The Formation of Exchanges: Risk Sharing and Information Aggregation*

Kei Kawakami†

November 19, 2009

Abstract

Why are some financial markets segmented and opaque? I propose a two-stage game framework to study how such a market structure arises as a result of interaction among traders and exchanges. In the first stage, traders and exchanges play a market formation game, which determines a number and size of exchanges. In the second stage, motivated by both risk sharing and speculation, traders in each exchange play a trading game. I show that the gain from trade is hump-shaped in the number of traders. When many traders speculate, prices become so informative that the ex ante gain from risk sharing is reduced. This can endogenously constrain the exchange size in the market formation game. I show that free entry determines the number and size of exchanges such that prices don't reveal much information, and thus each trader's gain from trade is maximized.

Keywords: Asymmetric information, CARA-normal, Imperfect competition, Information sharing, Market segmentation.

---

*I thank Andrew Atkeson, Antonio Bernardo, Francisco Buera, Bruce Carlin, Roger Farmer, Mark Garmaise, Shogo Hamasaki, Christian Hellwig, Hugo Hopenhayn, Ichiro Obara, Lee Ohanian, Marek Pycia, Avanidhar Subrahmanyam, Pierre-Olivier Weill, and participants in seminars at DBJ and UCLA for comments and suggestions. All errors are mine. The latest version is downloadable at http://kei.bol.ucla.edu/

†Department of Economics, University of California Los Angeles, e-mail: kei@ucla.edu

1 Introduction

In many financial markets, trades occur within a restricted, small group of traders, rather than in an open-access, large auction. Examples of such segmented markets include the markets for corporate bonds, foreign exchange, derivatives, and inter-bank lending. These markets are often described as opaque since only a small amount of trade-relevant information is shared among market participants. Empirical studies suggest that sometimes traders prefer segmentation and opaqueness to openness and transparency, and that some markets have indeed evolved in that direction (Pagano and Roell (1992); Board and Sutcliffe (1996); Gemmill (1996); Biais and Green (2007); Ready (2008)).

These observations raise the following questions: Why are segmentation and opaqueness attractive to traders? How does such a market structure arise? What is its welfare implication? I develop a theoretical market formation model to address these questions. The model shows that the emergence of segmented and opaque financial markets is consistent with traders' desire for risk sharing, which would be difficult in an open and transparent exchange. Existing empirical evidence offers support for the importance of risk sharing in financial markets (Reiss and Werner (1998, 2004)).

I analyze a two-stage game framework, where exchanges and traders play a market formation game at the first stage, and traders in each exchange play a trading game at the second stage. At time zero, referred to as the ex ante stage, all traders are identical. Exchanges provide costly trading service for traders and charge fixed entry fees. Traders decide in which exchange to participate. The equilibrium of market formation game determines the number and size of exchanges. At time one, referred to as the interim stage, traders in each exchange play a trading game. Traders trade a risky asset given their private information. The equilibrium of the trading game determines allocation of the asset. At time two, referred to as the ex post stage, all the uncertainty is resolved and traders enjoy realized profits.

In the trading game at the second stage, traders are motivated by both risk sharing and speculation. Each trader comes to the exchange with two pieces of private information.

First, his risky asset endowment determines his own risk sharing needs and provides more risk sharing opportunities for other traders. If this were the only factor each trader brings into the exchange, the gain from trade would increase as the exchange size increases. Second, each trader has a private signal about the risky asset return, which motivates informed speculation. In the trading game, each trader submits an order that is contingent on the market-clearing price ("limit order"). This allows each trader to condition his order in part on other traders' signals. Thus, the price not only clears the market but also serves as a communication tool among traders. Importantly, this information sharing *decreases* the benefit of risk sharing, because the risk to be shared has decreased due to information revealed through prices.

The second mechanism mentioned above is one variant of the Hirshleifer Effect, that is, that the revelation of information can prevent risk sharing and decrease welfare. In the trading game, the trade-off between the size of risk sharing pool and the Hirshleifer Effect produces a hump-shape gain. On the one hand, a large exchange increases risk sharing possibilities, and thus, trading volume. On the other hand, increased speculation makes prices more informative and decreases the benefit of risk sharing. The latter effect works independently of the quantity traded, because it is an informational force. Therefore, in a large exchange, gain from trade approaches zero even though each trader's trading volume is large. The hump-shape is illustrated in **Figure 1**.

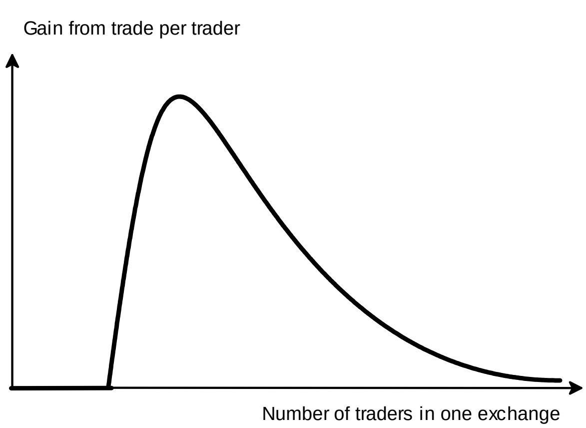

Figure 1

The hump-shape influences how exchanges form at the first stage. I begin my analysis of the market formation game by assuming that there is only one exchange. In this case, since traders will participate as long as the gain from trade exceeds the entry fee, the exchange can extract all the gain from trade as its profit. At the fee level that maximizes the monopoly exchange's profit, the exchange could further expand its size by lowering its fee, but there is no incentive to do so. This is the case because each additional trader creates negative externality and lowers the gain from trade of all other traders inside the exchange. At the monopoly optimum, the gain from each additional trader and the loss from a decreased gain for current traders in the exchange are equalized. Hence, the monopoly exchange limits the entry of traders at some finite size.

I then study how competition between two exchanges changes the analysis. At first glance, it would seem that Bertrand-like competition would induce the two exchanges to bid down the entry fee to marginal cost. Interestingly, this is not necessarily the case. If the potential number of traders is large enough, both exchanges set monopoly fees, and there is no incentive to lower the fee, just like in the monopoly case. When the potential number of traders is not large enough, the exchanges cannot reach monopoly size, and competition starts to matter. Still, as long as traders create negative externality at the shared market size, the exchanges make profits. This occurs because each exchange knows that setting a higher fee will not drive all of the traders to their competitor.

Competition between two exchanges becomes intense when the shared market size is small enough that the negative externality among traders disappears. In this case, there is strong incentive to cut fees, since a slightly lower fee allows each exchange to steal all of the traders from another exchange. In sum, competition has bite when it forces the exchanges to have small size. This creates the incentives to expand exchange sizes by cutting fees and, thus, decreases the exchanges' profits. I show that free entry increases the number of exchanges until each exchange size attains the peak of the hump-shaped gain derived in the trading game (i.e., until each trader's gain from trade is maximized).

I briefly review the related literature next. Following that, Sections 2 and 3 analyze the trading game for a given number of traders, which I take as an exchange size. Section 2 describes the environment of the game. Section 3 characterizes equilibrium and derives hump-shaped gain from trade. The exchange size is endogenized through a market formation game in Section 4. In particular, I study the implication of the hump-shape for competition among exchanges. Section 5 concludes. Appendix A contains all proofs. Appendix B compares endogenous asset supply in the trading game with exogenous supply.

1.1 Related literature

It is well documented that some financial markets are highly segmented and opaque. The over-the-counter (OTC) market for bonds is one extreme example where trades are often conducted as a bilateral negotiation between two traders. Biais and Green (2007) study the history of the U.S. bond market and show that bond trading was active on the New York Stock Exchange (NYSE) as a transparent limit-order market until the 1940s. Then it declined as trading migrated to the OTC market. Today, only a small fraction of bond trading is done through a centralized market on the NYSE.

Such migration has also been observed for stocks. For example, Pagano and Roell (1992) argue that when an opaque dealer market for cross-listed stocks was created in London in 1986, it drew "volume away from the auction markets of continental Europe." Another example of preference for opaqueness was found in the London Stock Exchange (LSE). According to Board and Sutcliffe (1996), delayed publication of trades was introduced in 1989 in response to pressure from market makers. Gemmill (1996) also studies the LSE and argues that "there will be a tendency for large transactions to gravitate toward the least transparent available market."[1]

More recently, private trading networks that match orders without routing them to a public exchange — sometimes called "dark pools"— have been developed and gained popu-

[1] Gemmill (1996) also states that, recognizing this tendency, the Paris Bourse introduced a delayed publication in 1994 and there have been suggestions that Frankfurt do the same.

larity.[2] In a study of three major dark pools, Ready (2008) states that their trading systems cater to a specific group of institutions and eliminate small orders to control information leakage.

Reiss and Werner (1998, 2004) study inter-dealer trades in the LSE to test whether risk sharing is an important motivation for trading. They conclude that dealers are risk-averse and use inter-dealer trades to share inventory risk. They also study anonymous brokered trading systems in the LSE, and state that the LSE rules restricted access to the anonymous systems, making the system a risk sharing device for registered market makers.

The trading game in this paper uses the noisy rational expectation equilibrium (REE) setup, which was initiated by three papers: Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981). My model environment is similar to Diamond and Verrecchia (1981), except that I study imperfect competition among traders. Following Kyle (1989), I characterize a Nash equilibrium in demand functions. As Reny and Perry (2006) criticize, most noisy REE models have the feature that no trade occurs without the presence of irrational agents. To get around this problem and facilitate welfare analysis, I replace noise trading in Kyle (1989) with rational traders' private endowments, as in Diamond and Verrecchia (1981).

Reny and Perry (2006) study the strategic foundation of an REE by considering a double auction with risk-neutral agents. They show that as the market size increases equilibrium converges to a fully revealing and efficient REE. Although their analysis is much more general than this paper in many dimensions, one important difference is that traders are risk-averse in my model. I show that when price aggregates private information, risk-averse traders' ex ante welfare may be reduced. This force, missing in many models with risk-neutral agents, has an important implication for the efficiency of large markets.

Welfare reducing information sharing in my model is one variant of the Hirshleifer Effect (Hirshleifer (1971)). Pithyachariyakul (1986) compares the social welfare of an exchange

[2] See the following article for dark pools: www.businessweek.com/investor/content/oct2007/pi2007102_394204.htm.

economy under two market structures: Walrasian and monopolistic market makers. In the Walrasian system, the asymmetry of information disappears because of the communication through prices. This allows for an efficient risk sharing against the remaining risk, while some insurance opportunities are destroyed. In the market maker system, preserved asymmetric information means that more risk sharing is possible but the allocation is less efficient. The relative efficiency of two systems depends on the nature of uncertainty.

Marin and Rahi (1999, 2000) analyze the trade-off between the adverse selection effect and the Hirshleifer Effect in the security design problem. They show that traders may be better off with the set of securities that leaves prices noisy. Hatchondo, Krusell and Schneider (2008) consider an asset market in which investors have private information about both payoffs and their own exposure to an aggregate risk. Similar to the result in my model, they show that there is less risk sharing than in a symmetric information benchmark.

Finally, Bernardo and Judd (1997) examine welfare of traders in the Grossman and Stiglitz (1980) exchange economy by replacing noise traders with rational traders whose risk aversions are not common knowledge. They show that information gathering leads to suboptimal risk sharing.

Compared with these works featuring the Hirshleifer Effect, the primary contribution of this paper is to analyze the Hirshleifer Effect under a fully strategic environment with a closed-form solution. In particular, the model identifies the mechanism through which the Hirshleifer Effect can shape a market structure.

My model also yields a version of no-trade result due to adverse selection, which is extensively studied in the literature. Glosten (1989), Bhattacharaya and Spiegel (1991), Subrahmanyam (1991), Bhattacharaya, Reny and Spiegel (1995), and Spiegel and Subrahmanyam (1992, 2000) all show no-trade result under various circumstances. Mailath and Noeldeke (2008) analyze two-dimensional asymmetric information, and provide necessary and sufficient conditions for market breakdowns. Compared with these works, no-trade result in my model is much weaker since I focus on the linear equilibrium. However, the papers cited above all assume a nested information structure, i.e., information asymmetry only between a group of symmetrically informed agents and uninformed agents. In contrast, I analyze an arbitrary number of traders with heterogeneous information. In this environment, if the asset supply is exogenous noise trading as in Kyle (1989), it is well known that in the limit as noise trading vanishes, markets become so illiquid that profit-maximizing quantities are driven to zero. In my model, asset supply is endogenous, and I show that the same symptom can arise away from the zero-noise limit.[3]

2 Trading game environment

Consider N traders indexed by $i = 1, ..., N$ in a single exchange, who have CARA preferences and trade a risky asset with return v. The return v is not known but traders have a common prior that it is a realization of a normal random variable with mean zero and variance τ_v^{-1}. Each trader has two types of private information: (i) risky asset endowment x_i, and (ii) a private signal s_i about the return v.

Endowment x_i can be interpreted as follows: Suppose that traders are professional and their endowments are created as a result of preceding trades with their customers. The professional traders then come to an inter-dealer market to adjust their positions. An implicit assumption here is that customers do not have direct access to the inter-dealer market. The sum of endowments $\sum_{i=1}^{N} x_i$ is the total amount of the risky asset in the market. Though each trader does not know the other traders' endowments, traders have a common prior that each trader's endowment is a realization of an independent normal random variable with mean zero and variance τ_x^{-1}.

It is assumed that the private signal takes the form $s_i = v + \varepsilon_i$, where ε_i is unobserved noise in the signal, and follows a normal distribution with mean zero and variance τ_ε^{-1}. To summarize, $2N + 1$ random variables $v, x_1, ..., x_N, \varepsilon_1, ..., \varepsilon_N$ are assumed to be normally and

[3] Also, since interim allocation is Pareto inefficient due to dispersed endowments, a simple no-trade theorem (e.g. Milgrom and Stokey (1982)) does not apply.

independently distributed with zero means, and variances

$$Var\left(v\right)=\tau_v^{-1},\ Var\left(x_i\right)=\tau_x^{-1},\ Var\left(\varepsilon_i\right)=\tau_\varepsilon^{-1}.$$

After observing the realization of private information (s_i, x_i), each trader chooses her order $q_i\left(p; s_i, x_i\right)$ which is explicitly conditioned on the market-clearing price p. The utility function of traders is

$$U\left(\pi_i\right)=-\exp\left(-\rho\pi_i\right),$$

where π_i is trader i's profit and ρ is the constant absolute risk aversion. Profit is the sum of return on the new position $q_i + x_i$ and payment or receipt for net trading q_i, and thus, $\pi_i = v\left(q_i + x_i\right) - pq_i$.

