THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Information Aggregation and Optimal Market Size

Kei Kawakami

May 2014

Research Paper Number 1182

ISSN: 0819 2642
ISBN: 978 0 7340 4532 4

Information Aggregation and Optimal Market Size*

Kei Kawakami†

May 25th, 2014

Abstract

This paper studies a rational expectations model of trading where strategic traders face information asymmetries and endowment shocks. We show that negative participation externalities arise due to an endogenous interaction between information aggregation and multiple trading motives. Moreover, the negative externalities are strong enough to make optimal market size finite. In a decentralized process of market formation, multiple markets can survive due to the negative externalities among traders. The model also predicts: (i) that only in a sufficiently large market the equilibrium multiplicity due to self-fulfilling trading motives can arise, (ii) that a high correlation in endowment shocks can make markets extremely illiquid.

Keywords: Asymmetric information, Aggregate shock, Imperfect competition, Market fragmentation, Multiple equilibria, Network externality puzzle, Price impact.

*I thank Antonio Bernardo, Bruno Biais, Bruce Carlin, Chris Edmond, Roger Farmer, Mark Garmaise, Shogo Hamasaki, Christian Hellwig, Ian King, Ichiro Obara, Lee Ohanian, Avanidhar Subrahmanyam, Lawrence Uren, Dimitri Vayanos, Xavier Vives, Pierre-Olivier Weill, Mark Wright, and seminar participants at ANU, DBJ, LSE, TSE, University of Auckland, UCLA and UNSW for comments and suggestions. All errors are mine.

†Department of Economics, University of Melbourne, e-mail: keik@unimelb.edu.au, website: http://sites.google.com/site/econkeikawakami/

1 Introduction

Many financial securities are traded on multiple platforms, rather than in a single large market. A growing body of empirical work suggests that such market fragmentation is a common and robust phenomenon.[1] However, this observation seems to contradict the common intuition that increasing the size of a market benefits market participants by providing better liquidity/depth/price discovery. In the market microstructure literature,[2] a popular view is that security trading should concentrate in a single venue due to positive liquidity externalities while intermediation costs (e.g., transaction costs, search costs etc) may prevent any market from becoming too large. From this perspective, the level of market fragmentation that we observe today looks puzzling if we believe that intermediation costs have been falling. Borrowing a phrase from Madhavan (2000), this is best summarized as *the network externality puzzle*: why is trading for the same security split across multiple trading venues?

We revisit this issue in a rational expectations model of trading with a finite number of risk averse traders, asymmetric information, and shocks to endowments. In the model, traders are ex ante uncertain about both the size of their endowments and the unit value of the endowment. Once the traders participate in a market, the quantity of their endowment is realized and they observe private signals about the unit value of the endowment. After endowment trading takes place according to the limit order trading rule of Kyle (1989), the value of their endowments is realized. There is no additional supply beyond the initial endowments, and risk sharing creates gains from trade.

We show that optimal market size is finite, because negative informational externalities are strong enough to countervail the positive externalities of risk sharing. The intuition behind the main result is as follows. Traders with a high endowment realization are greatly exposed to uncertainty about its value and would like to trade some of this risk away at the interim trading stage. With symmetric information, traders' beliefs are independent

[1] Brown, Mulherin, and Weidenmier (2008) and Cantillon and Yin (2010, 2011) study competition between exchanges. Ready (2009) and O'Hara and Ye (2011) study fragmentation of trades across trading platforms.

[2] See Madhavan (2000), Biais, Glosten, and Spatt (2005) and Vives (2008) for surveys.

of the number of traders, and the gains from risk sharing increase in market size. This is a standard source of positive externalities in financial markets. On the other hand, with asymmetric information, prices that clear the market can be good predictors of the actual value of the endowment. Because the informativeness of prices depend on the amount of *private* information aggregated by prices, traders' beliefs depend on the number of traders. When traders rationally anticipate this, their trading motives are also affected by market size. In particular, the hedging motive may be reduced in a large market where prices are expected to reveal more information, because the revealed risk cannot be traded away (the "Hirshleifer effect", as in Hirshleifer 1971). The problem of reduced risk sharing can be alleviated by having *less informative* prices at the interim trading stage. Thus, a small market where prices are expected to reveal less information may be more capable of providing risk sharing than a large market would. We show that this informational force is strong enough to make the optimal market size finite. **Figure 1** illustrates the main result.

