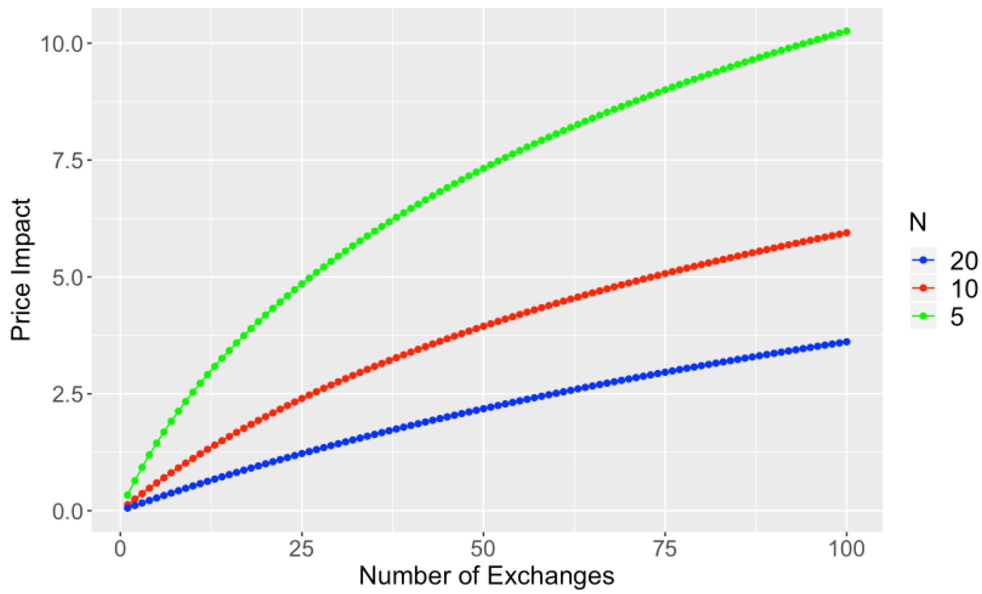

Figure 1: Variation of price impact Λ_E with the number E of exchanges, for various cases of N, the number of strategic traders. In all cases, the variance-aversion coefficient is $b = 1/2$ and a ratio σ_X^2/σ_Q^2 of strategic-trader asset endowment to total liquidity trade quantity of 10.

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 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 intuition of 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 $\frac{\sigma_X^2}{\sigma_Q^2}$.

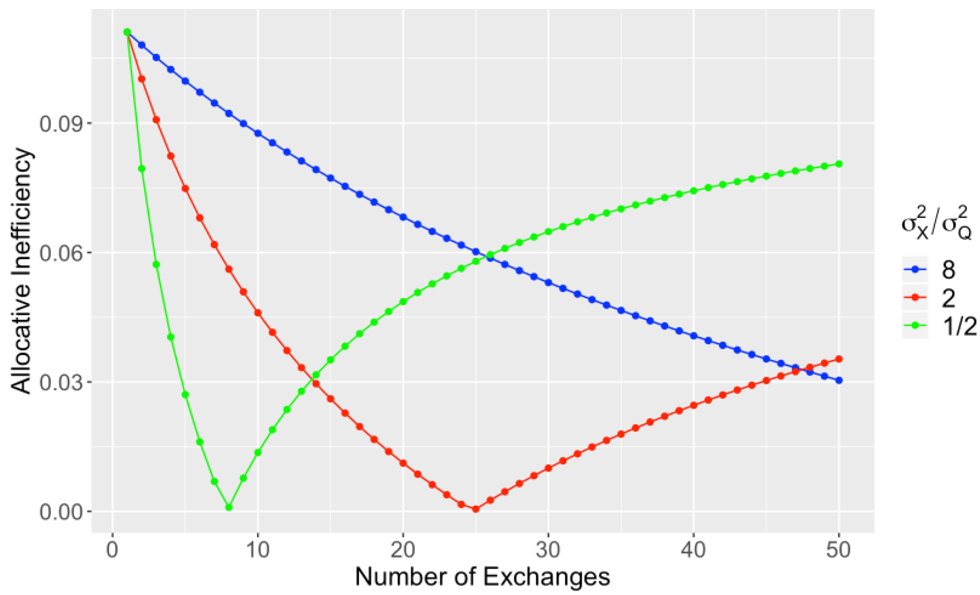

Figure 2: We plot equilibrium allocative inefficiency as measured by $|1 - E\alpha_E|$ against the number of exchanges for different values of the ratio σ_X^2/σ_Q^2 of the variance of the endowment of a strategic trader to the variance of the total amount of liquidity trade. In all cases, the number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

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.

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

2. $\mathrm{var}(\bar{X} \mid \{p_e^* : e \in E\})$ *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 this setting, because the equilibrium price of a given exchange is a linear combination of the aggregate endowment and of exchange-specific liquidity trade, traders do not need to make cross-exchange price inferences. That is, conditional on $\overline{X}$, prices in any two exchanges are uncorrelated. With no cross-exchange price inference, we can demonstrate the welfare benefits of fragmentation in a setting that does not require liquidity traders or Gaussian X_e and Q_e.

To this end, 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, (b) Q_e has zero mean, and (c) 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 $\mathbb{R}$, 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.*

[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 $f_{ie}(X_i, \sum_{j \in N} X_j, p)$.

2. *The price on exchange e is*

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

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

$$\frac{\Lambda_E}{\Lambda_E + 2bE}X_i + \frac{2bE}{\Lambda_E + 2bE}\bar{X} + \frac{Q}{N}.$$

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*

$$\frac{\mathrm{var}(Q)}{N-2}\frac{N-1}{N}.$$

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. Thus, 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 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.

[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.

7 Discussion of Model Extensions

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

7.1 Endogenous Liquidity Trade, Exchange by Exchange

In our first model extension, found in Appendix D, 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 that, provided c is sufficiently high, fragmentation always improves allocative efficiency, relative to a centralized market, in that the expected sum of all agents' holding costs is lower with $E > 1$ than with $E = 1$.

7.2 Private information about asset payoff

In a second extension, found in Appendix E, 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 so 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's on agents' payoff-relevant private information. Whether fragmentation is beneficial or harmful is shown to depend on the relative magnitudes of these two effects.

7.3 Correlated trade motives

In a third extension, found in Appendix F, 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.[9]

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.

8 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 is not a sign of worsening overall liquidity or market efficiency. We characterize the welfare-optimal number of exchanges in closed form.

Our stylized model abstracts from many important practical considerations. We do not consider some of the direct frictional costs of trade and order splitting, including the effects of 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,

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

specialists, and market makers (Baldauf and Mollner, 2019; Chao, Yao, and Ye, 2018; Colliard and 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).