Following Kyle (1989), a demand schedule $q_i\left(p; s_i, x_i\right)$ is allowed to be any convex-valued, upper-hemicontinuous correspondence which maps prices p into non-empty subsets of the closed infinite interval $[-\infty, \infty]$. The following market-clearing rule gives a well-defined market price for all strategy choices and a well-defined allocation given a finite price. An auctioneer calculates the set of market-clearing prices and quantity allocation which satisfy

$$\sum_{i=1}^{N} q_i\left(p; s_i, x_i\right)=0. \tag{1}$$

An allocation with infinite trade is assumed to be market-clearing if and only if there is at least one positive and one negative infinite quantity at that price. If a market-clearing price exists, the auctioneer chooses the price with minimum absolute value and the market-clearing quantity allocation that minimizes the sum of squared quantities traded. If there is positive excess demand at all prices, $p=\infty$ is announced and all buyers receive negative infinite utility. If there is negative excess demand at all prices, $p=-\infty$ is announced and all sellers receive negative infinite utility.

Suppose that the set of market-clearing prices that satisfy (1) is empty. From the restriction on the correspondence, it must be either: (i) there is positive excess demand at all prices, or (ii) that there is negative excess demand at all prices. Since positive (negative) infinite price will be announced in the first (second) case, some traders will receive negative infinite utility. Similarly, the situation with both positive and negative infinite quantity at some price level will give at least one trader negative infinite utility. Thus, infinite prices and quantities do not occur in equilibrium.

This market-clearing rule specifies how a market-clearing price depends on the strategies of traders. To make this dependence explicit, write

$$p = p(q), \; q_i = q_i(q),$$

where $q = (q_1, ..., q_N)$ is a vector of strategies. A rational expectations equilibrium with *imperfect competition* is defined as a q that satisfies

$$E[U((v - p(q))\, q_i(q) + vx_i)] \geq E[U((v - p(q'))\, q_i(q') + vx_i)] \tag{2}$$

for all $i = 1, ..., N$ and for any q' differing from q only in the i-th component. For comparison, a rational expectations equilibrium with *price-taking* is defined as a q that satisfies

$$E[U((v - p(q))\, q_i(q) + vx_i)] \geq E[U((v - p(q))\, q_i(q') + vx_i)] \tag{3}$$

for all $i = 1, ..., N$ and for any q' differing from q only in the i-th component. In (2), the expected utility under deviation from q is calculated by taking into account the effect on the price $p(q')$. In (3) trader i assumes that the price is not affected by his deviation.

An important feature of the trading game is illustrated in the market-clearing condition (1). In most models that use a CARA-normal framework, an independent random variable, often named noise trading, is added in (1). Noise trading is also referred to as a liquidity trader, since it offers a random trading opportunity for the other traders. In the same sense, each rational trader in this model is also a liquidity trader because each trader's private endowment creates a random trading opportunity for the other traders. However, while the randomness of noise trading is exogenously fixed, the impact of random endowments depends on how much traders change their initial positions to reduce the risk associated with the asset return. This in turn affects the amount of noise in the price, and thus the traders' risk assessments. This mutual dependence between the traders' risk sharing needs and price informativeness endogenizes liquidity supply in the trading game.

The absence of irrational traders facilitates welfare analysis. Here, I introduce a welfare measure based on each trader's ex ante utility. Let $E_i[\cdot]$ denote trader i's interim expectation, $E[\cdot|s_i, x_i, p]$, which is conditional on the information that he has at the time of order submission.

Definition 1 *Interim certainty equivalent (ICE) profit* Π_i^{ice} *is defined by* $E_i[\exp(-\rho\pi_i)] = \exp(-\rho\Pi_i^{ice})$*. Ex ante certainty equivalent (ACE) profit* Π^{ace} *is defined by* $E[\exp(-\rho\Pi_i^{ice})] = \exp(-\rho\Pi^{ace})$*.*

Definition 2 *Gain from trade (GFT) is defined by the difference in ACE profit between trading and not trading.*

When there is a trade, the profit is given by $\pi_i = v(q_i + x_i) - pq_i$, where quantity traded q_i and the price are determined in equilibrium. When there is no trade, profit is $\pi_i = vx_i$. Note that each trader can always achieve this no-trade profit by submitting zero demand. In fact, no trade is one equilibrium outcome. Let $Var_i[\cdot]$ denote trader i's conditional variance, conditional on the same information in $E_i[\cdot]$. Recognizing that conditional variance of v will be independent of i due to normality, I define the precision of each trader $\tau \equiv (Var_i[v])^{-1}$. Further, I assume $1 - \frac{\rho^2}{\tau_v\tau_x} > 0$ throughout the paper so that ex ante utility is well defined.

Lemma 1 (no-trade) *A no-trade equilibrium always exists. In this equilibrium,* $E_i[v] = \frac{\tau_\varepsilon}{\tau}s_i$ *and* $\tau = \tau_v + \tau_\varepsilon$*. ICE profit is* $\Pi_{iNT}^{ice} = E_i[v]x_i - \frac{\rho}{2\tau}x_i^2$*. ACE profit is* $\Pi_{NT}^{ace} = \frac{1}{2\rho}\log\left(1 - \frac{\rho^2}{\tau_v\tau_x}\right)$*.*

This lemma serves two purposes. First, the option of not trading, Π_{NT}^{ace}, sets a participation constraint for any trading arrangements. Note that neither ICE nor ACE profit depends on the number of traders N due to the autarkic nature. The next section derives ACE profit in equilibrium with trading. Gain from trade in **Definition 2** is then given by the derived ACE profit with trading minus Π_{NT}^{ace} in **Lemma 1**.

Second, the form of ACE profit is useful to understand an informational property of the trading game. No-trade ACE profit does not depend on precision of private signals τ_ε, while ICE profit does. In a no-trade equilibrium, private signal affects the interim stage valuation of the asset, but does not change the allocation. Informational difference at the interim stage can not affect ex ante utility, unless it changes trading behavior and thus allocation. On the other hand, τ_v and τ_x are relevant for no-trade ACE profit, since the former affects ex ante valuation and the latter directly affects interim and ex post allocation.

3 Trading game equilibrium

This section contains three subsections. The first subsection establishes the existence of equilibrium. As a benchmark, two special cases, price-taking equilibrium and no-speculation equilibrium, are also presented. Subsection 3.2 analyzes hump-shaped gain from trade. The last subsection provides numerical evaluation of the hump-shape.

3.1 Equilibrium characterization

I use a guess-and-verify method and start from the following conjecture.

Conjecture *There exists a symmetric linear equilibrium, in which the strategies are*

$$q_i\left(p; s_i, x_i\right) = \beta_s s_i - \beta_x x_i - \beta_p p \tag{4}$$

for some positive coefficients $(\beta_s, \beta_x, \beta_p)$.

Given a conjecture that the other traders use a strategy of the form (4), a maximization problem of trader i is solved. The solution, which is the best response to the conjectured strategy, must have the same form by symmetry. This defines a set of equations in coefficients β_s, β_x and β_p, which defines a fixed point in $\mathbb{R}^3$.

The key feature of the trading game is information sharing through prices. On the one hand, how much information is shared among traders and how much is kept private affect trading, because trading is partially driven by informational differences among traders. On the other hand, trading reveals information and changes the information distribution among traders. To capture this interaction between information and trading, I formally define the amount of information sharing. Recall that precision τ for any trader is defined by

$$\tau \equiv \left(Var_i[v]\right)^{-1}.$$

This is bounded below by its value in the no-trade equilibrium $\tau_v + \tau_\varepsilon$. At the other extreme, if traders can share all the available information about v, τ will be $\tau_v + N\tau_\varepsilon$. Thus, there exists a constant $\varphi \in [0, 1]$ such that

$$\tau = \tau_v + \tau_\varepsilon + \varphi \left(N - 1\right) \tau_\varepsilon. \tag{5}$$

The parameter φ measures the share of the other $N - 1$ traders' private signals that are revealed through prices. If φ is zero, prices do not reveal any information about v. If φ is one, prices reveal all the private signals in the market.

The following lemma characterizes φ given conjecture (4), and its relation to the valuation of the asset.

Lemma 2 (information sharing) *Given (4), the amount of information sharing is* $\varphi = \frac{1}{1+\left(\frac{\beta_x}{\beta_s}\right)^2 \frac{\tau_\varepsilon}{\tau_x}}$, *and* $E_i[v] = \frac{\tau_\varepsilon(1-\varphi)}{\tau} s_i + \frac{\tau_\varepsilon \varphi}{\tau} \frac{\beta_x}{\beta_s} x_i + \frac{\tau_\varepsilon \varphi N}{\tau} \frac{\beta_p}{\beta_s} p.$

First, it is useful to see how φ depends on β_s and β_x. Since β_s measures the sensitivity of orders to a private signal s_i, higher β_s implies more information sharing in the market. The relation between φ and β_x is negative, because as orders become more responsive to endowments, a market-clearing price becomes more noisy (recall that endowment contains no information about v).

Second, the conditional mean of the asset return is increasing in s_i, x_i and p, provided that β_s, β_x and β_p are positive in equilibrium. The signs of s_i and p are positive, since a higher private signal or market-clearing price implies a higher return v. The sign of x_i is positive for the following reason: Since large endowments lower prices through the market-clearing condition, a trader with high endowment rationally expects (for a given market-clearing price) that the other traders' private signals indicate a high return.

Finally, **Lemma 2** illustrates the interaction between information and trading. On the one hand, trading behavior captured by β_s and β_x affects information sharing φ. On the other hand, φ affects the traders' valuation of the asset $E_i[v]$ and $Var_i[v] = \tau^{-1}$. Since the traders maximize a mean-variance criterion due to CARA-normal set-up, $E_i[v]$ and τ characterize the optimal order q_i.

To derive the best response of trader i, I start from the market-clearing condition $\sum_{j \neq i} q_j + q_i = 0$. Given conjecture (4), $\sum_{j \neq i} q_j = \beta_s \sum_{j \neq i} s_j - \beta_x \sum_{j \neq i} x_j - (N-1)\beta_p p$. Solving for price, we obtain

$$p = c_i + \lambda q_i, \tag{6}$$

where $c_i \equiv \frac{\beta_s}{\beta_p}\overline{s}_{-i} - \frac{\beta_x}{\beta_p}\overline{x}_{-i}$, $\overline{s}_{-i}$ and $\overline{x}_{-i}$ are the average of private signals and endowments held by all traders except i, and $\lambda \equiv \frac{1}{(N-1)\beta_p}$. Equation (6) is trader i's residual supply curve with slope λ and an intercept c_i, where λ measures each trader's price impact. Although trader i can not directly condition her order on c_i, (s_i, x_i, p) provides the same information as (s_i, x_i, c_i) through the linear relationship (6).

The objective function at the interim stage is $E_i[-\exp(-\rho\pi_i)]$. Because of the normality of v conditional on each trader's information, the objective of the trader is to maximize the mean-variance criterion

$$E_i[v]\left(q_i + x_i\right) - \frac{\rho}{2} Var_i\left[v\right]\left(q_i + x_i\right)^2 - pq_i \tag{7}$$

subject to (6).

The first-order condition and the second-order condition are

$$E_i[v] - \rho Var_i\left[v\right]\left(q_i + x_i\right) = c_i + 2\lambda q_i = p + \lambda q_i, \tag{8}$$

$$2\lambda + \rho Var_i\left[v\right] > 0. \tag{9}$$

The first-order condition (8) equates the trader's marginal valuation of the asset (left-hand side) to his marginal cost of trading (right-hand side). Marginal valuation is decreasing in the trader's new position $q_i + x_i$ because of risk aversion ρ and conditional variance $Var_i\left[v\right]$. Marginal cost is increasing in his order q_i due to price impact λ. The second-order condition (9) needs to be verified since $\lambda \equiv \frac{1}{(N-1)\beta_p}$ is determined in equilibrium.[4]

From (8) and $Var_i\left[v\right] = \tau^{-1}$, we obtain

$$q_i^* = \frac{E_i[v] - p - \frac{\rho}{\tau} x_i}{\lambda + \frac{\rho}{\tau}}. \tag{10}$$

Using $E_i[v] = k_1 s_i + k_2 x_i + k_3 p$, where k_1, k_2, k_3 are given in **Lemma 2**, we obtain

$$q_i^* = \frac{k_1 s_i - \left(\frac{\rho}{\tau} - k_2\right) x_i - \left(1 - k_3\right) p}{\lambda + \frac{\rho}{\tau}}. \tag{11}$$

This is the best response of trader i when the other traders use (4). By equating coefficients of (4) and (11), we have three equations

$$\beta_s = \frac{k_1}{\lambda + \frac{\rho}{\tau}} = \frac{\tau_\varepsilon}{\lambda\tau + \rho}(1 - \varphi), \tag{12}$$

[4]The second-order condition (9) must hold with strict inequality. With equality, traders would like to order infinite quantity because $E_i[v] - \rho Var_i\left[v\right] x_i \neq c_i$.

$$\beta_x = \frac{\frac{\rho}{\tau} - k_2}{\lambda + \frac{\rho}{\tau}} = \frac{\rho}{\lambda\tau + \rho}\left(1 - \varphi\frac{\tau_\varepsilon}{\rho}\frac{\beta_x}{\beta_s}\right), \tag{13}$$

$$\beta_p = \frac{1 - k_3}{\lambda + \frac{\rho}{\tau}} = \frac{\tau}{\lambda\tau + \rho}\left(1 - \varphi\frac{N\tau_\varepsilon}{\tau}\frac{\beta_p}{\beta_s}\right), \tag{14}$$

where $\lambda \equiv \frac{1}{(N-1)\beta_p}$ and $\varphi = \frac{1}{1+\left(\frac{\beta_x}{\beta_s}\right)^2\frac{\tau_\varepsilon}{\tau_x}}$.

The following proposition establishes the existence and uniqueness of a symmetric linear equilibrium with trading.

Proposition 1 (trade equilibrium) (a) *A symmetric linear equilibrium with trading exists if and only if* $\frac{\tau_\varepsilon\tau_x}{\rho^2} < 1 - \frac{2}{N}$. *When it exists, it is unique.* (b) *Optimal order is* $q_i(p) = \left(1 - 2\varphi - \frac{1}{N-1}\right)\left\{\frac{\tau_\varepsilon}{\rho}s_i - x_i - \frac{\tau_v + \tau_\varepsilon + (\tau_v + \tau_\varepsilon N)I}{\rho(1+NI)}p\right\}$, *where* $\varphi = \frac{I}{1+I}$ *and an information parameter* $I \equiv \frac{\tau_\varepsilon\tau_x}{\rho^2}$.

Henceforth, a symmetric linear equilibrium with trading shall be called a *trade equilibrium*. A number of important observations follow from **Proposition 1**. First, a trade equilibrium is a partially revealing REE. In the equilibrium, not all of signals are revealed through prices, because $\varphi = \frac{I}{1+I}$ is smaller than one. The amount of information sharing is increasing in the information parameter $I \equiv \frac{\tau_\varepsilon\tau_x}{\rho^2}$. When τ_ε and τ_x are higher, or when risk aversion ρ is lower, more information is shared in equilibrium.