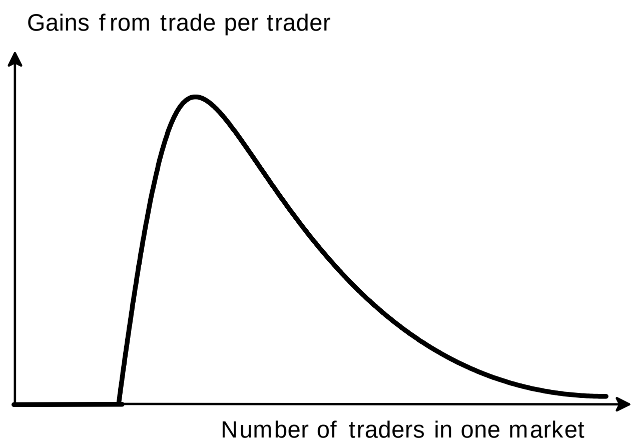


Figure 1. Hump-shaped gains from trade.

The paper makes two more contributions by identifying a novel connection between market size and (i) equilibrium multiplicity, and (ii) market illiquidity as measured by traders' price impact.[3] First, we show that the multiplicity arises due to self-fulfilling trading motives and that equilibria are Pareto-ranked. On the one hand, the balance between the two motives (hedging and speculation) determines price informativeness through market-clearing.

[3] The price impact is called "Kyle's lambda" in the market microstructure literature.

On the other hand, the price informativeness affects the two motives through rational expectations. When this two-way interaction becomes powerful enough, the *conjectured* balance (i.e. in strategies) can self-justify it, unless the *fundamental* balance of the two motives is too extreme (e.g. too high/low risk aversion, little private information to speculate on). Market size plays a crucial role in this context because the two-way interaction becomes strong enough only in a large market where an unobserved common shock to endowments has a large impact on prices. Because multiple equilibria are ordered by the relative importance of hedging to speculation in strategies, an equilibrium with a higher balance of hedging is Pareto-superior. These results allow us to identify an environment more susceptible to this type of multiplicity and to select one equilibrium over others.

Second, we show that the market illiquidity can *increase* in market size, when traders' incentive to learn from prices is high and risk sharing gains are small due to a large aggregate shock to endowments. We call this phenomenon *extreme illiquidity*, because as market size approaches a finite upper bound, the illiquidity increases and gains from trade disappear.[4] Importantly, the extreme illiquidity does not arise either with symmetric information or under a price-taking assumption. Therefore, it is caused by strategic traders learning from prices, and is NOT resolved by increasing the number of traders. We also show that the extreme illiquidity is not due to the Hirshleifer effect. It is a distinct phenomenon.

For all the findings, market size plays a crucial role through information aggregation. Therefore, two essential aspects of our model are: (i) dispersed private information that is of common concern to all, (ii) a finite number of rational traders. The informational impact of market size on welfare cannot be studied in models with a continuum of traders or with noise traders, as they obscure the meaning of market size and welfare.

In the final section, we propose a model of endogenous market formation based on our rational expectations model of trading. We study a market making game where each intermediary sets a fee for entering his market. Because traders are aware how market size affects

[4] Weretka (2011) is a general analysis of the price impact, but in his model it always decreases with the number of traders.

their trading motives and information aggregation in equilibrium, they might benefit more from a market with fewer participants. Therefore, the entry fee is not necessarily bid down to a zero-profit level and multiple markets can survive in equilibrium.

We review the related literature next. Section 2 outlines the model environment. Section 3 studies information aggregation, the multiplicity, and the extreme illiquidity, which all affect the analysis of ex ante welfare. Section 4 studies the optimal market size. We first show that in a symmetric information benchmark the optimal market size is infinite. Section 4.1 shows that the optimal market size is finite when information is asymmetric. Section 4.2 identifies four channels through which market size affects welfare. Section 4.3 studies robustness of the main result. Section 5 proposes a model of endogenous market formation. Section 6 concludes. The Appendix collects additional results and all proofs.