We also do not consider the endogenous entry of exchanges, a common theme in the literature, as reviewed by Pagnotta and Philippon (2018). We have also not captured the effect of high-frequency traders that 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,[10] 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).

[10] Babus and Parlatore (2017) focus on the role of divergent beliefs and the incentives of investors to trade on a venue mainly with dealers or on different venue facing other investors.

A Verification Theorem

In this section we prove a theorem which will be repeatedly used to verify that a candidate symmetric affine equilibrium is indeed an equilibrium. Before stating the theorem, we first clarify our notation. We denote a candidate symmetric affine equilibrium by a triple corresponding to the coefficients of strategic agents' demand schedules, (α, ζ, Δ), and the information set of agent i by $\mathcal{F}_i$. For example, in the baseline model, $\mathcal{F}_i$ is the σ-algebra generated by X_i while in the model of section 6, it is the σ-algebra jointly generated by X_i and $\sum_{j \in N} X_j$. Given a candidate equilibrium, for each $i \in N$ and $e \in E$, we let $q^c_{ie} : \Omega \times \mathbb{R} \to \mathbb{R}$ denote the corresponding measurable demand schedule, q^*_{ie} the random variable corresponding to the quantity purchased, and p^c_e the random variable corresponding to the market clearing price.

Theorem 3. *Let (α, ζ, Δ) be a candidate symmetric affine equilibrium with $y > 0$ such that α, ζ, and Δ are each $\cap_{i \in N} \mathcal{F}_i$-measurable random variables. For each $e \in E$ and $i \in N$ define*

$$r^i_e := \sum_{\{j \in N \,|\, j \neq i\}} -\alpha X_j + (N-1)\Delta - Q_e.$$

If for each $i \in N$ and $e \in E$, r^i_e has finite variance and

$$\mu_\pi - \mathbb{E}[2b(X_i + q^c_{ie}(\cdot, p^c_e) + \sum_{\{e' \in E \,|\, e' \neq e\}} q^*_{ie'}) \,|\, \mathcal{F}_i, r^i_e] = p^c_e + \Lambda q^c_{ie}(\cdot, p^c_e)$$

holds almost surely, then (α, ζ, Δ) is a symmetric affine equilibrium.

Proof. For each $i \in N$, let $\mathcal{M}_i$ be the set of all maps, f, from $\Omega \times \mathbb{R}$ into $\mathbb{R}$ such that for each $p \in \mathbb{R}$, $f(\cdot, p)$ is a $\mathcal{F}_i$-measurable map from Ω into $\mathbb{R}$. Let $\tilde{\mathcal{M}}_i$ be the subset of $\mathcal{M}_i$ of all maps f such that $f(\cdot, r^i_e(\cdot))$ is a finite variance random variable. Given a candidate symmetric affine equilibrium, (α, y, m) we have shown in the main body of the text that the market clearing price in e is

$$p_e = \frac{r^i_e + q_{ie}}{\zeta(N-1)} \tag{9}$$

if agent i purchases q_{ie} units and all other agents submit the candidate equilibrium demand schedules. It is clear that any demand schedule submitted by agent i to exchange e which conditions on the market clearing price can be implemented by a demand schedule which instead conditions on r^i_e. It is therefore convenient to reformulate agent i's optimization problem as maximizing

$$\mathbb{E}[\pi \sum_{e\in E} q_{ie}(\omega, r_e^i) - b(X_i + \sum_{e\in E} q_{ie}(\omega, r_e^i))^2 - \sum_{e\in E} \frac{r_e^i + q_{ie}(\omega, r_e^i)}{\zeta(N-1)} q_{ie}(\omega, r_e^i)] \tag{10}$$

over $(q_{i1}, ..., q_{iE}) \in \mathcal{M}^E$. Above, $\omega \in \Omega$. We have for convenience suppressed the dependence of each r_e^i on ω. We will first derive a necessary and sufficient condition of the form given in the statement of the theorem for a profile of demand schedules to be optimal in $\tilde{\mathcal{M}}^E$ and then argue that restricting attention to $\tilde{\mathcal{M}}^E$ is without loss of generality.

Take arbitrary $(h_1, ..., h_E) \in \tilde{\mathcal{M}}^E$ and consider

$$\begin{aligned}\mathbb{E}&\left[\pi \sum_{e\in E}(q_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)) - b(X_i + \sum_{e\in E} q_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i))^2\right] \\ &- \mathbb{E}\left[\sum_{e\in E} \frac{r_e^i + q_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i)}{\zeta(N-1)}(q_{ie}(\omega, r_e^i) + \nu h_e(\omega, r_e^i))\right],\end{aligned} \tag{11}$$

which is a function of ν, an argument taking values in $\mathbb{R}$. Differentiating (11) with respect to ν and evaluating at $\nu = 0$ gives:

$$\begin{aligned}\mathbb{E}&\left[\pi \sum_{e\in E} h_e(\omega, r_e^i) - 2b(X_i + \sum_{e\in E} q_{ie}(\omega, r_e^i)) \sum_{e\in E} h_e(\omega, r_e^i)\right] \\ &- \mathbb{E}\left[\sum_{e\in E}(\frac{q_{ie}(\omega, r_e^i)}{\zeta(N-1)} + \frac{r_e^i + q_{ie}(\omega, r_e^i)}{\zeta(N-1)}) h_e(\omega, r_e^i)\right] = 0.\end{aligned}$$

This is a necessary condition for optimality. It holds if

$$\mathbb{E}\left[-2b(X_i + \sum_{e'\in E} q_{ie'}(\omega, r_{e'}^i)) \,|\, \mathcal{F}_i, r_e^i\right] = \frac{q_{ie}(\omega, r_e^i)}{\zeta(N-1)} + \frac{r_e^i + q_{ie}(\omega, r_e^i)}{\zeta(N-1)} - \mu_\pi. \tag{12}$$

for each $e \in E$. □

We now show that (12) is a sufficient condition for optimality within $\tilde{\mathcal{M}}^E$. Differentiating (11) with respect to ν twice we derive

$$\mathbb{E}\left[-2b(\sum_{e\in E} h_{ie}(\omega, r_e))^2 - \frac{2}{\zeta(N-1)} \sum_{e\in E} h_{ie}(\omega, r_e)^2\right], \tag{13}$$

which is less than or equal to 0 for all $(h_1, ..., h_E) \in \tilde{\mathcal{M}}^E$. Note that this expression does not depend on ν. Moreover, it is strictly less than 0 if one of $h_1, ..., h_N$ is not equal to 0 on a set of positive $\mathbb{P}$-measure. Suppose $(q_{i1}, ..., q_{iE})$ satisfies (12). Suppose for contradiction that there exists $(h_1^*, ..., h_E^*) \in \tilde{\mathcal{M}}^E$ which achieves a strictly higher value of (10) than does $(q_{i1}, ..., q_{iE})$. Set $(h_1, ..., h_E) \equiv (h_1^* - q_{i1}, ..., h_E^* - q_{iE}) \in \tilde{\mathcal{M}}^E$. Then the function (11) achieves a higher value at $\nu = 1$ than at $\nu = 0$. However, by earlier analysis, (11) is a strictly concave function of ν and has global maximum at $\nu = 0$. This is a contradiction.