Second, the amount of information sharing is independent of N. This implies that a trade equilibrium is partially revealing even in the limit where N goes to infinity. However, since $\tau = \tau_v + \tau_\varepsilon + \varphi(N-1)\tau_\varepsilon$, conditional variance of v goes to zero in the limit.

Third, **Proposition 1** shows that $\frac{\beta_x}{\beta_s} = \frac{\rho}{\tau_\varepsilon}$. This ratio represents the balance between risk sharing and speculation. The more risk averse traders are, the more weight they put on their endowments in order to adjust their positions. The more precise their private signals are, the more weight they put on private signals.

Finally, three coefficients β_s, β_x, β_p have well-defined limit as N goes to infinity. This suggests that there will be non-trivial trade in a large market. In fact, it will be shown later that trading volume *per trader* is increasing in N. This is in part because price impact disappears in the large market, since $\lambda \equiv \frac{1}{(N-1)\beta_p}$ goes to zero as N goes to infinity.

To gain more insight, I study two special cases of the trade equilibrium. The first case is a price-taking equilibrium. This is obtained by setting $\lambda = 0$ in (8) and (9). This shuts down price impact for any market size. The second case is a no-speculation equilibrium, where a private signal is assumed to be useless and traders trade only for a risk sharing motive. Setting $\tau_\varepsilon = 0$ shuts down the speculation by making information about the asset return symmetric.

Corollary 1 (price-taking) *Assume $\lambda = 0$ in (8) and (9). In equilibrium, the amount of information sharing φ is same as in the strategic case. Optimal order is obtained by replacing $1 - 2\varphi - \frac{1}{N-1}$ in the strategic case with $1 - \varphi$.*

The result is immediate by setting $\lambda = 0$ in (12) – (14), so the proof is not provided. Notice that the price impact λ affects the level of coefficients through (12) – (14), but not the ratio. Recall from **Lemma 2** that φ depends only on $\frac{\beta_x}{\beta_s}$. Since this ratio is independent of price impact, so is information sharing. Moreover, from (5) and **Lemma 2**, neither τ nor $E_i[v]$ is affected by price-taking assumption. Where does the strategic behavior matter then? **Corollary 1** shows that three coefficients are larger under the price-taking assumption. Therefore, the strategic behavior reduces trading volume relative to the price-taking behavior, while still allowing each trader to have the same amount of information.

Corollary 2 (no-speculation) *Assume $\tau_\varepsilon = 0$.* (a) *$E_i[v] = 0$, $\tau = \tau_v$, and traders submit $q_i(p) = -\left(1 - \frac{1}{N-1}\right) x_i - \frac{\tau_v}{\rho}\left(1 - \frac{1}{N-1}\right) p$ in equilibrium.* (b) *ACE profit is increasing in N with a finite upper bound.*

Corollary 2 shows that, without asymmetric information, a larger exchange provides larger gain from trade *per trader*. This positive externality comes from the fact that each additional trader creates more risk sharing opportunity for the other traders. In this case, it is easy to verify that the market-clearing price is $p = -\frac{\rho}{\tau_v}\overline{x}$ and the corresponding quantity traded $q_i = -\left(1 - \frac{1}{N-1}\right)(x_i - \overline{x})$, where $\overline{x} = \frac{1}{N}\sum x_i$. In the limit where the market size approaches infinity, each trader unloads her initial position in the market ($q_i(p) = -x_i - \frac{\tau_v}{\rho} p$), and after the trade everyone holds average positions that are the ex ante mean ($q_i + x_i = \overline{x} = 0$).

Figure 2 illustrates gain from trade in the no-speculation case, which is monotonically increasing and approaching the perfect risk sharing upper bound as the exchange size increases. By definition, the no-trade case has zero gain.

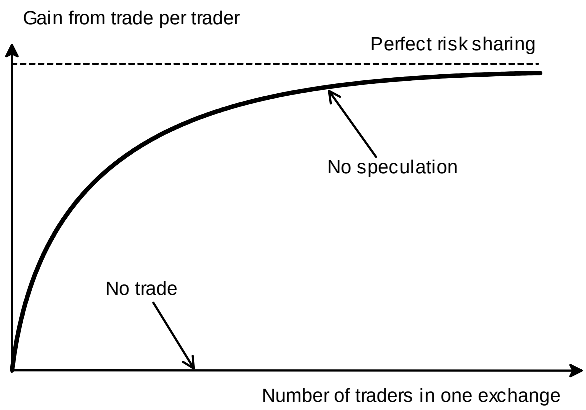

Figure 2

What happens if traders can use private signals? **Figure 3** illustrates the consequence of speculation. Introduction of private signals decreases gain from trade, and importantly, the negative effect does not disappear in a large market. The next subsection derives the hump-shape shown in **Figure 3**.

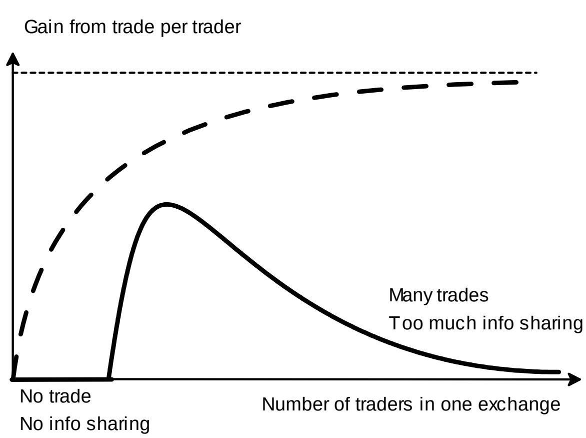

Figure 3

3.2 Hump-shaped gain from trade

In this subsection, I first establish the non-monotonicity of gain from trade by studying two polar cases: a small and large exchange. After that, I discuss the source of the hump-shape. Finally, the hump-shape is characterized.

I start with a small market. **Proposition 1** states that there is a parameter region where a trade equilibrium does not exist. The condition in **Proposition 1** can be written as,

$$\underline{N}(I) \equiv \frac{2}{1-I} < N \text{ for } I < 1. \tag{15}$$

To focus on a linear equilibrium, $I < 1$ is fixed for the rest and (15) is assumed to be satisfied. The next lemma shows what happens when the exchange size approaches the lower bound (15).

Lemma 3 (small exchange) *As condition (15) becomes close to equality, trading volume approaches zero and λ goes to infinity.*

This establishes the left end of **Figure 3**. When the exchange size is small, each trader's price impact becomes huge, and so does marginal cost of trading (right-hand side of (8)). Price impact can become arbitrarily large because $\lambda \equiv \frac{1}{(N-1)\beta_p}$, β_p is proportional to $1-2\varphi-\frac{1}{N-1}$, and thus, when (15) approaches equality, the denominator of λ approaches zero. When the exchange size is small, each trader anticipates a large price impact, and thus makes his order less sensitive to the price (smaller β_p). Moreover, when other traders' orders become less sensitive to price, each trader has a bigger price impact. Hence, traders' price impacts are mutually reinforcing through the market-clearing rule. Due to this mechanism, price impact and marginal cost can be so large that profit maximizing trading is driven down to zero. This argument does not hold under price-taking assumption.

It is well known in the literature that in models with exogenous noise trading, the symptom in **Lemma 3** arises when noise trading, and thus noise in prices, disappears. In contrast, as condition (15) becomes close to equality, price becomes more informative, but remains noisy. Nevertheless, trading volume goes down to zero and price impact goes to infinity. The key to this difference is that information sharing is independent of N in this model, while it positively depends on N in the model with noise trading. Note that β_p may become smaller either: (i) due to smaller N or (ii) due to larger φ. When φ is independent of N, the small market increases price impact through the first channel. When φ is positively related to N, the small market decreases φ. This increases β_p through the second channel and partially offsets the increased price impact through the first channel. In the model with noise trading, the impact of noise trading becomes large relative to rational traders' private signals in a small market, thus making the price less informative. This occurs because noise trading is exogenously fixed independent of N. In the current model, the impact of speculation and that of risk sharing are both proportional to N. This leaves φ independent of N in equilibrium.[5]

The following proposition expresses the properties of a large exchange.

Proposition 2 (large exchange) (a) *Let $E[v|p]$ be expected return without private information. It follows that $\frac{E[v|p]-p}{p} = \eta - 1$, where $\eta < 1$. Also, η is increasing in N and approaches* 1 *as N goes to infinity.* (b) *Expected per-trader trading volume is increasing in N and approaches a finite upper bound as N goes to infinity.* (c) *As $N \to \infty$, ACE profit approaches no-trade ACE profit.*

Proposition 2 characterizes the right end of **Figure 3** and establishes non-monotonicity of gain from trade. The first and second parts of **Proposition 2** show that as the exchange size increases, price becomes more informationally efficient and trading volume increases. In the meantime, however, the last part shows that gain from trade decreases to zero.

When the exchange size is finite, price is only a noisy signal of v. This price is not informationally efficient in the following sense. Note that $E[v|p] - p$ is the excess return forecast by an econometrician outside the exchange, who knows the rules of the trading game,

[5] Appendix B provides further discussion on this point.

observes prices, but cannot directly observe the traders' private information.[6] According to **Proposition 2**, the expected excess return is negative. This implies that if an uninformed trader without endowments enters the exchange, he has an incentive to short sell the asset.[7] In other words, there is arbitrage opportunity for outsiders *without risk sharing needs*, if there is no restriction on market entry and short-selling for such traders. **Proposition 2** shows that the expected excess return approaches zero when $N \to \infty$. This limiting exchange achieves the informationally efficient price that leaves no arbitrage opportunity for outsiders.

Proposition 2 also shows that expected trading volume *per trader* is increasing in N with a finite limit. Hence, the model provides a rationale for "liquidity externality," where liquidity here is measured by trading volume per trader. This result may appear to support a large exchange. However, this is not consistent with traders' welfare because trading volume does not go with gain from trade.

The last part of **Proposition 2** states that the large exchange with the informative price and large trading volume does not provide gain from trade. Why is there no gain from trade even though trading volume is large? It is easy to show that $p = \frac{\beta_s}{\beta_p}\left(\overline{s} - \frac{\rho}{\tau_\varepsilon}\overline{x}\right)$ approaches v as N goes to infinity. In other words, the price fully reveals the risky asset return. This necessarily decreases gain from trade, since (i) there is no return on private signal, and (ii) there is no risk to be shared. Recall that the marginal valuation of the asset (left-hand side of (8)) is $E_i[v] - \rho Var_i[v](q_i + x_i)$. We know that conditional variance of v goes to zero in the limit. It is easy to verify $E_i[v]$ goes to v. Hence, information sharing in a large exchange destroys dispersion of marginal valuation created by different endowment levels. In the limit, even though the physical allocation is still dispersed at interim stage, there is no dispersion in valuation.

To show where the hump-shape comes from, I analyze information aggregation through prices in more detail. In particular, I show that there is a sense in which risk sharing can be

[6] Hatchondo, Krusell and Schneider (2008) provide a similar argument.

[7] Of course, equilibrium behavior of informed traders would be different with uninformed traders in the market. This analysis is left for future work.

"most efficiently" done in the exchange of the intermediate size. Recall that each trader's profit is $\pi_i = v(q_i + x_i) - pq_i = vx_i + (v-p)q_i$. When traders submit orders, they face a constrained portfolio allocation problem: Given a fixed position x_i, which provides a risky return v, choose q_i, which provides a different risky return $v - p$.

What matters for this portfolio choice is correlation between two returns v and $v - p$. It can be shown that $v - p = (1 - \frac{\beta_s}{\beta_p})v - \frac{\beta_s}{\beta_p}\left(\overline{\varepsilon} - \frac{\rho}{\tau_\varepsilon}\overline{x}\right)$, with $\frac{\beta_s}{\beta_p} < 1$ and $\lim_{N\to\infty}\frac{\beta_s}{\beta_p} = 1$. Therefore, when the exchange size increases, there are two forces that affect correlation between v and $v - p$. On the one hand, variance of $\overline{\varepsilon} - \frac{\rho}{\tau_\varepsilon}\overline{x}$ decreases, which makes correlation stronger. On the other hand, price becomes more unbiased $(1 - \frac{\beta_s}{\beta_p} \to 0)$, which makes correlation weaker. The second force dominates in a large exchange and the marginal benefit of trading the "new security" with return $v - p$ decreases. This discussion is summarized in the following lemma.

Lemma 4 (correlation) *$Corr[v, v-p]$ is increasing in N for $N \leq \frac{1}{I}$, decreasing in N for $N \geq \frac{1}{I}$, and approaches zero as $N \to \infty$. Correlation is decreasing in I for all N.*

Lemma 4 implies that if we draw this ex ante correlation as a function of exchange size, it has a single peak at $N = \frac{1}{I}$ and approaches zero in the right tail. It can be shown that the slope of this hump-shape is positive and decreasing for small N, negative and decreasing for intermediate N, and negative and approaching zero for large N. The last part of **Lemma 4** states that higher I lowers this hump-shape. In the proof it is shown that only $\frac{\beta_p}{\beta_s}$ is relevant for this correlation pattern. Hence, the result holds both under the strategic and price-taking cases.

To fully characterize gain from trade for a different exchange size, it is convenient to relate the strategic case to the price-taking case. As explained above, the hump-shape comes from the correlation between the price and return of the asset, regardless of traders being strategic or not. In the strategic case, however, it will be shown that the hump-shape is shifted downward and to the right relative to the price-taking case. The following lemma relates the price-taking case to the strategic case.

Lemma 5 (ICE) *In a trade equilibrium, ICE profit is* $\Pi_i^{ice} = \widetilde{\lambda}\Pi_{iH}^{ice} + \left(1-\widetilde{\lambda}\right)\Pi_{iPT}^{ice} = \widetilde{\lambda}\left\{E_i[v]x_i - \frac{\rho}{2\tau}x_i^2\right\} + \left(1-\widetilde{\lambda}\right)\left\{px_i + \frac{\tau}{2\rho}(E_i[v]-p)^2\right\}$, *where* $\widetilde{\lambda} \equiv \left(\frac{\lambda\tau}{\lambda\tau+\rho}\right)^2 = \left(\frac{N}{N-1}I + \frac{1}{N-1}\right)^2$.

Lemma 5 decomposes ICE profit in an intuitive way. The first term is related to the value of holding entire endowment, denoted by Π_{iH}^{ice}. This appears to be exactly the same as a no-trade ICE ($E_i[v]x_i - \frac{\rho}{2\tau}x_i^2$), with the only difference in the information contained in $E_i[v]$ and τ due to information sharing through prices. Recall that a price-taking assumption is captured by setting $\lambda = 0$, which implies $\widetilde{\lambda} = 0$. Hence, the second term in ICE profit is related to the value under the price-taking assumption, denoted by Π_{iPT}^{ice}. This is the sum of the value of endowments evaluated at a market-clearing price and the value of speculative trading.[8]

Actual ICE profit Π_i^{ice} is a weighted sum of Π_{iH}^{ice} and Π_{iPT}^{ice}, where the weight is given by $\widetilde{\lambda} \equiv \left(\frac{\lambda\tau}{\lambda\tau+\rho}\right)^2$. As the exchange size decreases, price impact λ increases, and ICE profit approaches $E_i[v]x_i - \frac{\rho}{2\tau}x_i^2$. This is the case because traders trade less aggressively and hold more of initial endowments in a small market. On the other hand, as the exchange size increases, the weight $\widetilde{\lambda}$ goes down toward I^2.