1.1 Related literature

Understanding what economic forces limit market size is important. From an applied viewpoint, the securities industry is seeing the proliferation of new trading platforms, which seem to be causing more trading fragmentation.[5] A growing literature on competing exchanges (Santos and Scheinkman 2001; Foucault and Parlour 2004; Pagnotta 2013) provides explanations for this empirical observation by modelling the ways in which exchanges can differentiate product (e.g. collateral requirements, listing fees, trading speed). This paper does not consider product differentiation, but instead emphasizes information aggregation as a source of negative externalities, and hence provides a complementary rationale for market fragmentation. From a theoretical point of view, a market with a large number of traders is typically viewed as a "thicker" and ideal benchmark. For example, it is customary to consider the limit as the number of traders goes to infinity to obtain asset pricing implications.[6] Considering such a limit may be justified normatively, if a larger market is welfare-superior to a smaller market, or positively if someone has an incentive to create a large market. When

[5] O'Hara and Ye (2011) document the large variety of trading venues in US equity markets.

[6] For example, see Corollary 1 in Madhavan (1992).

negative externalities exist, the large market limit may not be the ideal benchmark.

Schlee (2001) analyses the Hirshleifer effect in a general exchange economy, but he assumes a competitive, public information environment. We study a strategic, private information environment and find a novel connection between market size, information aggregation, and welfare. Finding the Hirshleifer effect in the financial market context is hardly new. For example, Pithyachariyakul (1986) compares the Walrasian system and the monopolistic market-making system and shows that information revealed in the former system creates the welfare trade-off between the two systems. Naik, Neuberger, and Viswanathan (1999) show that trade disclosure can reduce welfare in a dealership environment. Marin and Rahi (2000) analyze a similar trade-off in a security design problem. But none of these models study market size. Our contribution is not that the Hirshleifer effect exists, but instead that it is *strong enough* to countervail the positive externalities of risk sharing and thereby make the optimal market size finite.[7]

Finally, Ganguli and Yang (2009) and Manzano and Vives (2011) study equilibrium multiplicity due to self-fulfilling trading motives in related models. However, their use of a continuum of traders precludes the analysis of market size. Our work complements theirs by providing (i) strategic foundations to the multiplicity, (ii) welfare implications, and (iii) an intuitive interpretation of the conditions that lead to the multiplicity.

2 Model Environment

A noisy rational expectation equilibrium (REE) setup is used to model trading. There are $n+1$ traders indexed by $i \in \{1, ..., n+1\}$ in a given market. Traders have identical preferences and trade endowments that have an unknown unit payoff v. Before trading, each trader i receives (i) the endowment e_i, and (ii) a private signal s_i about v. Both are privately known. There is no additional supply beyond the sum of endowments $\sum_{i=1}^{n+1} e_i$ in the market.

[7] In fact, in Section 4 we show that there are four distinct channels through which market size affects welfare, and the Hirshleifer effect is only one of them.

The random payoff v has two components, v_0 and v_1:

$$v = \sqrt{1-t}v_0 + \sqrt{t}v_1, \;\; t \in [0,1], \tag{1}$$

where v_0 and v_1 are independently normally distributed with mean zero and variance τ_v^{-1}. Each trader observes a signal about v_1:

$$s_i = v_1 + \varepsilon_i,$$

where ε_i is unobserved noise normally distributed with mean zero and variance τ_ε^{-1}. A constant t in (1) measures the degree of information asymmetry. If t is zero, the signal s_i provides no information about the payoff $v = v_0$. When t is positive, there is something to learn from signals. If t is one, the entire payoff $v = v_1$ is subject to information asymmetry.