We now argue that it is without loss of generality to restrict attention to $\tilde{\mathcal{M}}^E$. Consider, for arbitrary $e \in E$,

$$\begin{aligned}
\mathbb{E}\left[-\pi q_{ie}(\omega, r_e^i) + \frac{r_e^i q_{ie}(\omega, r_e^i) + q_{ie}(\omega, r_e^i)^2}{(N-1)\zeta}\right] = \\
\int_{\{\omega : |r_e| > \frac{1}{2}|q_{ie}(\omega, r_e^i)|\}} \left[-\pi q_{ie}(\omega, r_e^i) + \frac{r_e q_{ie}(\omega, r_e^i) + q_{ie}(\omega, r_e^i)^2}{\zeta(N-1)}\right] \mathbb{P}\{d\omega\} \\
+ \int_{\{\omega : |r_e| \le \frac{1}{2}|q_{ie}(\omega, r_e^i)|\}} \left[-\pi q_{ie}(\omega, r_e^i) + \frac{r_e q_{ie}(\omega, r_e^i) + q_{ie}(\omega, r_e^i)^2}{\zeta(N-1)}\right] \mathbb{P}\{d\omega\}. \quad (14)
\end{aligned}$$

Suppose that $q_{ie}(\cdot, r_e^i(\cdot))$ is not a finite variance random variable. The first integral is by construction finite since r_e^i is a finite variance random variable. The second integral by construction must exceed

$$\int_{\{\omega : |r_e| \le \frac{1}{2}|q_{ie}(\omega, r_e^i)|\}} \left[-\pi q_{ie}(\omega, r_e^i) + \frac{1}{2\zeta(N-1)} q_{ie}(\omega, r_e^i)^2\right] \mathbb{P}\{d\omega\}.$$

However, this second integral must be infinite since $q_{ie}(\cdot, r_e^i)$ is not a finite variance random variable (this can be seen by using the fact that π is assumed to be of finite variance and partitioning the integral to separately consider the cases when $|q_{ie}| > 3y(N-1)|\pi|$ and $|q_{ie}| \le 3y(N-1)|\pi|$). Thus, (14) is ∞ which, by inspecting (10), implies that the objective is $-\infty$. It is therefore without loss of generality to only consider optimality within $\tilde{\mathcal{M}}^E$.

Finally, putting together (12) with (9) we see that the condition in the statement of the theorem is a sufficient condition for each agent to optimize by submitting $\{q_{ie}^c\}_{e \in E}$ if each of the other agents act analogously.

B Proofs for Section 4

We give proofs for all results in section 4. First we formally define what it means for an allocation to be "more efficient" and then prove Proposition 4 which will be used in the proof of Theorem 1.

Definition 1. We say that an allocation, is *more efficient* than another allocation if the sum of strategic agents' holding costs is weakly lower at that allocation than at the other allocation for each $\omega \in \Omega$ and strictly lower almost surely. We say that two allocations are equally efficient if the sum of strategic agents' holding costs are equal at both allocations for each $\omega \in \Omega$.

Proposition 4. *Fix E and E' in $\mathbb{N}$ and let $(\alpha_E, \zeta_E, \Delta_E)$ and $(\alpha_{E'}, \zeta_{E'}, \Delta_{E'})$ be symmetric affine equilibria corresponding to E and E' respectively. If*

$$|\,1 - E\alpha_E\,| \;>\; |\,1 - E'\alpha_{E'}\,|$$

then the allocation corresponding to $(\alpha_{E'}, \zeta_{E'}, \Delta_{E'})$ is more efficient than the allocation corresponding to $(\alpha_E, \zeta_E, \Delta_E)$. The allocations corresponding to $(\alpha_E, \zeta_E, \Delta_E)$ and $(\alpha_{E'}, \zeta_{E'}, \Delta_{E'})$ are equally efficient if and only if

$$|\,1 - E\alpha_E\,| \;=\; |\,1 - E'\alpha_{E'}\,|\,.$$

Proof. The sum of strategic agents' holding costs at the equilibrium allocation corresponding to $(\alpha_E, \zeta_E, \Delta_E)$ is

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

Expanding, rearranging, and combining like terms we obtain

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

The result is an implication of the above expression and Jensen's inequality. $\square$

Proof of Theorem 1. The proof proceeds in three steps. In step one, we derive equations (21), (22), and (23) and show that a necessary and sufficient condition for a symmetric affine equilibrium is that α_E, ζ_E and Δ_E solve these three equations. In step two we prove that there is a unique solution to establish existencess and uniqueness of a symmetric affine equilibrium. This completes a proof of the preamble of the theorem. In the final third step, 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

$$q_{ie}^{eq} = -\alpha_E X_i - \zeta_E p_e + \Delta_E \tag{15}$$

to each $e \in E$ and $i \in N$, where α_E, ζ_E, and Δ_E are constants. Market clearing in exchange e implies that the equilibrium market clearing price is

$$p_e^* = \frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_e}{\zeta_E N}. \tag{16}$$

Price impact can also be determined from the market clearing condition. If agent i purchases q_{ie} units at the market clearing price when all other agents submit the conjectured equilibrium demand schedules, then

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

which implies that

$$p_e = \frac{-\alpha_E(\sum_{j \neq i} X_i) + q_{ie} + \Delta_E(N-1) - Q_e}{\zeta_E(N-1)}. \tag{17}$$

That is, 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 faces in all exchanges. In determining his or her optimal demand schedule for exchange e, agent i equates his or her expected marginal utility conditional on $p_e - \frac{q_{ie}}{\zeta_E(N-1)}$ and X_i, with his marginal cost. The optimality condition is

$$-2b\left(X_i + q_{ie} + (E-1)\mathbb{E}\left[q_{ik}^* \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i \right] \right) = p_e + \Lambda_E q_{ie} - \mu_\pi. \tag{18}$$

Here above, k is an arbitrary exchange in E, distinct from e and q_{ik}^* is the quantity agent i purchases in exchange k if all agents submit their equilibrium demand schedules to k. In (18), we have used symmetry of the exchanges. Our next step is to compute

$$\mathbb{E}\left[q_{ik}^* \;\middle|\; p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i \right]$$

using the projection theorem. We obtain

$$\begin{aligned} \mathbb{E}\left[q_{ik}^* \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i \right] = -\alpha_E X_i \frac{N-1}{N} - \left(1 - \frac{N-1}{N}\gamma\right)(-\alpha_E \mu_X + \Delta_E) \\ - \frac{N-1}{N}\gamma \zeta_E p_e + \gamma \frac{q_{ie}}{N} + \Delta_E, \end{aligned} \tag{19}$$

where

$$\gamma_E = corr_{X_i}(p_e^*, p_k^*) = \frac{E\alpha_E^2(N-1)\sigma_X^2}{E\alpha_E^2(N-1)\sigma_X^2 + \sigma_Q^2}. \tag{20}$$