The decomposition is useful to study ex ante gain from trade. From the proof of **Proposition 2(c)**, ACE profit is $\Pi^{ace} = \frac{1}{2\rho}\log\left(\det(I_3 + 2\rho\Sigma A)\right)$, where I_3 is a 3-by-3 identity matrix and $A \equiv \widetilde{\lambda}A_H + (1-\widetilde{\lambda})A_{PT}$ is also a 3-by-3 matrix. Hence, ACE profit is monotonic transformation of $I_3 + 2\rho\Sigma A = \widetilde{\lambda}(I_3 + 2\rho\Sigma A_H) + (1-\widetilde{\lambda})(I_3 + 2\rho\Sigma A_{PT})$. The 3-by-3 matrices A_H and A_{PT} contain coefficients of quadratic representation of ICE profit Π_{iH}^{ice} and Π_{iPT}^{ice}.

First, note that ACE profit of holding endowment is same with no-trade ACE profit, because the allocation is same. Therefore, $\det(I_3 + 2\rho\Sigma A_H) = 1 - \frac{\rho^2}{\tau_v\tau_x}$, which corresponds to no-trade case. Second, the value associated with price-taking case $I_3 + 2\rho\Sigma A_{PT}$ is greater than no-trade case. Hence, as long as $\widetilde{\lambda} > 0$, the gain from trade is lower in the strategic case than in the price-taking case. As $\widetilde{\lambda}$ approaches 1, ACE profit approaches no-trade ACE

[8] px_i appears in Π_{iPT}^{ice} since, because of CARA utility, an initial position is irrelevant for the choice of new position except its informational use in $E_i[v]$.

profit, while as $\widetilde{\lambda}$ decreases, ACE profit approaches price-taking ACE profit. Since the weight $\widetilde{\lambda}$ is decreasing in N, the shape of gain from trade is the down-and-rightward shift of the price-taking gain from trade.

Finally, the shape of gain from trade is summarized in the following proposition:

Proposition 3 (hump-shape) (a) *The exchange size that maximizes gain from trade $N^*(I)$ is decreasing in I for small I, and increasing in I for I close to* 1*.* (b) *Gain from trade is increasing in N for $N \leq N^*(I)$, decreasing in N for $N \geq N^*(I)$, and decreasing in I uniformly over N.* (c) *In the limit $N \to \infty$, gain from trade converges to zero by the speed of $\frac{1}{N}$.*

Properties presented in **Proposition 3** are graphically shown in the next subsection.

3.3 Numerical evaluation

This subsection provides numerical evaluation of gain from trade. **Figure 4** shows each trader's gain from trade $GFT \equiv \Pi^{ace} - \Pi^{ace}_{NT}$ as a function of N for different levels of $I < 1$. **Figure 5** shows a total gain from trade $TG \equiv N(\Pi^{ace} - \Pi^{ace}_{NT})$ similarly.

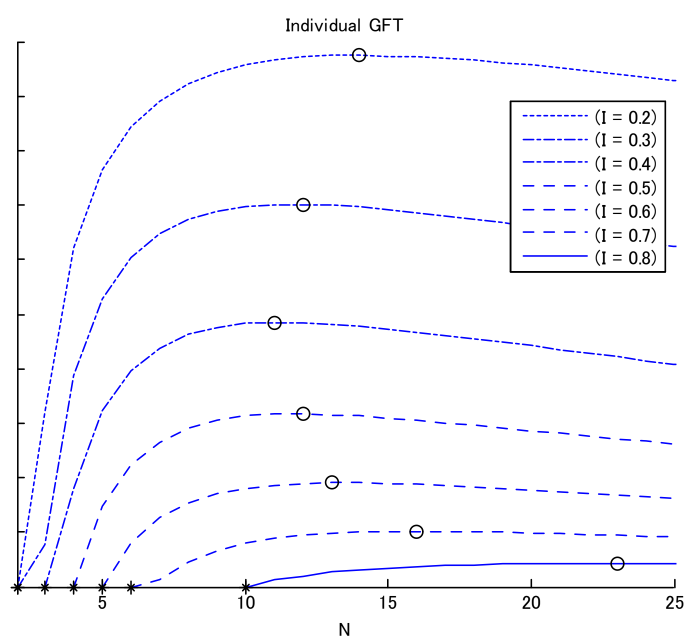

Figure 4

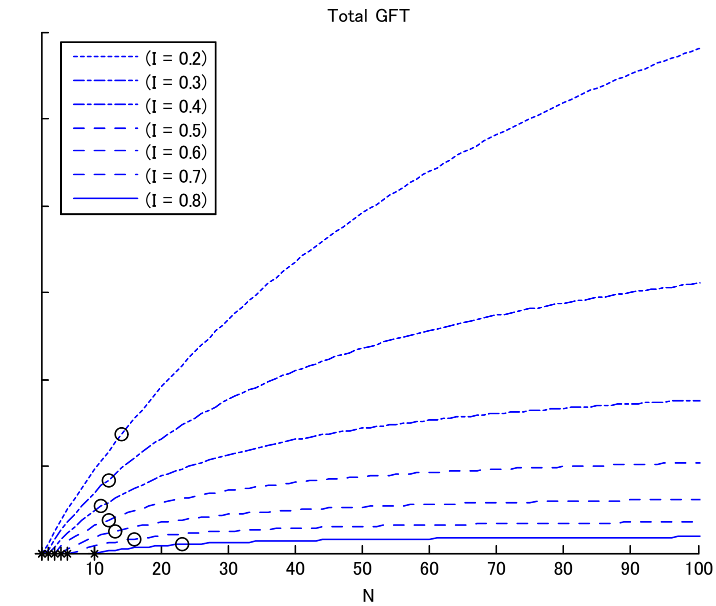

Figure 5

Circle markers represent $N^*(I)$ in **Figures 4** and **5**. From **Proposition 3(c)**, total gain from trade has an upper bound, which is illustrated in **Figure 5**. The exchange size which maximizes an individual gain from trade may seem "too small" since a total gain from trade is still increasing at that size. I study the welfare implication of exchange size in the next section.

Figure 6 shows the relation between the price-taking and strategic case for $I = 0.7$. As shown in the previous subsection, gain from trade in the strategic case is lower and distorted to the right relative to the price-taking case.

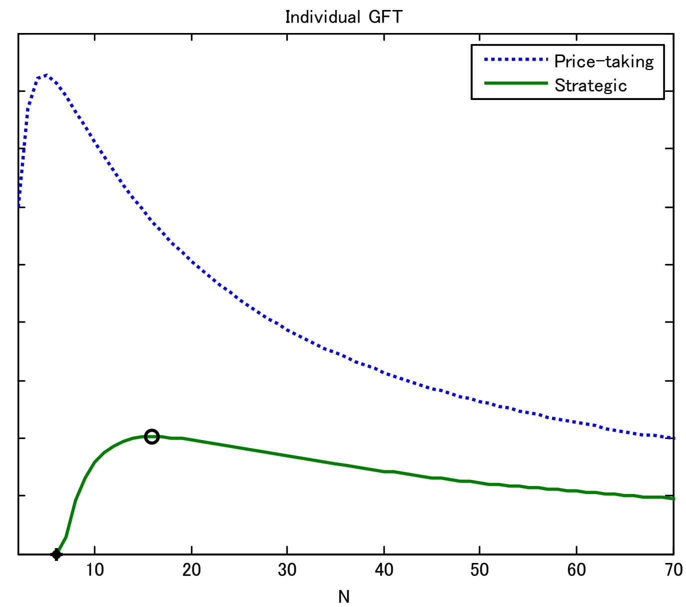

Figure 6

4 Market formation game

This section endogenizes a number and size of exchanges through a market formation game. A market formation game is played at the ex ante stage, i.e., before realization of any private information. Hence, all traders are identical. The game can be interpreted as follows. Suppose that a new financial asset is created. Traders, probably dealers, know that they can potentially make profit by trading the asset with their customers. However, to the extent that the amount of trading with customers is uncertain, dealers will be exposed to risk associated with the random position of the risky asset. Since there is potential for mutually beneficial risk sharing among these dealers, exchanges may make profits by providing a service that facilitates trading (i.e., a system that conducts order submission, market-clearing, and settlement).

To focus on the study of competition among exchanges, I abstract from ownership structure of exchanges, and assume that exchanges are independent, for-profit institutions. Each exchange must cover an operation cost that is proportional to the number of traders in the exchange. Each exchange charges a fixed entry fee for each trader in order to maximize profit. I assume that each exchange can charge only a fixed fee, and can not charge a variable fee per usage (e.g., per trading volume or transaction value). These exchanges are the only places where traders can play the trading game analyzed in the preceding sections. Also, it is assumed that each trader can enter only one exchange, and that the trading game in each exchange occurs simultaneously so that there is no information leakage across exchanges.

I borrowed the view that exchanges compete to get traders "on board" from the literature on two-sided markets.[9] Note that, for exchanges to be non-trivial players in the market formation game, it must be that traders cannot directly interact to realize gains from trading. The validity of this assumption depends on who, in reality, owns the technology that implements trading. For example, see Reedy (2008) for the type of technology that typical dark pools provide.

The market formation game proceeds in two steps. First, each exchange simultaneously sets an entry fee taking other exchanges' fees as given. Second, each trader simultaneously decides which exchange to join, or not to join any exchanges, taking all fees and other traders' choices as given. It is common knowledge that, after this market formation game, the trading game is played independently in each exchange.

First, I introduce a measure of total surplus for a given number of exchanges.

Definition 3 (total surplus) *Total surplus with K exchanges is* $\sum_{j=1}^{K} \{GFT(N_j) - c\} N_j$, *where N_j is the size of exchange j and c is the operation cost of each exchange per trader.*

Given this definition, I ask the following questions: Suppose there are $\overline{N}$ potential traders in the economy. How many exchanges should be set up to maximize total surplus? How large should each exchange be? Can exchange competition achieve the maximum total surplus? In the subsections that follow, I analyze three cases: (i) a monopoly exchange ($K = 1$), (ii) two exchanges ($K = 2$), and (iii) free entry of exchanges (endogenous K).

[9] See Rochet and Tirole (2006).

4.1 Monopoly exchange

When there is only one exchange, traders come to the exchange as long as the expected gain from trade exceeds an entry fee. Monopoly exchanges can choose the fee and number of traders that achieve maximum profit. The monopoly exchange's problem is

$$\max_{\phi} \ (\phi - c)N \tag{16}$$

$$\text{s.t. } GFT(N) - \phi \geq 0, \tag{17}$$

$$N \leq \overline{N}. \tag{18}$$

where ϕ is the fee, c is the operation cost per trader, N is the number of traders who join the exchange, and $\overline{N}$ is the potential number of traders in the economy. The notation for gain from trade makes it explicit that it depends on the size of the exchange. Assuming that the maximum gain from trade exceeds the marginal cost c, the constraint (17) binds in equilibrium. Hence, solution is

$$N_m = \arg\max_{N \leq \overline{N}} \ (GFT(N) - c)N, \tag{19}$$

$$\phi_m = GFT(N_m). \tag{20}$$

Figure 7 illustrates the monopoly exchange size (19) and fee (20), assuming that c is 5% of maximum gain from trade and that $\overline{N}$ is large enough to avoid corner solution $N_m = \overline{N}$.

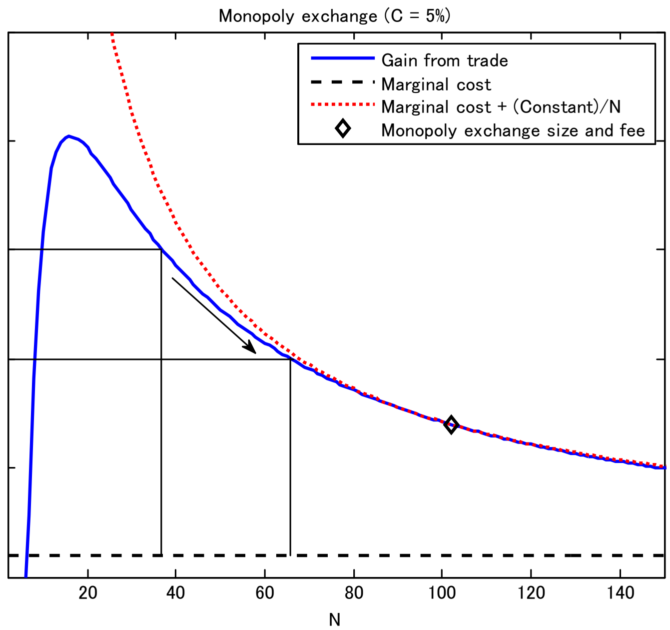

Figure 7

The monopoly exchange's profit for each pair of fee and market size is represented by a rectangle on a marginal cost line (black dashed line), keeping its top-right corner on a gain from trade curve (blue hump-shape). Hence, maximum profit is achieved when a rectangular hyperbola (red dotted line) defined above the marginal cost line is tangent to gain from trade. From **Proposition 3(c)**, this monopoly size exists and finite for any $c > 0$. Also, the monopoly exchange size is greater than $N^* \equiv \arg\max_{N \leq \overline{N}} GFT(N)$, since $(GFT(N_m) - c)N^* \leq (GFT(N^*) - c)N^* \leq (GFT(N_m) - c)N_m$.

The monopoly exchange extracts all the surplus as its profit. Traders are made indifferent between paying the entry fee and not entering the exchange. Notice that the objective in (19) is total surplus for $K = 1$. Hence, under the restriction that there can be only one exchange, total surplus is maximized. A lower or higher fee decreases total surplus compared to the monopoly exchange's profit.

However, there are potential problems with the monopoly exchange. If there are more traders than the monopoly exchange is willing to accept, some traders are excluded from trading. Also, as long as $N^* < \overline{N}$, traders who join the monopoly exchange create negative externality, because the exchange size is at a downward sloping part of the gain from trade curve. These problems point to the possibility that competition among multiple exchanges may provide a more socially beneficial outcome.

4.2 Two exchanges

To simplify the analysis, I assume that two exchanges must be separate in the following sense: Once traders join one exchange, they can submit orders only in that exchange and can not observe the other exchange's activity. This may seem unrealistic if you think about major exchanges such as the NYSE. However, independent market organizers such as Instinet actually provide services that (i) limit the membership to a certain group of institutional traders, (ii) keep the information within their members, and (iii) minimize information sharing among the members. The model should be interpreted as competition among such trading service providers.