The endowment e_i also has two components, x_0 and x_i:

$$e_i = \sqrt{1-u}x_0 + \sqrt{u}x_i, \;\; u \in [0,1], \tag{2}$$

where x_0 and $\{x_i\}_{i=1}^{n+1}$ are independently normally distributed with mean zero and variance τ_x^{-1}. The first component, x_0, represents an aggregate shock to the endowment, while the second component, x_i, is an idiosyncratic shock. Trader i knows the realized e_i but does not observe x_0 and x_i separately. A constant u in (2) determines the relative importance of two shocks to the endowment. If u is zero, there is no diversifiable risk and no trade can happen. Therefore, we focus on a case $u > 0$ where a risk-sharing opportunity exists. If u is one, endowments $e_i = x_i$ are i.i.d., and there are maximum potential gains from risk sharing.[8]

We also allow for the cross-sectional correlation in ε_i:

$$\varepsilon_i = \sqrt{1-w}\epsilon_0 + \sqrt{w}\epsilon_i, \;\; w \in [0,1], \tag{3}$$

[8]For example, $1-u$ can be interpreted as a degree of correlation in traders' hedging needs due to a common risk factor (e.g. business cycles).

where ϵ_0 and $\{\epsilon_i\}_{i=1}^{n+1}$ are independently normally distributed with mean zero and variance τ_ε^{-1}. A constant w in (3) determines the cross-sectional correlation in the noise in signals. If w is zero, all the signals $\{s_i\}_{i=1}^{n+1}$ are identical, and the information is symmetric. When w is positive, there is something to learn from other traders' signals. As w increases, the amount of information traders want to learn each other increases.

The parameters $(t, u, w) \in [0, 1]^3$ determine the nature of the trading environment.[9] The parameterization (1), (2), (3) ensures that ex ante variances of v, e_i and s_i do not depend on (t, u, w), but the values of (t, u, w) make a difference at the interim trading stage. For a fixed $w > 0$, a larger t leads to a higher degree of information asymmetry and a smaller u results in a higher importance of the aggregate shock to the endowment. **Figure 2** illustrates this.

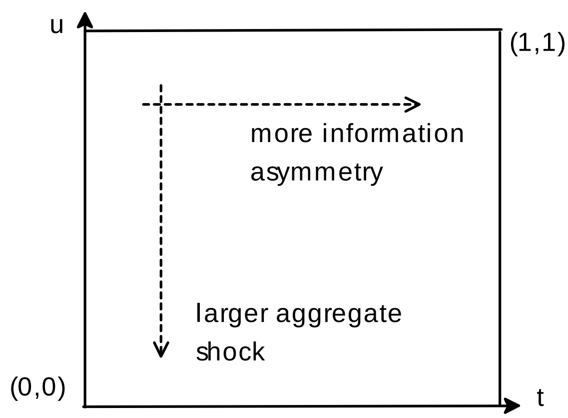


Figure 2. Trading environment parameterized by (t, u).

While the model identifies some phenomena that occur for particular parameter values, our main result is that optimal market size is finite as long as $t > 0$ and $w > 0$.

To summarize, the $4 + 2(n + 1)$ random variables $\left(v_0, v_1, x_0, \{x_i\}_{i=1}^{n+1}, \epsilon_0, \{\epsilon_i\}_{i=1}^{n+1}\right)$ are normally and independently distributed with zero means, and variances

$$Var[v_0] = Var[v_1] = \tau_v^{-1},\ Var[x_0] = Var[x_i] = \tau_x^{-1},\ Var[\epsilon_0] = Var[\epsilon_i] = \tau_\varepsilon^{-1}.$$

[9] Our framework nests informational environments of Diamond and Verrecchia (1981), Pagano (1989), Madhavan (1992), Ganguli and Yang (2009), and Manzano and Vives (2011).

Each trader has an exponential utility function with a risk-aversion coefficient $\rho > 0$

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

where π_i is trader i's profit. Profits are the sum of the payoff from the new position $q_i + e_i$ and the payment or the receipt for trading q_i, and hence $\pi_i = v(q_i + e_i) - pq_i$. After observing the private information $H_i = (e_i, s_i)$, each trader chooses her order $q_i(p; H_i)$.