We substitute the expression for γ_E and (15) into (18) and match coefficients to derive 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. We obtain

$$\zeta_E = \frac{1}{2b((E-1)\gamma_E + 1)} \frac{N-2}{N-1} \tag{21}$$

$$\alpha_E = \frac{1}{1 + \frac{\gamma(E-1)}{N} + \frac{(E-1)\gamma+1}{N-2} + (E-1)\frac{N-1}{N}}, \tag{22}$$

and

$$\Delta_E = -\frac{(E-1)(1 - \frac{N-1}{N}\gamma)\alpha_E\mu_X - \frac{\mu_\pi}{2b}}{1 + \frac{\gamma(E-1)}{N} + \frac{(\gamma(E-1)+1)}{N-2} + (E-1)(\frac{N-1}{N}\gamma)}. \tag{23}$$

It is clear that any symmetric affine equilibrium must solve equations (21), (22), and (23) since (18) must be satisfied for all but possibly a measure zero set of prices for optimality. We now argue that this is sufficient. Suppose there exists a solution $(\alpha_E, \zeta_E, \Delta_E)$ to the system of equations (21), (22), and (23). From (18), we see that this solution satisfies the conditions of Theorem 3 so it is indeed a symmetric affine equilibrium.

Step 2. We now prove existence of a symmetric affine equilibrium. It is straightforward to substitute (20) into (22) 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 (20), (21), and (23). Thus a symmetric affine equilibrium exists.

We now prove there is a unique symmetric affine equilibrium. 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 (20) 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 (20) 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 (16) and the fact that $\Lambda_E = \frac{1}{(N-1)\zeta_E}$. Part 2 follows immediately from (21). Part 3 follows by substituting (16) into (15). Part 4 can be seen from the fact that when $\sigma_Q^2 = 0$, γ_E is equal to 1 so that inspecting equations (21), (22), and (23) 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 (22), 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. We first prove that $E\alpha_E \to \frac{N}{N-1}$. 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 (22) with some rearrangement we write

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

Taking limits we see that $E\alpha_E \to \frac{N}{N-1}$. We now 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 (20), 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.

We now prove that the equilibrium allocation in a fragmented market ($E > 1$) is more efficient than the equilibrium allocation in a centralized market ($E = 1$). 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

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

The result follows from Proposition 4.

□

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 the remainder of this proof assume that $\sigma_Q^2 > 0$.

We now prove that Λ_E is strictly monotone increasing. By Theorem 1 we have $\Lambda_E = \frac{2b(1+\gamma_E(E-1))}{N-2}$. It therefore 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 $\gamma_{E+1} < \gamma_E$. 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

$$\frac{(E+1)num_{E+1}}{(E+1)num_{E+1} + \sigma_\xi^2} > \frac{Enum_E}{Enum_E + \sigma_\xi^2},$$

which in turn implies that

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

Thus we have shown that $(E-1)\gamma_E$ is strictly monotone increasing so that Λ_E is strictly monotone increasing.

We now 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

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

By Theorem 1, we have an explicit expression for the limit point of $E\alpha_E$ which implies also that $E\alpha_E^2 \to 0$. Thus, we can directly take limits of the right-hand side of the above equation to obtain $\Lambda_E \to \frac{2b}{N-2}(1+\frac{N^2\sigma_X^2}{(N-1)\sigma_Q^2})$.

We next prove that $\gamma_E \to 0$. By inspecting (22), we 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 the equation (20), we see that for large E, the numerator of γ_E is $O(\frac{1}{E})$ since by Theorem 1 we know that $E\alpha_E$ 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. By inspecting (20) it suffices to prove that $E\alpha_E^2$ is strictly monotone decreasing. Suppose for the moment that $(E-1)\gamma_E$ is constant. If we treat E as potentially taking non-integer values in $\mathbb{R}$ using (22) we can compute

$$\frac{d(E\alpha_E^2)}{dE} = \alpha_E^2 - 2\alpha_E\frac{d\alpha_E}{dE}E < 0$$

$$\Leftrightarrow$$

$$1 - 2\alpha_E \frac{N-1}{N} E < 0.$$

In the last inequality, we have used the fact that $-E\alpha_E = -\frac{N-2}{N-1}$ when $E = 1$ and is monotone decreasing by Theorem 1. The computation above is assuming $(E-1)\gamma_E$ is constant. We showed earlier that it is in fact increasing. But that it increases only serves to ensure that α_E decreases at a higher rate so it must be that $E\alpha_E^2$ is strictly monotone decreasing as aglklain seen by inspecting (22). Thus γ_E is strictly monotone decreasing. $\square$

Proof of proposition 2. Substituting (20) into (22) 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 (20) and (22) is such that $E\alpha_E$ is strictly monotone increasing. Combining this observation with Proposition 4 gives the result. $\square$

Proof of Proposition 3. We first prove Part 1. Recall that

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

We therefore have

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

and

$$cov(\sum_{i\in N} X_i, p_e^*) = \frac{N-1}{N}\Lambda_E\alpha_E^2 Var[\sum_{i\in N} X_i].$$

By the rules of conditional Gaussians

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

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

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

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

so that

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

Next we have,

$$cov(\sum_{i\in N} X_i, \sum_{e\in E} p_e^{*^E}) = \frac{N-1}{N}\Lambda_E E\alpha_E Var[\sum_{i\in N} X_i].$$

By the rules of conditional Gaussians,

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

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

Proposition 5. *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.*

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 proven in Proposition 1. □

C Proofs for Section 5

Proof of Theorem 2. The proof proceeds in three steps. In the first step we derive a candidate equilibrium. In the second step 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 the third step 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). We consider agent i's optimal choice of demand schedule to submit to exchange $e \in E$ assuming he submits the conjectured equilibrium demand schedule to all other exchanges and all other agents submit the conjectured equilibrium demand schedule to all exchanges. Under this conjecture, the market clearing condition in exchange e is

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

which implies that

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

if agent i purchases q_{ie} units on exchange e. The market clearing price in exchange $k \in E$ where $k \neq e$ is, by analogous steps,

$$p^*_k = \frac{(\sum_{j \in N} -\alpha X_j) + N\Delta - Q_k}{N\zeta}. \tag{24}$$