Two exchanges A and B compete by each setting fee ϕ_A and ϕ_B. Exchange A's problem is

$$\max_{\phi_A} \ (\phi_A - c)N_A, \tag{21}$$

$$\text{s.t. } GFT(N_A) - \phi_A \geq \max \ \{GFT(N_B) - \phi_B, 0\} \tag{22}$$

$$N_A + N_B \leq \overline{N}, \tag{23}$$

where N_A and N_B are the numbers of traders in each exchange. Exchange B's problem is symmetric. Constraint (22) reflects traders' choice at the second stage.

Note that at the second stage there is always a trivial equilibrium where no trader participates in any exchange no matter what the fees are. Also, asymmetric equilibria where all traders concentrate in the one exchange always exist. Since this is due to the self-fulfilling belief of traders that "I'd be the only one in that exchange," I assume that each exchange has the minimum number of loyal traders to exclude these cases. Still, there may be non-trivial asymmetric equilibria where each exchange has a different fee and size. For simplicity, I focus on a symmetric equilibrium where two exchanges set the same fee and have the same size. In this equilibrium, the one group of traders chooses exchange A and the other group with the same number of traders chooses exchange B.[10] No trader has an incentive to change an exchange in equilibrium. When the number of potential traders is odd, one exchange can be larger than the other by one trader, but it still needs to be the case that no trader has an incentive to change an exchange.

The choice of ϕ_A may affect both N_A and N_B through traders' choice at the second stage. Suppose that (22) binds and $GFT'(N_A) > 0$ and $GFT'(N_B) > 0$. If the exchange A lowers its fee, it will attract some traders. This will make the exchange A more attractive, since $GFT(N_A)$ increases. Moreover, if (23) is also binding, the exchange B becomes less attractive, since $GFT(N_B)$ decreases. This argument does not hold if two exchanges are competing on the downward sloping part of the gain from trade curve. If $GFT'(N_A) < 0$, lowering ϕ_A may attract some traders, but not all, because additional traders lower $GFT(N_A)$ and possibly raise $GFT(N_B)$.

In sum, slope of $GFT(N)$ changes nature of competition:

$$GFT'(N_A) \geq 0 \text{ ... Lower } \phi_A \text{ will seize all traders.}$$
$$GFT'(N_A) < 0 \text{ ... Lower } \phi_A \text{ will gather some traders.}$$

The following lemma characterizes competition between two exchanges.

Lemma 6 (two exchanges) [*Case* 1. $N_m \leq \frac{\overline{N}}{2}$] *Both exchanges set monopoly fee and each has* N_m *traders.* $\overline{N} - 2N_m$ *traders are excluded from trading.* [*Case* 2. $N^* \leq \frac{\overline{N}}{2} < N_m$] *Both set* $\phi = \min\{c - GFT'(\frac{\overline{N}}{2})N, GFT(\frac{\overline{N}}{2})\}$ *and have* $\frac{\overline{N}}{2}$ *traders. If* $\phi = GFT(\frac{\overline{N}}{2})$, *exchanges extract all the surplus. If* $\phi = c - GFT'(\frac{\overline{N}}{2})\overline{N}$, *exchanges and traders share the total surplus.* [*Case* 3. $\frac{\overline{N}}{2} < N^*$] *There is no symmetric equilibrium with traders' participation.*

[10] There may also be a mixed strategy equilibrium where traders randomly choose one exchange. This analysis is complicated due to the hump-shaped gain from trade, and is left for future work.

Lemma 6 shows that two exchanges do not necessarily compete away profits even though the competition is of Bertrand type. When the number of potential traders is large enough, two exchanges can afford each being a monopoly. Competition has some bite when the number of potential traders is not large enough, such that both exchanges cannot enjoy a monopoly size. Still, exchanges can make profits as long as a shared market size is on the downward sloping part of the gain from trade curve.

Exchanges' market power comes from negative externality among traders. Knowing that raising its fee will not lose all the traders due to negative externality, each exchange sets fee strategically. In Case 3, competition becomes so severe that two exchanges cannot survive (at least as a symmetric equilibrium). Positive externality among traders creates strong incentive to cut the fee, since a slightly lower fee allows each exchange to steal all of the traders from the other exchange. **Figures 8** and **9** illustrate Cases 1 and 2.

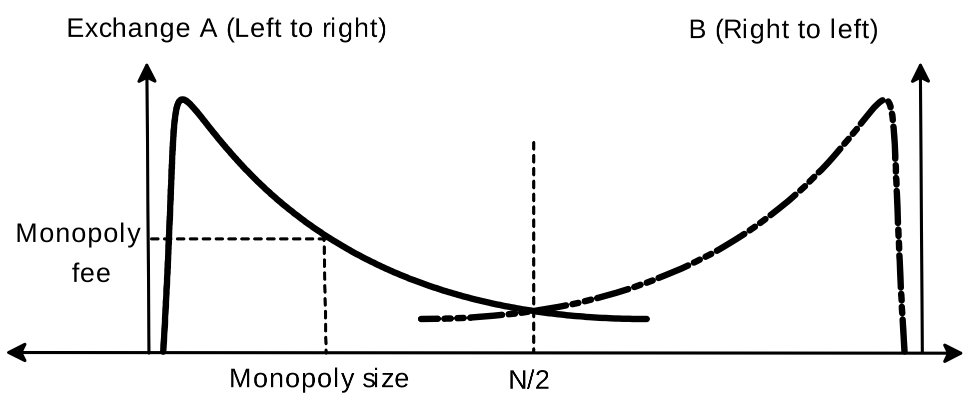

Figure 8

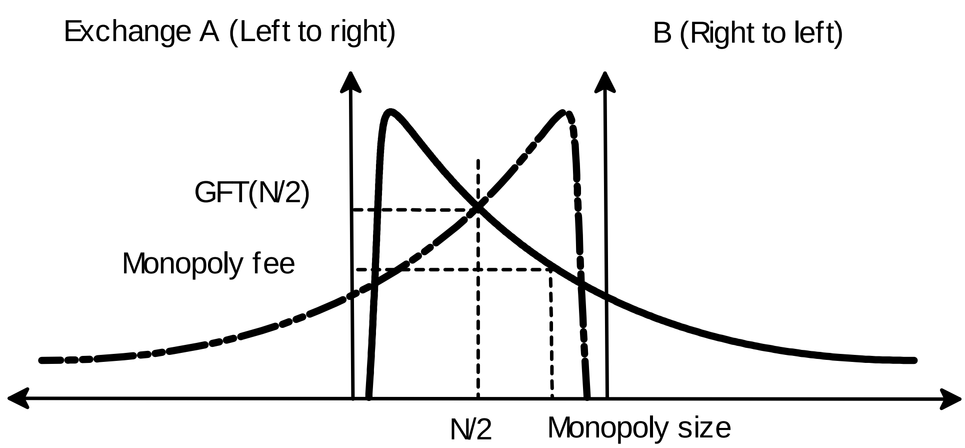

Figure 9

In **Figures 8** and **9**, the length between the two vertical axes represents the number of potential traders $\overline{N}$. The size of exchange A is measured from the left axis to the right, while the size of exchange B is measured to the opposite direction. In **Figure 8**, the number of potential traders is large enough that each exchange can behave as if it is a monopoly. In equilibrium, each exchange could expand its size by lowering its fee, but does not have such incentive.

In **Figure 9**, since the number of potential traders is not large enough, the shared market size is smaller than the monopoly size. In this case, there may be an incentive to set the fee lower than $GFT(\frac{\overline{N}}{2})$, because expanding market size by attracting more traders might increase profit. However, **Lemma 6** shows that as long as the slope of gain from trade curve at the shared market size is negative, the equilibrium fee is higher than marginal cost, and thus exchanges earn profit.

4.3 Free entry of exchanges

The analysis with two exchanges extends to a general case with K exchanges. I focus on Case 2 where each exchange sets $\phi = \min\{c - 2GFT'(\frac{\overline{N}}{K})\frac{\overline{N}}{K}, GFT(\frac{\overline{N}}{K})\}$ and has $\frac{\overline{N}}{K}$ traders. When the number of exchanges K satisfies $N^* < \frac{\overline{N}}{K}$, the fee exceeds marginal cost and thus exchanges make profits. Entry of exchanges seeking profits will make a shared market size $\frac{\overline{N}}{K}$ smaller, and there will be no profit when $\frac{\overline{N}}{K} = N^*$.[11]

It is shown below that free entry of exchanges, which achieves $\frac{\overline{N}}{K} = N^*$, also achieves the maximum total surplus.

Proposition 4 (free entry) *For given $\overline{N}$, free entry induces $K = \frac{\overline{N}}{N^*}$ exchanges. All $\overline{N}$ traders participate in one of the exchanges and the exchanges earn zero profit. This achieves*

[11]Strictly speaking, entry could stop after $K = \left\lfloor \frac{\overline{N}}{N^*} \right\rfloor$ exchanges have entered, where $\left\lfloor \frac{\overline{N}}{N^*} \right\rfloor$ denotes the largest integer that does not exceed $\frac{\overline{N}}{N^*}$. A potential entrant could rationally expect that its entry would trigger fierce competition in Case 3.

the maximum total surplus $\{GFT(N^*) - c\}\overline{N}$.

Proposition 4 shows that, ignoring discreteness, free entry of exchanges can achieve the maximum total surplus. At the social optimum, there are $K = \frac{\overline{N}}{N^*}$ exchanges. The exchanges set $\phi = c$ and compete away profit.

First, the resulting market structure from free entry is characterized by segmentation and opaqueness. Each exchange size N^* is small relative to the size of a monopoly exchange, and each preserves a high degree of asymmetric information with respect to the asset return. Moreover, none of the exchanges and traders has an incentive to create a larger market, or share information.

Second, as the number of exchanges increases, the equilibrium fee does not necessarily decrease. It can be shown that $\phi = c - 2GFT'(\frac{N}{K})\frac{N}{K}$ decreases in K if and only if $-\frac{GFT''(\frac{N}{K})}{GFT'(\frac{N}{K})}\frac{N}{K} < 1$.[12] If $GFT''(\frac{N}{K}) > 0$, the equilibrium fee might increase with more exchanges. This could happen when entry moves a shared market size from a relatively flat part of the gain from trade curve to a steep part. However, a sufficiently large number of exchanges eventually drives the fee down to the marginal cost.

5 Conclusion

This paper presented a two-stage game framework to study the formation of exchanges. The second stage trading game captured the three aspects of financial markets: (i) strategic behavior, (ii) private knowledge regarding the uncertain return of assets, and (iii) private endowments. It was shown that gain from trade is hump-shaped in the number of traders. Each trader affects the trading game in two ways: he (i) increases the size of the risk sharing pool, and (ii) increases the informativeness of prices. The key for the hump-shape is that the second force decreases the benefit of the first force through the Hirshleifer Effect. In the small exchange, the first factor dominates and gain from trade increases in the number of

[12] Discrete version is $\frac{GFT'(\frac{N}{K+1})}{GFT'(\frac{N}{K})}\frac{K}{K+1} < 1$.

traders. On the other hand, the Hirshleifer Effect becomes dominant in the large exchange and gain from trade decreases in the number of traders. In other words, traders face a trade-off between risk sharing and information sharing. Since the latter tends to decrease the former, either one must be sacrificed for the other.

The first stage marker formation game formalized the interaction among exchanges and traders. It was shown that the negative externality among traders due to the Hirshleifer Effect may enable exchanges to enjoy market powers. The increase in the number of exchanges makes each exchange size smaller and mitigates the negative externality. With free entry of exchanges, each trader's gain from trade is maximized, because prices don't reveal much information when each exchange is small. The prediction of the model is consistent with historical evidence that some traders have preferred less transparency, and that market structures have evolved in that direction. In particular, the model may provide a rationale for the recent development of "dark pools".

To extend the model presented in this paper, an environment in which traders have correlated endowments was analyzed in a separate paper. For example, when the economy experiences a macro shock (e.g., crash of a bubble in some markets), professional traders may face a selling or buying pressure from their customers who are reacting to this shock. Within this environment, it can be shown that there could be three types of trade equilibria, ranked in decreasing order by the amount of information sharing and in increasing order by trading volume. With this extension, however, the parameter region for the existence of linear equilibria becomes significantly small. These results may explain the high volatility and fragility of financial markets when traders face a macro shock.[13]

Another extension of the model is to incorporate heterogeneity among traders, such as different risk aversions, introduction of uninformed traders, and unequal access to some trading technology (for example, some traders can be restricted to use market orders). This will clarify who benefits from an exchange and who does not.

[13] In the limit where N goes to infinity, the extended model is equivalent to that of Ganguli and Yang (2009).

More generally, this raises the interesting issue of how different types of trading games attract different types of traders. As documented in Biais and Green (2007), trading did not just disappear when a centralized bond market declined—it migrated to the OTC market, where much less information is shared. Thus, the OTC market structure could also be an endogenous reaction to avoid welfare-reducing information sharing. To further investigate this issue, one could explore a unified framework in which different trading games can be compared. A mechanism design approach might be useful in exploring this possibility.

Finally, it would be interesting to apply this model in a macroeconomic context. While the current paper focused on the welfare of an exchange economy, the welfare implication for a production economy is not obvious. On the one hand, reduced risk sharing due to information sharing may slow down production and innovation. On the other hand, informative prices may guide better productive decisions. Another application of the model is to study a central bank's market operations by adding the central bank as a trader in financial markets. In the model environment, a central bank can change not only the mean of return (interest rate) but also the variance of return. The current analysis suggests that this market operation may have a significant impact on the viability of financial markets. The nature of the central bank's private information, asset position, and public announcement may be important in such an analysis. These are all left for future work.

APPENDIX A

Proof of Lemma 1

When each trader submits zero quantity for all price levels (i.e., $q_i(p) = 0$), any price can clear the market. Following the market-clearing rule, $p = 0$ is announced. Given that other traders submit $q_i(p) = 0$, each trader is indifferent between submitting $q_i(p) = 0$ and any price-contingent order. If he submits a price-contingent order, it will only change the price so that his net trading is zero, and he still receives $\pi_i = vx_i$. Also, each trader strictly prefers submitting $q_i(p) = 0$ to any non-zero market order, because the latter deviation would give him negative infinite utility.[14] $E_i[v]$ and τ are obtained by Bayes' rule. Without trading, each trader's interim utility is $E_i[-\exp(-\rho v x_i)]$. Using a characteristic function of normal distribution, $\Pi^{ice}_{iNT} = E_i[v]x_i - \frac{\rho}{2\tau}x_i^2$. Write this in a matrix form $\Pi^{ice}_{iNT} = [s_i, x_i]A_{NT}[s_i, x_i]'$, where $A_{NT} \equiv \begin{bmatrix} 0 & \frac{\tau_\varepsilon}{2\tau} \\ & -\frac{\rho}{2\tau} \end{bmatrix}$.[15] To obtain ACE profit, apply the following formula: $E[\exp(-\rho X'AX)] = \frac{1}{\sqrt{\det\Sigma\det(\Sigma^{-1}+2\rho A)}}$, where X follows normal distribution $N(0,\Sigma)$. Substituting $\Sigma_{NT} = \begin{bmatrix} \frac{1}{\tau_v}+\frac{1}{\tau_\varepsilon} & 0 \\ & \frac{1}{\tau_x} \end{bmatrix}$ and A_{NT} yields $\det\Sigma_{NT}\det(\Sigma^{-1}_{NT}+2\rho A_{NT}) = 1-\frac{\rho^2}{\tau_v\tau_x}$.