Following Kyle (1989), we characterize an equilibrium in demand functions. Because an order can be explicitly conditioned on a market-clearing price p, traders' strategies internalize informational contents of p. This solution concept captures the idea that rational traders learn from prices by exploiting the systematic dependence of market-clearing prices on their strategies. The market-clearing price satisfies

$$\sum_{i=1}^{n+1} q_i(p; H_i) = 0$$

subject to a market-clearing rule described in the Appendix. To make explicit the dependence of the market-clearing price and an allocation on the strategies of traders, write $p = p(q)$ and $q_i = q_i(q)$, where $q = (q_1, ..., q_{n+1})$ is a vector of strategies. A rational expectations equilibrium with imperfect competition is defined as a q that satisfies

$$\forall i \in \{1, ..., n+1\}, \; E[U((v - p(q)) q_i(q) + ve_i)] \geq E[U((v - p(q')) q_i(q') + ve_i)] \quad (4)$$

for any q' differing from q only in the i-th component. We call this equilibrium *a trade equilibrium* in this paper. In Section 4, we also study *a price-taking equilibrium*, which is defined by replacing $p(q')$ with $p(q)$ in (4).

The absence of noise traders facilitates a welfare analysis based on the *ex ante gains from trade* (henceforth GFT). To define the GFT, let $E_i[\cdot]$ denote trader i's conditional expectation $E[\cdot|H_i, p]$ based on his private information and the market-clearing price.

Definition 1 (gains from trade)

The interim profit is $\Pi_i \equiv -\frac{1}{\rho}\ln\left(E_i[\exp\left(-\rho\pi_i\right)]\right)$.

The interim no-trade profit is $\Pi_i^{nt} \equiv -\frac{1}{\rho}\ln\left(E_i[\exp\left(-\rho v e_i\right)]\right)$.

The interim gains from trade is $G_i \equiv \Pi_i - \Pi_i^{nt}$.

The GFT is $G \equiv -\frac{1}{\rho}\ln\left(E[\exp\left(-\rho G_i\right)]\right)$.

Each trader's reservation value Π_i^{nt} of not trading is defined by $\exp\left(-\rho\Pi_i^{nt}\right) = E_i[\exp\left(-\rho v e_i\right)]$, i.e., the certainty equivalent value of no-trade profit ve_i conditional on private information and the market clearing-price. The monetary value of trading is captured by the interim gains from trade, $G_i \equiv \Pi_i - \Pi_i^{nt}$. The GFT, G, is defined by the ex ante certainty equivalent value of G_i. We define the optimal market size by the number of traders that maximizes G.

Definition 2 (optimal market size)

$$n^* \equiv \arg\max_{n\geq 1} G. \tag{5}$$

Remark on n^*. Our definition of the optimal market size does not take into account the costs of organizing a market. This reflects our emphasis on the role of information structure and trading rules in shaping market structures. A popular view is that (5) is infinite due to positive externalities but costs of intermediation prevent any market from becoming too large. We do not necessarily disagree with this view. However, in this paper we argue that the optimal market size can be finite *without exogenously imposed intermediation costs.* The definition (5) makes it clear that negative externalities endogenously arise in our model.

3 Information Aggregation, Multiplicity and Liquidity

We characterize a trade equilibrium of the linear form

$$q_i\left(p; H_i\right) = \beta_s s_i - \beta_e e_i - \beta_p p \tag{6}$$

by a guess-and-verify method: Given the conjecture that everyone else is using a strategy (6), derive the best response and solve a fixed point problem in $(\beta_s, \beta_e, \beta_p)$.[10] Details of the equilibrium characterization are gathered in the Appendix. Here, we focus on the information aggregation properties of the equilibrium.

Key feature of the model. Our model identifies how market size affects (i) traders' welfare, (ii) equilibrium multiplicity, (iii) market illiquidity, through information aggregation. The connection between information aggregation and market size is the key to all the results.

First, information aggregation depends on two factors: (i) traders' motives (the intensive margin) and (ii) market size (the extensive margin). To see how traders' motives affect information aggregation, suppose that traders ignore private information that is useful for other traders when submitting their orders. Then trading would be driven by other non-informational trading motives only, and the market-clearing price will not aggregate any useful information. In our model, this corresponds to $\beta_s = 0$ in (6). At the opposite extreme, if traders ignore the non-informational trading motives, the market-clearing price would accurately reflect traders' private information. In our model, this corresponds to $\beta_e = 0$. In general, information aggregation depends on how traders balance informational and non-informational trading motives in equilibrium. To see how market size affects information aggregation, notice that market-clearing prices allow traders to internalize information contained in the *average* of submitted orders. When averaged, idiosyncratic shocks in the orders tend to cancel out in a large market, while a shock common to all orders remains. Therefore, information aggregation becomes more sensitive to the aggregate shock in a larger market.