If agent i purchases q_{ie} units on exchange e he can infer the realization of Q_e. Agent i therefore seeks to maximize

$$\mathbb{E}[\pi q_{ie} + \pi \sum_{\{j \in N \,|\, j \neq i\}} q^*_{ij} - b(X_i + \sum_{\{j \in N \,|\, j \neq i\}} q^*_{ij} + q_{ie})^2 - p_e q_{ie} \,|\, Q_e, \mathcal{F}_i] - \mathbb{E}[\sum_{\{j \in N \,|\, j \neq i\}} p^*_j q^*_{ij} \,|\, \mathcal{F}_i]$$

by choosing q_{ie} for each realization of Q_e. Above q^*_{ij} is the equilibrium quantity agent i purchases on exchange j. The expectation on the right hand side does not condition on Q_e because it is is independent of q^*_{ij} and p^*_j for each $\{j \in N \,|\, j \neq i\}$. The first order condition is

$$-2b(X_i + q_{ie} + (E-1)\mathbb{E}[q^*_{ik} \,|\, \mathcal{F}_i]) = p_e + q_{ie}\Lambda - \mu_\pi \tag{25}$$

where $\Lambda \equiv \frac{1}{(N-1)y}$ and we have used symmetry. Rearranging, we have

$$q_{ie} = \frac{-2bX_i - 2b(E-1)\mathbb{E}[q_{ik}^* \,|\, \mathcal{F}_i] - p_e + \mu_\pi}{\Lambda + 2b}.$$

Substituting the equilibrium price p_k^* into the equilibrium demand schedule, we have

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

so that

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

We therefore have

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

We now match coefficients with our conjecture that $q_{ie}^{eq} = -\alpha X_i - yp_e + m$ to determine that

$$\zeta = \frac{N-2}{N-1}\frac{1}{2b}, \tag{26}$$

$$\Lambda = \frac{2b}{N-2}, \tag{27}$$

$$\alpha = \frac{2b}{\Lambda + 2bE}, \tag{28}$$

and

$$\Delta = \frac{-2b(E-1)[\frac{2b}{\Lambda+2bE}\frac{\sum_{i \in N} X_i}{N} + \frac{\mu_Q}{EN}] + \mu_\pi}{2b + \Lambda}. \tag{29}$$

Step 2. To complete step two we appeal to Theorem 3. By construction the condition in the theorem is satisfied (see (25)). 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 (26), (28), and (29) are not satisfied. Then equation (25) is violated for some realization of the price in some exchange $k \in E$ for some agent $j \in N$. Continuity implies that the condition (25) must be violated for realizations of p_e in an open neighborhood of positive Lebesgue measure, denoted $\mathcal{B}$. Note that it is a necessary condition that (25) holds almost surely. Since each Q_e has full support over the real line and is independent of $\mathcal{F}_i$, agent i can not be maximizing his expected utility by submitting the affine demand schedule, $-\alpha X_i - \zeta p_e + \Delta$ to exchange e.

Step 3. Part 1. was shown in equation (27). Part 2. follows from substituting equations (26), (28), and (29) into (24). Part 3 follows from substituting the equation for price in part 2. in to the equilibrium demand schedule and adding X_i. Part 4 follows from direct computation:

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

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

D Extension: Endogenous 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.

D.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

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

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 $\mu_\pi = 0$.

D.2 Analysis

Theorem 4. *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*

$$\gamma_E = \frac{(E\alpha)^2 \sigma_X^2 (N-1)}{(E\alpha)^2 \sigma_X^2 (N-1) + (\frac{c}{c+\Lambda_e \frac{N-1}{N}})^2 \sigma_H^2 E} \tag{30}$$

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

$$\Lambda_E = \frac{2b((E-1)\gamma_E + 1)}{N-2} \tag{31}$$

while the price impact of a liquidity trader is

$$\frac{N-1}{N}\Lambda_E. \tag{32}$$

4. *$E\alpha_E$ satisfies*

$$E\alpha_E = \frac{1}{\gamma_E(\frac{1}{N} + \frac{1}{N-2}) + (1-\gamma_E)\frac{1}{E}(\frac{1}{N} + \frac{1}{N-2}) + \frac{N-1}{N}}. \tag{33}$$

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 in the equilibrium is

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

if all agents $i \in$ 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 or she seeks to maximize

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

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 (30), (31) and (33) are satisfied. Then the analysis of the baseline model therefore ensures that provided there exists α_E and γ_E which satisfies (30), (31), and (33), 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 (31) and (33) into (30) 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 6. *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, the symmetric affine equilibrium allocation is more efficient than that of the symmetric affine equilibrium when there*

is a single exchange in the sense that the expected sum of all agents' holding costs is lower.

Proof. Centralized markets correspond to the case when γ_E is one and E is one. To prove Part 1, it is clear by inspecting (33) 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 the Part 2 recognize that, using part 1 of Theorem 4, 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 (30). 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 4, 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. □

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 7. *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 (33). Suppose $E\alpha_E$ is decreasing in E then it must be that γ_E is increasing. Consider now (30). 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. □

E 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.

E.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 that 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.

E.2 Analysis

Theorem 5. *In any symmetric affine equilibrium*

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

$$q_{ie} = -\alpha X_i - y p_e + w S_i + m.$$

where α, y, w, and m are defined by the system of equations (35)—(42).

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 (35) and (37).

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

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

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

$$q_{ie} = -\alpha X_i - y p_e + w S_i + m$$

to exchange $e \in E$ for each $i \in N$ and $e \in E$ where α, y, w, and m are each $\cap_{i \in N} \sigma(X_i, \sum_{j \in N} X_j, S_i)$-measurable random variables. If agent i purchases q_{ie} units at the market clearing price while all other individuals submit the conjectured demand schedule, then the market clearing price is

$$p_e = \frac{1}{(N-1)y}\left[\sum_{j \neq i}(-\alpha X_j + w S_j + m) - Q_e + q_{ie}\right].$$

The equilibrium price is

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

Let us study an agent i's optimal choice of demand schedule to submit to exchange 1. He or she must equate marginal utility with marginal cost for every realization of the price:

$$-2b(X_i + q_{i1} + (E-1)\mathbb{E}[q_{i2} \mid p_1 - \frac{1}{(N-1)y}q_{i1}, X_i, S_i]) = p_1 - \mathbb{E}[\pi \mid p_1 - \frac{1}{(N-1)y}q_{i1}, X_i, S_i] + \frac{1}{(N-1)y}q_{i1}. \quad (34)$$

The above condition is a necessary condition since it is easy to see that the second order condition is satisfied provided $y > 0$ (which is a condition a symmetric affine equilibrium must satisfy by definition). We now compute the two conditional moments $\mathbb{E}[q_{i2}\,|\,p_1 - \frac{1}{(N-1)y}q_{i1}, S_i, X_i]$ and $\mathbb{E}[\pi\,|\,p_1 - \frac{1}{(N-1)y}q_{i1}, S_i, X_i]$ by using the projection theorem. We begin with the former. We can, using the projection theorem, express