Proof of Lemma 2

From the market-clearing condition, $0 = \sum_{j\neq i} q_j + q_i = \beta_s \sum_{j\neq i} s_j - \beta_x \sum_{j\neq i} x_j - (N-1)\beta_p p + q_i$. Observe that (s_i, x_i, p) provides the same amount of information about v as (s_i, x_i, h_i) does, where $h_i \equiv \frac{(N-1)\beta_p p - q_i}{(N-1)\beta_s} = v + \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j - \frac{\beta_x}{\beta_s}\frac{1}{N-1}\sum_{j\neq i} x_j = v + \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j - \frac{\beta_x}{\beta_s}\frac{1}{N-1}\sum_{j\neq i} x_j$. By Bayes' rule, $\tau = \tau_v + \tau_\varepsilon + (Var[v|h_i])^{-1}$, where $(Var[v|h_i])^{-1} = (N-1)\left[\tau_\varepsilon^{-1} + \left(\frac{\beta_x}{\beta_s}\right)^2 \tau_x^{-1}\right]^{-1} = (N-1)\tau_\varepsilon\varphi$. Plug this and $h_i = \frac{(N-1)\beta_p p - q_i}{(N-1)\beta_s} = \frac{N\beta_p}{(N-1)\beta_s}p - \frac{1}{N-1}s_i + \frac{\beta_x}{(N-1)\beta_s}x_i$ into $E_i[v] = \frac{\tau_\varepsilon}{\tau}s_i + \frac{(Var[v|h_i])^{-1}}{\tau}h_i$ to get the last result.

Proof of Proposition 1

From $\frac{\beta_x}{\beta_s} = \frac{1}{1-\varphi}\frac{\rho}{\tau_\varepsilon} - \frac{\varphi}{1-\varphi}\frac{\beta_x}{\beta_s}$, $K \equiv \frac{\beta_x}{\beta_s} = \frac{1}{1-\varphi(K)}\frac{\rho}{\tau_\varepsilon} - \frac{\varphi(K)}{1-\varphi(K)}K$, where $\varphi(K) \equiv \frac{1}{1+K^2\frac{\tau_\varepsilon}{\tau_x}}$. This defines a cubic equation $\left\{K^2 + \frac{\tau_x}{\tau_\varepsilon}\left(1+\frac{1}{(N-1)\beta_s^2}\frac{\tau_\varepsilon}{\tau_z}\right)\right\}\left(K - \frac{\rho}{\tau_\varepsilon}\right) = 0$ with a unique solution $K = \frac{\rho}{\tau_\varepsilon}$. It follows that $\varphi = \frac{1}{1+\frac{\rho^2}{\tau_\varepsilon\tau_x}} = \frac{I}{1+I}$. Similarly, solving $\frac{\beta_p}{\beta_s} = \frac{1}{1-\varphi}\frac{\tau}{\tau_\varepsilon} - \frac{N\varphi}{1-\varphi}\frac{\beta_p}{\beta_s}$ for $\frac{\beta_p}{\beta_s}$ shows that $\frac{\beta_p}{\beta_s} = \frac{\tau}{\tau_\varepsilon + \varphi(N-1)\tau_\varepsilon}$. Define $C_s \equiv corr(s_i, s_j) = \frac{\tau_\varepsilon}{\tau_\varepsilon+\tau_v}$ and $N_s \equiv 1 + C_s(N-1)$. Note that

[14] The market-clearing rule has some bite here. If traders are not punished by infinite prices, there may be incentive to throw away positive endowments. This happens if $x_i > 0$ and $\frac{\tau_\varepsilon}{\rho}s_i < x_i$. If an alternative market-clearing rule is used for this "free disposal" to be allowed in equilibrium, no-trade utility will be higher, and the level of gain from trade will be lower, but the main results in this paper remain unchanged.

[15] Throughout the paper, a symmetric matrix is represented by its upper triangular part.

$\frac{\tau_\varepsilon \varphi N}{\tau} = \frac{\tau_\varepsilon}{\tau_v+\tau_\varepsilon+\tau_\varepsilon(N-1)\varphi}\frac{NI}{1+I} = \frac{C_s(1+I)}{1+N_sI}\frac{NI}{1+I} = \frac{C_sNI}{1+N_sI}$ and $\frac{\tau_\varepsilon(1-\varphi)}{\tau} = \frac{\tau_\varepsilon}{\tau_v+\tau_\varepsilon+\tau_\varepsilon(N-1)\varphi}\frac{1}{1+I} = \frac{C_s}{1+N_sI}$. Hence, $\frac{\beta_p}{\beta_s} = \frac{1}{C_s}\frac{1+N_sI}{1+NI}$. To show $\beta_s = \frac{\tau_\varepsilon}{\rho}\left(1-2\varphi-\frac{1}{N-1}\right)$, plug $\lambda = \frac{1}{(N-1)\beta_p} = \frac{\rho\beta_s+\varphi\tau_\varepsilon N}{(N-2)\tau\beta_s}$ into $\beta_s = \frac{(1-\varphi)\tau_\varepsilon}{\lambda\tau+\rho}$ and simplify it. This establishes the existence and uniqueness of a fixed point defined by (12),(13),(14). From the second order condition (9), $2\lambda+\frac{\rho}{\tau} > 0 \Leftrightarrow \beta_s = \frac{(1-\varphi)\tau_\varepsilon}{\lambda\tau+\rho} > 0 \Leftrightarrow \beta_s = \frac{\tau_\varepsilon}{\rho}\left(1-2\varphi-\frac{1}{N-1}\right) > 0 \Leftrightarrow \varphi < \frac{1}{2}\left(1-\frac{1}{N-1}\right)$. Use $\varphi = \frac{I}{1+I}$ to obtain $I < 1-\frac{2}{N}$. If $I = 1-\frac{2}{N}$, then $\beta_s = \beta_x = \beta_p = 0$, i.e., a unique symmetric linear equilibrium is a no-trade equilibrium.

Proof of Corollary 2

The result for $E_i[v]$ and τ is trivial. A trader's objective is $E_i[v](q_i+x_i)-\frac{\rho}{2\tau}(q_i+x_i)^2-pq_i = -\frac{\rho}{2\tau}(q_i+x_i)^2-pq_i$. After taking price impact into account, it can be shown that $q_i = -\left(1-\frac{1}{N-1}\right)x_i-\frac{\tau_v}{\rho}\left(1-\frac{1}{N-1}\right)p$. This derivation is similar to the proof of **Proposition 1**, so it is omitted here. ICE profit is $[x_i,p]A_{NS}[x_i,p]'$, where $A_{NS} \equiv \frac{1}{2N^2}\begin{bmatrix} -\frac{\rho}{\tau_v} & (N-1)(N+1) \\ & \frac{\tau_v}{\rho}(N-1)(N+1)\end{bmatrix}$.

Using A_{NS} and $\Sigma_{NS} = \begin{bmatrix} \frac{1}{\tau_x} & -\frac{\rho}{\tau_v\tau_x}\frac{1}{N} \\ & \frac{\rho^2}{\tau_v\tau_x}\frac{1}{N}\end{bmatrix}$, obtain $\det\Sigma_{NS}\det(\Sigma_{NS}^{-1}+2\rho A_{NS}) = 1-\frac{\rho^2}{\tau_v\tau_x}K(N,\frac{\rho^2}{\tau_v\tau_x})$, where $K(N,\frac{\rho^2}{\tau_v\tau_x}) \equiv \frac{(N-1)^2(N+1)}{N^4}\frac{\rho^2}{\tau_v\tau_x}+\frac{N^2+N-1}{N^3}$. Note that $K(1,\frac{\rho^2}{\tau_v\tau_x}) = 1$. It is straightforward to show that $K(N,\frac{\rho^2}{\tau_v\tau_x})$ is decreasing in N for $\frac{\rho^2}{\tau_v\tau_x} \leq 1$. Finally, $\lim_{N\to\infty} K(N,\frac{\rho^2}{\tau_v\tau_x}) = 0$.

Proof of Lemma 3

In a trade equilibrium, a market-clearing condition yields $p = \frac{\beta_s}{\beta_p}\left(\overline{s}-\frac{\rho}{\tau_\varepsilon}\overline{x}\right)$. Then $q_i = \beta_s\left\{s_i-\overline{s}-\frac{\rho}{\tau_\varepsilon}(x_i-\overline{x})\right\} = \beta_s\{\varepsilon_i-\frac{\rho}{\tau_\varepsilon}x_i-(\overline{\varepsilon}-\frac{\rho}{\tau_\varepsilon}\overline{x})\}$ and $TV = \frac{\beta_s}{2}\sum\left|\varepsilon_i-\frac{\rho}{\tau_\varepsilon}x_i-(\overline{\varepsilon}-\frac{\rho}{\tau_\varepsilon}\overline{x})\right|$. Since $\beta_s = \frac{\tau_\varepsilon}{\rho}\left(1-2\varphi-\frac{1}{N-1}\right)$, β_s approaches zero from above as $I < 1-\frac{2}{N}$ approaches equality. Also, $\lambda = \frac{1}{(N-1)\beta_p} = \frac{\tau-\tau_v}{(N-1)\tau\beta_s}$ goes to infinity.

Proof of Proposition 2

(a) Recall that $p = \frac{\beta_s}{\beta_p}\left(\overline{s} - \frac{\rho}{\tau_\varepsilon}\overline{x}\right)$. Hence, $\frac{\beta_p}{\beta_s}p = v + \overline{\varepsilon} - \frac{\rho}{\tau_\varepsilon}\overline{x}$ is an unbiased signal of v with precision $N\tau_\varepsilon\varphi$. By Bayes rule, $E\left[v|p\right] = \frac{N\tau_\varepsilon\varphi}{\tau_v + N\tau_\varepsilon\varphi}\frac{\beta_p}{\beta_s}p$, where $\frac{\beta_p}{\beta_s} = \frac{\tau}{\tau_\varepsilon + \varphi(N-1)\tau_\varepsilon}$ from the proof of **Proposition 1**. Substituting $\tau = \tau_v + \tau_\varepsilon\left(1 + (N-1)\varphi\right)$ and simplifying it yields $\eta = \frac{\tau_v + \tau_\varepsilon(1+(N-1)\varphi)}{\tau_v\left(1+\frac{1-\varphi}{N\varphi}\right) + \tau_\varepsilon(1+(N-1)\varphi)} \leq 1$. With $\varphi = \frac{I}{1+I}$, $\frac{N\tau_\varepsilon\varphi}{\tau_v + N\tau_\varepsilon\varphi} = \left(1 + \frac{1-C_s}{C_s}\frac{1+I}{NI}\right)^{-1} = \frac{C_s NI}{1+I+C_s(NI-1-I)} = \frac{C_s NI}{1-C_s+N_s I}$. Recall that $\frac{\tau}{\tau_\varepsilon + \varphi(N-1)\tau_\varepsilon} = \frac{1}{C_s}\frac{1+N_s I}{1+NI}$. Hence, $E\left[v|p\right] = \frac{NI}{1+NI}\frac{1+N_s I}{1-C_s+N_s I}p$. It is easy to verify $\lim_{N\to\infty}\frac{NI}{1+NI}\frac{1+N_s I}{1-C_s+N_s I} = 1$.

(b) From the proof of **Lemma 3**, $\frac{TV}{N} = \frac{\beta_s}{2N}\sum\left|\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i - \frac{1}{N}\sum(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i)\right|$. Note that $\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i - \frac{1}{N}\sum(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) = \frac{N-1}{N}(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) - \frac{N-1}{N}\frac{1}{N-1}\sum_{j\neq i}(\varepsilon_j - \frac{\rho}{\tau_\varepsilon}x_j)$. Hence, $E\left[\frac{TV}{N}\right] = \frac{\beta_s}{2N}\frac{N-1}{N}\sum E\left[\left|(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) - \frac{1}{N-1}\sum_{j\neq i}(\varepsilon_j - \frac{\rho}{\tau_\varepsilon}x_j)\right|\right] = \frac{\beta_s}{2}\frac{N-1}{N}E\left[\left|(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) - \frac{1}{N-1}\sum_{j\neq i}(\varepsilon_j - \frac{\rho}{\tau_\varepsilon}x_j)\right|\right]$. Here, $Var\left[(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) - \frac{1}{N-1}\sum_{j\neq i}(\varepsilon_j - \frac{\rho}{\tau_\varepsilon}x_j)\right] = \frac{N}{N-1}\left\{\tau_\varepsilon^{-1} + (\frac{\rho}{\tau_\varepsilon})^2\tau_x^{-1}\right\} = (\frac{\rho}{\tau_\varepsilon})^2\frac{N}{N-1}\frac{1}{\tau_x}(1+I)$. Apply the following fact: $E[|X|] = \sqrt{\frac{2}{\pi}\tau_X^{-1}}$, where X is a normal random variable with mean zero and variance τ_X^{-1}, to obtain $E\left[\left|(\varepsilon_i - \frac{\rho}{\tau_\varepsilon}x_i) - \frac{1}{N-1}\sum_{j\neq i}(\varepsilon_j - \frac{\rho}{\tau_\varepsilon}x_j)\right|\right] = \frac{\rho}{\tau_\varepsilon}\sqrt{\frac{2}{\pi}\frac{N}{N-1}\frac{1}{\tau_x}(1+I)}$. Substituting $\beta_s = \frac{\tau_\varepsilon}{\rho}\left(1 - 2\varphi - \frac{1}{N-1}\right)$ and $1 - 2\varphi - \frac{1}{N-1} = \frac{(1-I)N-2}{(1+I)(N-1)}$ yields $E\left[\frac{TV}{N}\right] = \frac{(1-I)N-2}{\sqrt{N(N-1)2\pi\tau_x(1+I)}}$. This is increasing in N and $\lim_{N\to\infty}\frac{(1-I)N-2}{\sqrt{N(N-1)2\pi\tau_x(1+I)}} = \frac{1-I}{\sqrt{2\pi\tau_x(1+I)}}$.