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

Here, γ_1 and γ_2 are derived as follows. The variables, $\sum_{j\neq i} S_j$,S_i, $p_1 - \frac{1}{(N-1)y}$ are jointly Gaussian with variance matrix

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

Define

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

with

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

Define

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

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

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

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

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

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

$$\mathbb{E}[\pi\,|\,p_1 - \frac{q_{i1}}{(N-1)y}, S_i] = \mu_\pi + \gamma_3(p_1 - \frac{q_{i1}}{(N-1)y} - \frac{w\mu_\pi}{y} + \alpha\frac{Z-X_i}{(N-1)y} - \frac{m}{y}) + \gamma_4(S_i - \mu_\pi).$$

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

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

Define

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

Denote

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

Then

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

We obtain,

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

and

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

Note that $\gamma_3 \frac{w}{y(N-1)} \in [0,1]$ 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 $\tilde{\gamma}_1 = \frac{w}{y(N-1)}\gamma_1$, $\tilde{\gamma}_2 = \frac{1}{N-1}\gamma_2$, $\tilde{\gamma}_3 \equiv \frac{w}{y(N-1)}\gamma_3$, and $\tilde{\gamma}_4 = \gamma_4$. Then

$$\tilde{\gamma}_1 = \frac{w^2\sigma_\pi^2\sigma_\epsilon^2 + \frac{w^2\sigma_\epsilon^2(\sigma_\pi^2+\sigma_\epsilon^2)}{N-1}}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)} \tag{35}$$

$$\tilde{\gamma}_2 = \frac{\frac{\sigma_\pi^2\sigma_Q^2}{E(N-1)^2}}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)}. \tag{36}$$

$$\tilde{\gamma}_3 = \frac{w^2\sigma_\pi^2\frac{\sigma_\epsilon^2}{N-1}}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)}. \tag{37}$$

$$\tilde{\gamma}_4 = \frac{\sigma_\pi^2[w^2(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{(N-1)}) + \frac{\sigma_Q^2}{(E(N-1)^2}] - w^2\sigma_\pi^4}{w^2[(\sigma_\pi^2+\sigma_\epsilon^2)(\sigma_\pi^2+\frac{\sigma_\epsilon^2}{N-1})-\sigma_\pi^4] + \frac{\sigma_Q^2}{E(N-1)^2}(\sigma_\pi^2+\sigma_\epsilon^2)}. \tag{38}$$

We can now use the equation (34) together with the conditional moments we just computed, to match coefficients and pin down α, y, w, and m. The coefficient of q_{i1} gathered on to the LHS is

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

The coefficient of p_1 gathered on to the RHS is

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

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 $\tilde{\gamma}_3$:

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

Price impact is therefore

$$\frac{1}{(N-1)y} = \frac{(2b[(E-1)\tilde{\gamma}_1+1] + N\frac{\tilde{\gamma}_3}{w})}{N-2}. \tag{39}$$

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:

$$\begin{aligned}&-2b(E-1)w\frac{N-1}{N}(1-\tilde{\gamma}_2) + \gamma_4 = \\ &\qquad w[2b + \frac{(2b[(E-1)\tilde{\gamma}_1+1] + \frac{N}{w}\tilde{\gamma}_3)}{N-2} + 2b(E-1)\frac{1}{N}\tilde{\gamma}_1 + \frac{\tilde{\gamma}_3}{w}]\end{aligned} \tag{40}$$

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

$$\alpha = \frac{2b}{2b + \frac{(2b[(E-1)\tilde{\gamma}_1+1]+N\frac{\tilde{\gamma}_3}{w})}{N-2} + 2b(E-1)\frac{1}{N}\tilde{\gamma}_1 + 2b(E-1)(1-\frac{\tilde{\gamma}_1}{N})}. \tag{41}$$

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

$$\begin{aligned}m = -[&\frac{2b(E-1)[\frac{\alpha Z}{N} - \frac{w}{N}(\mu_\pi(N-1) + \frac{\tilde{\gamma}_1\alpha Z}{w} - \tilde{\gamma}_2(N-1)\mu_\pi - \tilde{\gamma}_1(N-1)\mu_\pi)]}{2b + \frac{1}{(N-1)y} + \frac{2b(E-1)\tilde{\gamma}_1}{N} + \frac{\tilde{\gamma}_3}{w} + \frac{2b(E-1)\tilde{\gamma}_1(N-1)}{N} + \frac{\tilde{\gamma}_3(N-1)}{w}} \\ &+ \frac{-\mu_\pi + \tilde{\gamma}_3\mu_\pi(N-1) - \frac{\tilde{\gamma}_3\alpha Z}{w} + \tilde{\gamma}_4\mu_\pi}{2b + \frac{1}{(N-1)y} + \frac{2b(E-1)\tilde{\gamma}_1}{N} + \frac{\tilde{\gamma}_3}{w} + \frac{2b(E-1)\tilde{\gamma}_1(N-1)}{N} + \frac{\tilde{\gamma}_3(N-1)}{w}}]\end{aligned} \tag{42}$$

Thus equations (41), (39), (40), (42), (35), (36), (37), and (38) are necessary conditions that any symmetric affine equilibrium must satisfy. An argument analogous to that of Theorem 3 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 (39). This completes the proof of Parts 1 and 2. We omit the proof of Part 3 since it is a straightforward computation. $\square$

Proposition 8. *For any value of E, if there exists a symmetric affine equilibrium w must be positive.*

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

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

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. □

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 5, perfect allocative efficiency is acheived if $Ew = 0$ and $E\alpha = 1$.

Proposition 9. *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,

$$w_1 = \frac{\tilde{\gamma}_4 - (1 + \frac{N}{N-2})\tilde{\gamma}_3}{2b(1 + \frac{1}{N-2})}. \tag{43}$$

Rearranging (43), we derive

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

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

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

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

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

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

Since the coefficient of w^2 is positive, the coefficient of w^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 $pqr = -\frac{constant\ coefficient}{coefficient\ of\ w^3} > 0$. 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, $p + q + r = -\frac{coefficient\ of\ w^2}{coefficient\ of\ w^3} < 0$ 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, by (39), y is positive. A theorem analogous to Theorem 3 can then be used to verify that there is a symmetric affine equilibrium corresponding to this value of w. Since it is the unique positive real root, the equilibrium must be unique since (43) is a necessary condition which must be satisfied in any symmetric affine equilibrium.

Part 2. We rearrange (38) 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 (43) together with the above inequality gives the result.

Parts 3 and 4. Rearranging equation (40), 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 $|w|$ 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 must be positive, which by (39) implies that y is positive for any real root. An argument analogous to Theorem 3 can then be used to verify that there is a symmetric affine equilibrium corresponding to any positive root of the cubic equation for w. 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 $E = 1$,

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

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 10. *If $\frac{\sigma_\pi^4}{\sigma_\epsilon^2}$ is sufficiently small, then in the limit as $E \to \infty$, the allocation of any symmetric affine equilibrium is more efficient than the allocation of the unique symmetric affine equilibrium when E is one.*

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

$$\mathbb{E}[b((1 - E\alpha_E)X_i + E\alpha_E\frac{Z}{N} + Ew(S_i - \frac{1}{N}\sum_{j\in N} S_j) + \frac{\sum_{e\in E} Q_e}{N})^2] =$$

$$b(((1 - E\alpha_E)X_i + E\alpha_E\frac{Z}{N})^2 + (Ew)^2((\frac{N-1}{N})^2 + \frac{N-1}{N^2})\sigma_\epsilon^2 + \frac{\sigma_Q^2}{N^2}]$$

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 large amounts of fragmentation is preferred to centralized exchange. $\square$

F 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.

F.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 $\mathbb{E}[C] = \mu_Q$, $Var[C] = \rho$, $cov(X_i, X_j) = \Sigma$ for all $i, j \in N$ such that $i \neq j$, and $cov(X_i, C) = \eta$, $\mathbb{E}[X_i] = \mu_X$, and $Var[X_i] = \sigma_X^2$ 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.

G Analysis

Lemma 6. The condition, $\sigma_X^2 + (N-1)\Sigma > 0$, holds.

Proof of Lemma 6. 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 7. *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 (58), (59), and (60) given in the Appendix. Moreover:*

1. *For each $e \in E$,*

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

   where

$$\gamma_E \equiv corr_{X_i}(p_e^*, p_{e'}^*)$$

   for $e' \neq e$ s.t $e' \in E$.

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 inventory of agent $i \in N$ is*

$$X_i + \sum_{e \in E} q_{ie}^* = (1 - E\alpha_E)X_i + E\alpha_E \frac{\sum_{i \in N} X_j}{N} + \frac{Q}{N}.$$

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 7. We first conjecture that there exists a symmetric affine equilibrium $(\alpha_E,\ \zeta_E,\ \Delta_E)$ and derive conditions of the equilibrium in two steps. We then show the various claims in the statement of the theorem hold. In the first step we compute $\mathbb{E}[q_{ie'}^* \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i]$ corresponding to a symmetric affine equilibrium, $(\alpha_E, \zeta_E, \Delta_E)$, which will be used in the second step. In the second step, we substitute the derived moment from step one into the optimality condition and match coefficients to derive a system of three equations for α_E, ζ_E, and Δ_E.

Step 1: To begin we conjecture an arbitrary symmetric affine equilibrium $(\alpha_E, \zeta_E, \Delta_E)$ and first compute the following unconditional moments.

$$\mathbb{E}[\frac{-\alpha_E(\sum_i X_i) + mN - Q_{e'}}{yN}] = \frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E N} \tag{44}$$

$$\mathbb{E}[\frac{\sum_{j \neq i} -\alpha_E X_j}{\zeta_E(N-1)} - \frac{Q_e}{\zeta_E(N-1)} + \frac{\Delta_E}{\zeta_E}] = \frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E(N-1)} \tag{45}$$

$$Var[\sum_i X_i] = N\sigma_X^2 + 2\Sigma\sum_{i=1}^{N}(i-1) = N\sigma_X^2 + \Sigma(N-1)N \tag{46}$$

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

$$\mathbb{E}[\frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_{e'}}{\zeta_E N} \,|\, X_i] = \\ \frac{-\alpha_E \mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E N} + \frac{\frac{1}{\zeta_E N}(-\alpha_E(N-1)\Sigma - \alpha_E \sigma_X^2 - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) \quad (47)$$

$$\mathbb{E}[\frac{(\sum_{j\neq i} -\alpha_E X_j) - Q_e + \Delta_E(N-1)}{\zeta_E(N-1)} \,|\, X_i] = \\ \frac{-\alpha_E \mu_X + \Delta_E}{\zeta_E} - \frac{\mu_Q}{E\zeta_E(N-1)} + \frac{\frac{1}{\zeta_E(N-1)}(-\alpha_E\Sigma(N-1) - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) \quad (48)$$

$$Var[-\alpha_E(\sum_{j\neq i} X_j) + \Delta_E(N-1) - Q_{e'} \,|\, X_i] = \\ \alpha_E^2(N-1)\sigma_X^2 + \alpha_E^2\Sigma(N-2)(N-1) + \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} + \frac{2\eta\alpha_E(N-1)}{E} - \frac{[(-\alpha_E\Sigma(N-1) - \frac{\eta}{E})]^2}{\sigma_X^2} \quad (49)$$

$$cov_{X_i}(\sum_j -\alpha_E X_j - Q_{e'}, \sum_{j\neq i} -\alpha_E X_j - Q_e) = \\ Var[\sum_{j\neq i} -\alpha_E X_j \,|\, X_i] - 2cov_{X_i}(Q_{e'}, \sum_{j\neq i} -\alpha_E X_j) + cov_{X_i}(Q_e', Q_e) \quad (50)$$

Using the above moments, we compute the following moments, conditional on X_i and $p_e - \Lambda q_{ie}$, (the portion of price in exchange e which is unknown to agent i—see equation (56)) by using the projection theorem. The equations below depend on γ_E, which we will define later.

$$\mathbb{E}[p^*_{ie'} \,|\, p_e - \frac{q_{ie}}{y(N-1)}, X_i] = \\ (1 - \frac{N-1}{N}\gamma_E)\frac{-\alpha_E\mu_X + \Delta_E}{\zeta_E} - (1-\gamma_E)\frac{\mu_Q}{E\zeta_E N} + (1-\gamma_E)\frac{\frac{1}{\zeta_E N}(-\alpha_E(N-1)\Sigma - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) \\ + \frac{-\alpha_E X_i}{\zeta_E N} + \frac{N-1}{N}\gamma_E p_e - \gamma_E \frac{q_{ie}}{\zeta_E N} \quad (51)$$

$$\mathbb{E}[q^*_{ie'} \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i] =$$
$$- \alpha_E X_i \frac{N-1}{N} - (1 - \frac{N-1}{N}\gamma_E)(-\alpha_E \mu_X + \Delta_E) + (1-\gamma_E)\frac{\mu_Q}{EN}$$
$$- (1-\gamma_E)\frac{\frac{1}{N}(-\alpha_E(N-1)\Sigma - \frac{\eta}{E})}{\sigma_X^2}(X_i - \mu_X) - \frac{N-1}{N}\gamma_E \zeta_E p_e + \gamma_E \frac{q_{ie}}{N} + \Delta_E \quad (52)$$