(c) Write ICE profit given in **Lemma 5** in a matrix form. Let $X_i \equiv [s_i, x_i, p]'$. ICE profit of holding endowment is $\Pi_{iH}^{ice} = X_i'A_H X_i$ and that of price-taking case is $\Pi_{iPT}^{ice} = X_i'A_{PT}X_i$, where $A_H \equiv \frac{C_s}{2(1+N_s I)}\begin{bmatrix} 0 & 1 & 0 \\ & -\frac{\rho}{\tau_\varepsilon}(1-I) & \frac{NI}{C_s}\frac{1+N_s I}{1+NI} \\ & & 0 \end{bmatrix}$, $A_{PT} \equiv \frac{C_s}{2(1+N_s I)}\frac{I}{1+I}\begin{bmatrix} \frac{\rho}{\tau_x} & 1 & -\frac{\rho}{\tau_x}\frac{1+N_s I}{1+NI}\frac{1}{C_s} \\ & \frac{\tau_x}{\rho} & \frac{1+NI+NI^2}{I}\frac{1+N_s I}{1+NI}\frac{1}{C_s} \\ & & \frac{\rho}{\tau_x}\left(\frac{1+N_s I}{1+NI}\frac{1}{C_s}\right)^2 \end{bmatrix}$

and $\Sigma = \begin{bmatrix} \frac{1}{\tau_v C_s} & 0 & \frac{1}{\tau_v}\frac{N_s}{N}\frac{1+NI}{1+N_s I} \\ & \frac{1}{\tau_x} & -\frac{\rho}{\tau_\varepsilon\tau_x}\frac{C_s}{N}\frac{1+NI}{1+N_s I} \\ & & \frac{1}{N}\left(\frac{N_s}{\tau_v C_s} + \frac{1}{\tau_\varepsilon I}\right)\left(\frac{1+NI}{1+N_s I}C_s\right)^2 \end{bmatrix}$. First, $\Pi_{iH}^{ice} = E_i[v]x_i - \frac{\rho}{2\tau}x_i^2 = \frac{\tau_\varepsilon(1-\varphi)}{\tau}s_i x_i + (\frac{\rho\varphi}{\tau} - \frac{\rho}{2\tau})x_i^2 + \frac{\varphi N}{\varphi N + 1 - \varphi}x_i p$. Derivation of the coefficient on $s_i x_i$ is as given in the proof of **Proposition 1**: $\frac{\tau_\varepsilon(1-\varphi)}{\tau} = \frac{C_s}{1+N_s I}$. The other coefficients can be evaluated similarly. Next, $\Pi_{iPT}^{ice} = px_i + \frac{\tau}{2\rho}(E_i[v] - p)^2 = \frac{\tau}{2\rho}k_1^2 s_i^2 + \frac{\tau}{\rho}k_1 k_2 s_i x_i + \frac{\tau}{\rho}k_1(k_3 - 1)s_i p + \frac{\tau}{2\rho}k_2^2 x_i^2 + \{\frac{\tau}{\rho}k_2(k_3 - 1) + 1\}x_i p + \frac{\tau}{2\rho}(k_3 - 1)^2 p^2$. Substitute k_1, k_2, k_3 in **Lemma 2** and algebra yields the result.

Finally, $\Sigma_{11} = Var[s_i]$, $\Sigma_{22} = Var[x_i]$, $\Sigma_{33} = Var[p]$, $\Sigma_{12} = cov[s_i, x_i]$, $\Sigma_{13} = cov[s_i, p]$, $\Sigma_{23} = cov[x_i, p]$. Σ_{11}, Σ_{12}, Σ_{22} are trivial. I show the derivation of Σ_{13}. Σ_{23} and Σ_{33} are evaluated similarly. $cov[s_i, p] = \frac{\beta_s}{N\beta_p}(Var[s_i] + (N-1)Cov[s_i, s_j]) = \frac{\beta_s}{\beta_p}(\tau_v^{-1} + \frac{1}{N}\tau_\varepsilon^{-1})$. From the proof of **Proposition 1**, $\frac{\beta_s}{\beta_p} = C_s\frac{1+NI}{1+N_sI}$. Since $(\tau_v^{-1} + \frac{1}{N}\tau_\varepsilon^{-1}) = \frac{1}{N\tau_v}(\frac{1}{C_s} + N - 1) = \frac{N_s}{C_s\tau_v N}$, the desired result is obtained. ICE profit is a weighted sum of these two, $\Pi_i^{ice} = \widetilde{\lambda} X_i' A_H X_i + (1-\widetilde{\lambda}) X_i' A_T X_i = X_i'\{\widetilde{\lambda} A_H + (1-\widetilde{\lambda}) A_T\} X_i$. Finally, it is straight forward to show that $\lim_{N\to\infty} A_H = \lim_{N\to\infty} A_{PT} = \begin{bmatrix} 0 & 0 & 0 \\ & 0 & \frac{1}{2} \\ & & 0 \end{bmatrix}$ and $\lim_{N\to\infty}\Sigma = \begin{bmatrix} \frac{1}{\tau_v C_s} & 0 & \frac{1}{\tau_v} \\ & \frac{1}{\tau_x} & 0 \\ & & \frac{1}{\tau_v} \end{bmatrix}$. Also, $\lim_{N\to\infty}\widetilde{\lambda} = I^2$. Calculating $\Pi^{ace} = \frac{1}{2\rho}\log\left(\det\Sigma\det\left(\Sigma^{-1} + 2\rho A\right)\right)$ with these matrices shows that $E\left[\exp\left(-\rho\lim_{N\to\infty}\Pi_i^{ice}\right)\right] = \exp\left(-\rho\Pi_{NT}^{ace}\right)$. By dominated convergence theorem, the result holds.

Proof of Lemma 4

First, $v - p = (1-\frac{\beta_s}{\beta_p})v - \frac{\beta_s}{\beta_p}\left(\overline{\varepsilon} - \frac{\rho}{\tau_\varepsilon}\overline{x}\right)$. Hence, $Var[v-p] = (1-\frac{\beta_s}{\beta_p})^2\tau_v^{-1} + \frac{1}{N}(\frac{\beta_s}{\beta_p})^2\{\tau_\varepsilon^{-1} + (\frac{\rho}{\tau_\varepsilon})^2\tau_x^{-1}\}$. $Cov[v, v-p] = (1-\frac{\beta_s}{\beta_p})\tau_v^{-1}$. Thus, $Corr[v, v-p] = \frac{(1-\frac{\beta_s}{\beta_p})}{\sqrt{(1-\frac{\beta_s}{\beta_p})^2 + \frac{\tau_v}{N}(\frac{\beta_s}{\beta_p})^2\{\tau_\varepsilon^{-1} + (\frac{\rho}{\tau_\varepsilon})^2\tau_x^{-1}\}}} = \frac{1}{\sqrt{1+\frac{\tau_v}{\tau_\varepsilon}\frac{1+I}{N}(\frac{\beta_p}{\beta_s}-1)^{-2}}}$. From **Proposition 1**, $\frac{\beta_p}{\beta_s} - 1 = \frac{1}{C_s}\frac{1+N_sI}{1+NI} - 1 = \frac{1-C_s}{C_s}\frac{1+I}{1+NI} = \frac{\tau_v}{\tau_\varepsilon}\frac{1+I}{1+NI}$. This establishes $Corr[v, v-p] = \frac{1}{\sqrt{1+\frac{\tau_\varepsilon}{\tau_v}H(N)}}$, where $H(N) \equiv \frac{(1+IN)^2}{(1+I)N}$. Correlation is monotonically and inversely related to $H(N)$. The second part follows form the fact that $H'(N) = \frac{1}{1+I}\frac{2(1+IN)IN-(1+IN)^2}{N^2} = \frac{1}{1+I}\frac{(IN)^2-1}{N^2}$ and $\lim_{N\to\infty} H'(N) = \frac{I^2}{1+I}$. Finally, $\frac{\partial H}{\partial I} = -\frac{(1+IN)(1-IN)^2}{(1+I)^2} \leq 0$ with equality only at $I = \frac{1}{N}$.

Proof of Lemma 5

First, $\Pi_i^{ice} = (E_i[v] - p)(q_i + x_i) - \frac{\rho}{2\tau}(q_i + x_i)^2 + px_i$. Substitute $q_i = \frac{\tau}{\lambda\tau+\rho}(E_i[v] - p_i - \frac{\rho}{\tau}x_i)$ from (10) and simplify it to obtain the result. To show $\frac{\lambda\tau}{\lambda\tau+\rho} = \frac{N}{N-1}I + \frac{1}{N-1}$, first note that $\frac{\lambda\tau}{\lambda\tau+\rho} = \frac{\tau}{\tau+\rho(N-1)\beta_p}$ from the definition of λ. Next, $\frac{\beta_p}{\beta_s} = \frac{\tau}{\tau_\varepsilon+\varphi(N-1)\tau_\varepsilon}$ and

$\beta_s = \frac{\tau_\varepsilon}{\rho}\left(1 - 2\varphi - \frac{1}{N-1}\right)$ from **Proposition 1**. This implies $\rho\beta_p = \frac{\tau(1-2\varphi-\frac{1}{N-1})}{1+\varphi(N-1)}$. Hence, $\frac{\lambda\tau}{\lambda\tau+\rho} = \frac{1}{1+\frac{(1-2\varphi-\frac{1}{N-1})(N-1)}{1+\varphi(N-1)}} = \frac{1+\varphi(N-1)}{1+\varphi(N-1)+(1-2\varphi)(N-1)-1} = \frac{1+\varphi(N-1)}{(N-1)(1-\varphi)} = \frac{\varphi+\frac{1}{N-1}}{1-\varphi} = I + \frac{1+I}{N-1} = \frac{N}{N-1}I + \frac{1}{N-1}$.

Proof of Proposition 3

(a) As I approaches 0, $N^*(I)$ becomes arbitrarily large, hence it is decreasing in I for small I. Recall that both $\widetilde{\lambda}$ and lower bound $\underline{N}(I) \equiv \frac{2}{1-I}$ in (15) are increasing in I. As I increases, higher $\widetilde{\lambda}$ lowers GFT. As I approaches 1, $\underline{N}(I)$ goes to infinity and so does $N^*(I)$.

(b) From **Corollary 2**, GFT is monotonically increasing in N with $I = 0$. When $I > 0$, precision increases linearly in the market size since $\tau = \tau_v + \tau_\varepsilon + \varphi(N-1)\tau_\varepsilon$. This creates uni-modal shape of GFT.

(c) To show the convergence speed of GFT, it suffices to show that $\det(I_3+2\rho\Sigma A_{PT}) = 1 - \frac{\rho^2}{\tau_v\tau_x} + O(\frac{1}{N})$. Element-by-element evaluation of ΣA_{PT} shows that $\Sigma A_{PT} = \begin{bmatrix} O(\frac{1}{N}) & O(0) & O(\frac{1}{N}) \\ O(\frac{1}{N}) & O(\frac{1}{N}) & O(0) \\ O(\frac{1}{N^2}) & O(0) & O(\frac{1}{N}) \end{bmatrix}$. I show the derivation of (2,3) element. The other elements can be derived similarly. $(\Sigma A_{PT})_{(2,3)} = \frac{\varphi}{2}\frac{C_s}{1+N_sI}\left(\frac{1+NI+NI^2}{\tau_x C_s I}\frac{1+N_sI}{1+NI} - \frac{1}{\tau_x C_s IN}\frac{1+N_sI}{1+NI}\right) = \frac{\varphi}{2\tau_x I}\left(1 + \frac{NI^2}{1+NI} - \frac{1}{1+NI}\frac{1}{N}\right) = \frac{1}{2\tau_x} - O(\frac{1}{N^2})$. It can be shown that $(\Sigma A_{PT})_{(3,2)} = \frac{1}{2\tau_v} + O(\frac{1}{N})$. Hence, $\det(I_3 + 2\rho\Sigma A_{PT}) = 1 - \frac{\rho^2}{\tau_v\tau_x} + O(\frac{1}{N})$.

Proof of Lemma 6

Case 1 is trivial. In Case 2, first note that the fee cannot be greater than $GFT(\frac{\overline{N}}{2})$. If the fee were higher than $GFT(\frac{\overline{N}}{2})$, each exchange would have a market size smaller than $\frac{\overline{N}}{2}$, which is smaller than N_m. Hence, there is an incentive to set a lower fee to attract additional traders. Given binding (22) and (23), the exchange A's problem becomes $\max_{N_A} (GFT(N_A) - GFT(\overline{N} - N_A) + \phi_B - c)N_A$. First order condition is $(GFT'(N_A) + GFT'(\overline{N} - N_A))N_A + GFT(N_A) - GFT(\overline{N} - N_A) + \phi_B - c = 0$. By imposing symmetry, obtain $2GFT'(\frac{\overline{N}}{2})\frac{\overline{N}}{2} + \phi - c = 0$. Hence, $\phi = c - GFT'(\frac{\overline{N}}{2})\overline{N}$. If this is greater than $GFT(\frac{\overline{N}}{2})$, traders' participation constraint binds and $\phi = GFT(\frac{\overline{N}}{2})$, and each exchange extracts all the surplus. If $\phi = c - GFT'(\frac{\overline{N}}{2})\overline{N} < GFT(\frac{\overline{N}}{2})$, total surplus $GFT(\frac{\overline{N}}{2})\overline{N}$ is split between the exchanges and traders, since $\phi \in [c, GFT(\frac{\overline{N}}{2}))$. As long as two exchanges have equal size (or with one trader difference when $\overline{N}$ is odd), traders have no incentive to change their exchange since doing so will lower the net benefit (when $\overline{N}$ is odd, traders in the larger exchange are indifferent). In Case 3, first note that any fee higher than c cannot be in equilibrium, since there is incentive to lower the fee and seize traders. Even if two exchanges set $\phi = c$, equal market size cannot be in equilibrium, since traders have an incentive to change the exchange.

Proof of Proposition 4

Focus on the case $N^* \leq \frac{\overline{N}}{K} < N_m$. Similarly to the two exchanges case, a typical exchange's first order condition, combined with symmetry, yields $2GFT'(\frac{\overline{N}}{K})\frac{\overline{N}}{K} + \phi - c = 0 \Leftrightarrow \phi = c - 2GFT'(\frac{\overline{N}}{K})\frac{\overline{N}}{K}$. If this is greater than $GFT(\frac{\overline{N}}{K})$, traders' participation constraint binds, exchanges set $\phi = GFT(\frac{\overline{N}}{K})$ and extract all the surplus. If $\phi = c - 2GFT'(\frac{\overline{N}}{K})\frac{\overline{N}}{K} < GFT(\frac{\overline{N}}{K})$, total surplus $GFT(\frac{\overline{N}}{K})\overline{N}$ is split between K exchanges and traders, since $\phi \in [c, GFT(\frac{\overline{N}}{K}))$. Profit is zero only if $K = \frac{\overline{N}}{N^*}$. When K exchanges equally split $\overline{N}$ traders, total surplus is given by $K\left\{GFT\left(\frac{\overline{N}}{K}\right) - c\right\}\frac{\overline{N}}{K}$. Maximize this with respect to K: $\max_K K\left\{GFT\left(\frac{\overline{N}}{K}\right) - c\right\}\frac{\overline{N}}{K}$. This is equivalent to $\max_K GFT\left(\frac{\overline{N}}{K}\right)$. Hence, $\frac{\overline{N}}{K} = N^* \Leftrightarrow K = \frac{\overline{N}}{N^*}$. Maximized total surplus is $\{GFT(N^*) - c\}\overline{N}$.