Above, γ_E denotes

$$\frac{cov_{X_i}(\sum_i -\alpha_E X_i - Q_e, \sum_{j\neq i} -\alpha_E X_j - Q_{e'})}{Var[\sum_{j\neq i} -\alpha_E X_j - Q_e \,|\, X_i]}. \quad (53)$$

Of course, $\mathbb{E}[q^*_{ie'} \,|\, p_e - \frac{q_{ie}}{y(N-1)}, X_i]$ 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 $\mathbb{E}[q^*_{ie'} \,|\, p_e - \frac{q_{ie}}{y(N-1)}, X_i]$, 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 (50). 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

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

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

We finally derive that

$$\gamma_E = \frac{\alpha_E^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1) - \frac{\Sigma^2(N-1)^2}{\sigma_X^2}) + 2\alpha_E\frac{\eta}{E}(N-1)(1-\frac{\Sigma}{\sigma_X^2}) + \frac{\rho}{E^2} - \frac{\eta^2}{E^2\sigma_X^2}}{\alpha_E^2((N-1)\sigma_X^2 + \Sigma(N-2)(N-1)) + \frac{\rho}{E^2} + \frac{\sigma_\xi^2}{E} + 2\frac{\eta}{E}\alpha_E(N-1) - \frac{[(-\alpha_E\Sigma(N-1)-\frac{\eta}{E})]^2}{\sigma_X^2}}. \tag{54}$$

This concludes step one.

Step 2. Conjecture that there exists a symmetric affine equilibrium, $(\alpha_E, \zeta_E, \Delta_E)$. Under this conjecture, each agent $i \in N$ submits

$$q_{ie}^{eq} = -\alpha_E X_i - \zeta_E p_e + \Delta_E$$

to each $e \in E$ and $i \in N$, where α_E, ζ_E, and Δ_E are constants. Market clearing in exchange e implies that

$$-\alpha_E(\sum_i X_i) - \zeta_E N p_e + \Delta_E N = Q_e.$$

Solving for p_e yields

$$p_e = \frac{-\alpha_E(\sum_i X_i) + \Delta_E N - Q_e}{\zeta_E N}. \tag{55}$$

Price impact can also be determined from the market clearing condition as well:

$$-\alpha_E(\sum_{j\neq i} X_j) - \zeta_E(N-1)p_e + \Delta_E(N-1) + q_{ie} = Q_e$$

Solving for p_e yields

$$p_e = \frac{-\alpha_E(\sum_{j\neq i} X_i) + q_{ie} + \Delta_E(N-1) - Q_e}{\zeta_E(N-1)}. \tag{56}$$

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 faces in all exchanges. Let p_e^* denote the market clearing price on exchange e when all agents submit their equilibrium demand schedules—that is, the price in equation (55). Let p_e denote the market clearing price in exchange e when agent i submits a demand schedule such that he purchases quantity q_{ie} at the market clearing price as in (56). In determining his optimal demand schedule for exchange e, agent i equates his expected marginal utility conditional on $p_e - \frac{q_{ie}}{\zeta_E(N-1)}$ and X_i, with his marginal cost. Thus, he must compute $\mathbb{E}[q_{ik}^* \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i]$ where $k \neq e$ and q_{ik}^* denotes the equilibrium quantity purchased in exchange k when agent i follows his equilibrium strategy. The optimality condition is

$$\mu_\pi - 2b(X_i + q_{ie} + (E-1)\mathbb{E}[q_{i2}^* \,|\, p_e - \frac{q_{ie}}{\zeta_E(N-1)}, X_i]) = p_e + \Lambda q_{ie}. \tag{57}$$

Applying equation (52) and matching coefficients we can 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

$$\zeta_E = \frac{1}{2b((E-1)\gamma_E + 1)} \frac{N-2}{N-1}. \tag{58}$$

Matching the coefficients on X_i we obtain

$$\alpha_E = \frac{1 + (E-1)\frac{(1-\gamma_E)\frac{\eta}{E}}{N\sigma_X^2}}{1 + \frac{\gamma_E(E-1)}{N} + \frac{(E-1)\gamma_E+1}{N-2} + (E-1)\frac{N-1}{N} - (1-\gamma_E)(E-1)\frac{N-1}{N}\frac{\Sigma}{\sigma_X^2}}. \tag{59}$$

Matching the constant coefficients, we obtain

$$\Delta_E = \frac{N-2}{N-1}\frac{\mu_\pi - 2b(E-1)\mu_X(\frac{(1-\gamma_E)\mu_Q}{EN} - \frac{(1-\gamma_E)\frac{1}{N}(\alpha_E(N-1)\Sigma + \frac{\eta}{E})}{\sigma_X^2} + \alpha_E(1 - \frac{N-1}{N}\gamma_E))}{2b(1+\gamma_E(E-1))} \tag{60}$$

Above, γ_E, as we saw in equation (54) is dependent on α_E, so we have not derived a closed form solution for a candidate equilibrium. By inspecting (59) and (54) 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 (57) is violated on a set of strictly positive $\mathbb{P}$-measure). This concludes step 2.

To prove that the derived necessary conditions are in fact sufficient we appeal to Theorem 3 since, by construction, the symmetric affine equilibria we derived satisfies the condition in Theorem 3. To prove existence of at least one and up to three such symmetric affine equilibria, it suffices to observe from (59) and (54) 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 (54) 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 (58). Part 2 follows immediately from (56). 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 exactly analogous to that of Theorem 1. Part 4 follows from Part 3 and (59) 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 12, $\gamma_E \to 0$. By equation (59),

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

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 7.1. *Let $\{E\alpha_E\}_{E\in\mathbb{N}}$ be defined as in Theorem 7. 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 7.1. Theorem 7 supplies a closed form expression for the limiting value of

$E\alpha_E$ as $E \to \infty$. The rest of the proof follows from some simple computations.

$\square$

Proposition 11. *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 7, 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 11. Let $(\alpha_E, \zeta_E, \Delta_E)$ denote an arbitrary symmetric affine equilibrium. Define $g_E \equiv E\alpha_E$. Substituting equation (54) into (59) 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 15. □

Proposition 12. *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 12. The claims when $\sigma_\xi^2 = 0$ are obvious in light of Theorem 7. We prove the claims when $\sigma_\xi^2 > 0$. By inspecting equation (59), and recognizing that Lemma 6 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 (54), 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 7, $-El_E$ 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

7 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 13. *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 13. The proof is analogous to that of Proposition 1 so we omit it. □

Proposition 14. *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 14. 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 15. *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 15. 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.

Babus, A., and C. Parlatore. 2017. Strategic fragmented markets. Working paper, New York University, Stern School.

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.

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.

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.

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 markets? *American Economic Journal: Microeconomics* forthcoming.

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