APPENDIX B

In this appendix, I compare the trading game in this paper with an alternative game with exogenous noise trading. The key distinction between the two models is whether a random supply of the asset (liquidity supply) is endogenous or exogenous. In this paper, the source of the random supply is the traders' private endowments. In the alternative model, it is noise trading z and endowments are known to be zero ($x_i = 0$ for all i, $\tau_x = \infty$).

Accordingly, a normal random variable z with $\tau_z < \infty$ appears in a market clearing condition: $\sum_i q_i(p) + z = 0$. These two models are compared under two environments: strategic and price-taking. Table 1 summarizes the four cases.

Table 1

	Endowments	Noise trading
Strategic	Case A	Case B
Price-taking	Case C	Case D

Case A corresponds to the model in this paper. In Section 3, I showed a comparative statics where trading decreases to zero and the price impact goes to infinity with noisy prices. In Case B, these symptoms arise asymptotically as noise trading disappears.[16] It can be shown that these symptoms are possible also only asymptotically in a price-taking environment. In Case C, these are possible only in the limit where endowment uncertainty vanishes. In Case D, these symptoms happen only in the limit where noise trading vanishes.[17] Thus, it is a combination of strategic behavior and endogenous supply of the asset that makes the condition (15) relevant.

To see why (15) is relevant only in Case A, it is useful to compare the four cases in terms of endogenous parameters (β_s, φ). The difference between a strategic model and a price-taking model is in β_s;

$$\beta_s^S = \frac{\tau_\varepsilon}{\rho}\left(1 - 2\varphi - \frac{1}{N-1}\right) \tag{24}$$

$$\beta_s^{PT} = \frac{\tau_\varepsilon}{\rho}\left(1 - \varphi\right). \tag{25}$$

The difference between a model with endowments and a model with noise trading is in φ;

$$\varphi^E = \frac{1}{1 + \frac{\rho^2}{\tau_\varepsilon \tau_x}} \tag{26}$$

[16] See Kyle (1989) for details.

[17] Proof is in the previous version of this paper available at my webpage.

$$\varphi^N = \frac{1}{1 + \frac{1}{(N-1)\beta_s^2}\frac{\tau_\varepsilon}{\tau_z}}. \tag{27}$$

Superscripts S, PT, E and N denote strategic, price taking, endowments, and noise trading, respectively. Condition (15) is relevant only with a combination $\left(\beta_s^S, \varphi^E\right)$.

First, β_s^S is necessary since β_s^S can take negative value with sufficiently high information sharing (φ being close to 1), while β_s^{PT} is always non negative for any $\varphi \in [0,1]$. Second, φ^E is necessary since it mutes a feedback effect through noise trading in the following sense. As β_s becomes large, price becomes more informative, i.e., φ^N becomes large. However, there is a feedback effect from φ^N which makes β_s small. Because the amount of noise trading is not affected by the traders' choice of β_s, large β_s means the market-clearing price provides more information. Since the informational value of prices is higher, the traders put more weight on prices, and thus, less weight on private signals. This partially offsets the initial increase in β_s. In a model with endowments, there is no such feedback. The absence of feedback is due to the balance between risk sharing and speculation: $\frac{\beta_x}{\beta_s} = \frac{\rho}{\tau_\varepsilon}$. When φ^E becomes small, β_s becomes large, but so does β_x. The first effect is the same as the information-based effect described above, while the second effect is based on *the risk sharing motive*–the less informative a market-clearing price is, the greater the need for sharing risk against an uncertain payoff. More risk-sharing-based trades imply less informative prices. In equilibrium, the two effects exactly cancel out each other, which leaves φ^E independent of β_s and β_x.

The following simulation illustrates a comparison of the four cases A through D. Note that total supply of the asset is $\sum_{i=1}^{N} x_i$ in Case A and C, while it is z in Case B and D. In the simulation, I set $\tau_x = (N-1)\tau_z$ such that ex ante variance of total supply of the asset is the same from each trader's perspective. In the simulation, N and τ_x are increased, keeping $\frac{\tau_x}{N-1} = 1$ for the model with endowments. For the model with noise trading, N is increased, keeping $\tau_z = 1$. Other parameters are set equally in the two settings: $\tau_\varepsilon = \tau_v = 1$, and $\rho = 5$. Given these numbers, small N satisfies (15) but large N does not. Therefore we observe for Case A that trade decreases to zero and price impact goes to infinity as N approaches some finite value. In the other three cases, β_s goes to zero in the limit as N goes to infinity, but λ also decreases and is finite in the limit. **Figures B1** through **B3** show the behavior of $(\beta_s, \varphi, \lambda)$ as N increases.

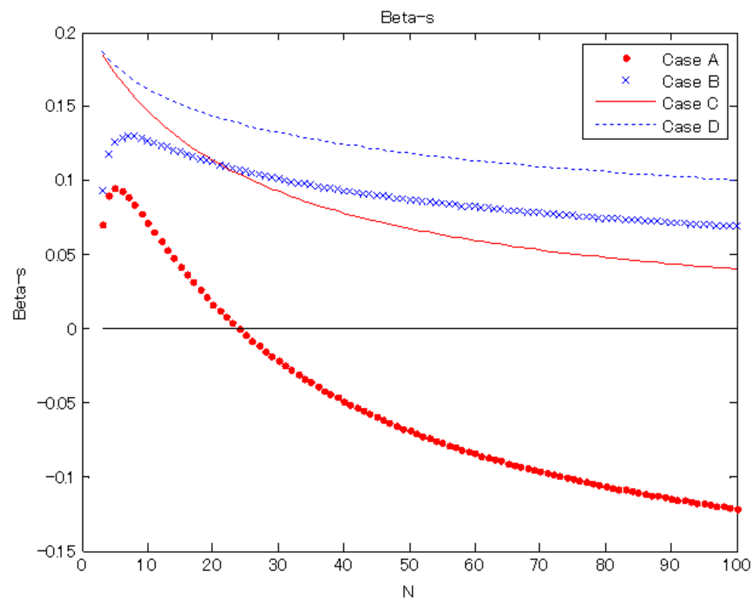

Figure B1

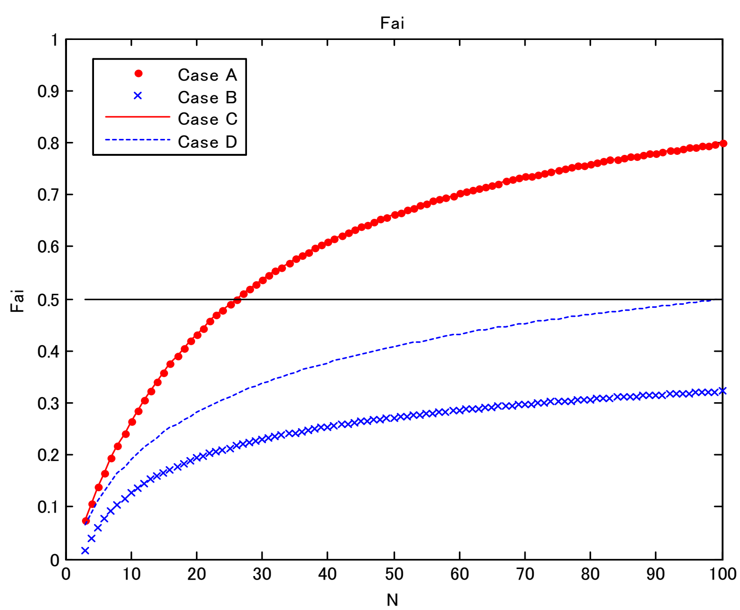

Figure B2

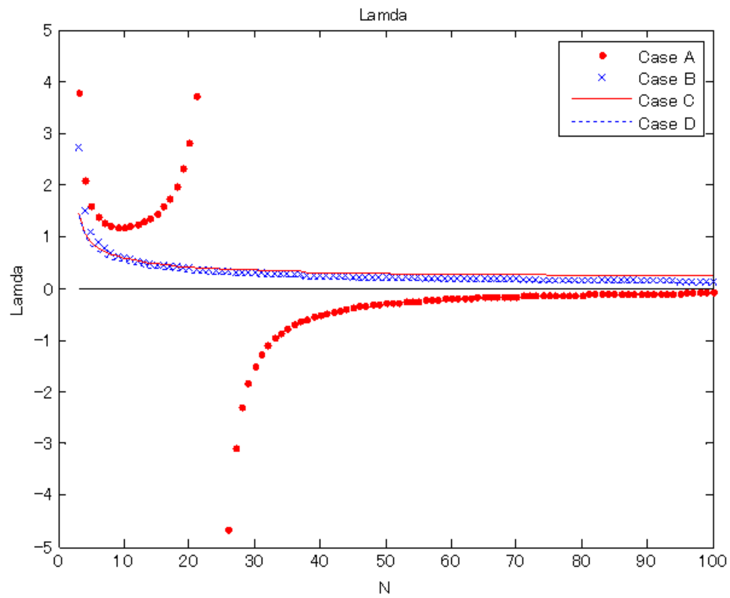

Figure B3

Figure B1 compares β_s in the four cases. It shows that β_s can be negative in Case A for finite N. This is when the second-order condition (9) is violated. For the other three cases, β_s is always positive. Note also that β_s is smaller in the strategic cases (A and B) than in price taking cases (C and D), and more significantly so in a model with endowments (A compared with C) than in a model with noise trading (B compared with D). **Figure B2** compares φ. First it shows that Case A and C have the same φ, which approaches 1 as τ_x goes to infinity (recall that $\frac{\tau_x}{N-1} = 1$ is fixed). This causes problem in Case A when φ approaches value close to $\frac{1}{2}$, violating the second order condition (9), while in Case C the same value of φ never violates (9). The figure also shows that information sharing is greater in a model with endowments than in a model with noise trading, even though ex ante total quantity uncertainty is the same in two markets. As mentioned above, the hedging motive endogenously decreases quantity uncertainty in a model with endowments, while there is no such mechanism with exogenous noise trading.[18] Finally, **Figure B3** shows the behavior of λ. In Case A, the market becomes illiquid for finite N. It can be proved that λ approaches

[18]Theorem 6.1. and 7.2. in Kyle (1989) show that φ approaches $\frac{1}{2}$ in Case B and 1 in Case D.

$\lim \frac{\tau_x}{\rho(N-1)} = 0.2$ in Case C and zero in Case B and D. This shows that, even though ex ante aggregate uncertainty is the same, the liquidity implication can be quite different between a model with endowments and a model with noise trading.

References

[1] Bernardo, A., and K. Judd (1997): "Efficiency of Asset Markets with Asymmetric Information," Working Paper.

[2] Biais, B., and R. C. Green (2007): "The Microstructure of the Bond Market in the 20th Century," Working Paper.

[3] Biais, B., Glosten L. R., and C. Spatt (2005): "Market Microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications," *Journal of Financial Markets*, 8, 217-264.

[4] Bhattacharya, U., Reny, P. J., and M. Spiegel (1995): "Destructive Interference in an Imperfectly Competitive Multi-Security Market," *Journal of Economic Theory*, 65, 136-170.

[5] Bhattacharya, U., and M. Spiegel (1991): "Insiders, Outsiders, and Market Breakdowns," *Review of Financial Studies*, 4, 2, 252-282.

[6] Board, J., C. Sutcliffe (1996): "Trade Transparency and the London Stock Exchange," *European Financial Management*, 2, 3, 355-365.

[7] Diamond, D. W., and R. E. Verrecchia (1981): "Information Aggregation in a Noisy Rational Expectations Economy," *Journal of Financial Economics*, 9, 221-235.

[8] Ganguli, J. V., and L. Yang (2009): "Complementarities, Multiplicity, and Supply Information," *Journal of the European Economic Association*, 7, 1, 90-115.

[9] Gemmill, G. (1996): "Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange under Different Publication Rules," *Journal of Finance*, 51, 5, 1765-1790.

[10] Glosten, L. R. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62, 2, 211-235.

[11] Grossman, S. J., and J. E. Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, 3, 393-408.

[12] Hatchondo, J. C., P. Krusell, and M. Schneider (2008): "Asset Trading and Valuation with Uncertain Exposure," Working Paper.

[13] Hellwig, M. F. (1980): "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 22, 477-498.

[14] Hirshleifer, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561-574.

[15] Kyle, A. S. (1989): "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 56, 317-356.

[16] Mailath, G. J., and G. Noeldeke (2008): "Does Competitive Pricing Cause Market Breakdown under Extreme Adverse Selection?" *Journal of Economic Theory*, 140, 97-125.

[17] Marin, J. M., and R. Rahi (1999): "Speculative Securities," *Economic Theory*, 14, 653-668.

[18] ——— (2000): "Information Revelation and Market Incompleteness," *Review of Economic Studies*, 67, 563-579.

[19] Milgrom, P. R., and N. Stokey (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17-27.

[20] Pagano, M., and A. Roell (1992): "Auction Markets and Dealership Markets. What is the Difference?" *European Economics Review*, 36, 613-623.

[21] Pithyachariyakul, P. (1986): "Exchange Markets: A Welfare Comparison of Market Maker and Walrasian Systems," *Quarterly Journal of Economics*, 69-84.

[22] Ready, M. J. (2008): "Determinant of Volume in Dark Pools," Working Paper.

[23] Reiss, P. C., and I. M. Werner (1998): "Does Risk Sharing Motivate Interdealer Trading?" *Journal of Finance*, 53, 5, 1657-1703.

[24] ——— (2004): "Anonymity, Adverse Selection, and the Sorting of Interdealer Trades," *Review of Financial Studies*, 18, 2, 599-636.

[25] Reny, P. J., and M. Perry (2006): "Toward a Strategic Foundation for Rational Expectations Equilibrium," *Econometrica*, 74, 1231-1269.

[26] Rochet, J. C., and J. Tirole (2006): "Two-sided Markets: a Progress Report," *Rand Journal of Economics*, 37, 3, 645-667.

[27] Spiegel, M., and A. Subrahmanyam (1992): "Informed Speculation and Hedging in a Noncompetitive Securities Market," *Review of Financial Studies*, 5, 2, 307-329.

[28] ——— (2000): "Asymmetric Information and News Disclosure Rules," *Journal of Financial Intermediation*, 9, 363-403.

[29] Subrahmanyam, A. (1991): "Risk Aversion, Market Liquidity, and Price Efficiency," *Review of Financial Studies*, 4, 3, 417-442.