Conversely, traders' motives depend on information aggregation via rational expectations. The balance of different trading motives is endogenously determined when traders choose orders to equate the marginal value of the asset position to its marginal cost. The marginal value depends on beliefs, which are affected by the information contained in prices. In

[10] The linearity is not an exogenous constraint on equilibria because the best response is not constrained to be linear. We follow the market microstructure literature by focusing on this class of equilibria. To the best of our knowlege, the existence (or non-existence) of other types of equilibria is an open question.

our model, information aggregation changes traders' risk assessment of v_1, and hence their marginal valuations through risk aversion.

In sum, market size affects the two-way interaction between information aggregation and traders' motives by changing the balance of idiosyncratic shocks and the aggregate shock in prices. We investigate this mechanism in our model in the next two subsections.

3.1 Information aggregation

To characterize information aggregation in the model, let $Var_i[\cdot]$ denote trader i's conditional variance operator $Var[\cdot|H_i, p]$. Because the normality and symmetry of random variables make conditional variances independent of i and the realizations of (H_i, p), we define

$$\tau_1 \equiv (Var_i[v_1])^{-1}, \, \tau \equiv (Var_i[v])^{-1} = \left(\frac{1-t}{\tau_v} + \frac{t}{\tau_1}\right)^{-1}.$$

Because τ_1 is endogenously determined, so is τ as long as $t > 1$. The precision τ_1 is bounded from below by $\tau_v + \tau_\varepsilon$: a case where traders use prior beliefs and only one signal. On the other hand, τ_1 is bounded from above by a case where each trader observes $(s_1, .., s_{n+1})$. In this case, the Bayes rule implies

$$(Var_i[v_1|s_1, .., s_{n+1}])^{-1} = \tau_v + \tau_\varepsilon \frac{1+n}{1+(1-w)\,n}. \tag{7}$$

With $w = 1$, signals are conditionally independent and the upper bound (7) is $\tau_1 = \tau_v + \tau_\varepsilon (1+n)$.[11] The value of τ_1 in equilibrium depends on how much information about v_1 is revealed by the market-clearing price, but it must be between $\tau_v + \tau_\varepsilon$ and (7). In the Appendix, we show that there exists an endogenously determined $\varphi \in [0,1]$ such that

$$\tau_1 = \tau_v + \tau_\varepsilon \frac{1-\varphi+(1+n)\,w\varphi}{1-\varphi+\{1+(1-w)\,n\}\,w\varphi}, \tag{8}$$

[11] With $t = 0$, $\tau = \tau_v$ and τ_1 is irrelevant. With $w = 0$, the upper bound (7) is equivalent to the lower bound $\tau_v + \tau_\varepsilon$, because all the signals are identical. For both cases, traders' beliefs are independent of n.

$$\varphi = \left[1 + \frac{\tau_\varepsilon}{\tau_x} u \left\{ 1 + (1-u)n \right\} \left(\frac{\beta_e}{\beta_s} \right)^2 \right]^{-1}, \tag{9}$$

where $\frac{\beta_e}{\beta_s}$ in equilibrium depends on model parameters. The endogenous variable φ measures the fraction of information held by a single trader that is revealed by prices. If φ is zero, equilibrium prices do not reveal any information about $(s_1, .., s_{n+1})$ and (8) is $\tau_v + \tau_\varepsilon$. If φ is one, prices allow traders to share information $(s_1, .., s_{n+1})$ perfectly and (8) is identical to (7). With an endogenously determined $\varphi \in (0, 1)$, $n\varphi$ measures the total amount of information each trader learns from prices. In the following analysis φ plays a central role.

Strategic discount amplifies learning discount. There are two important aspects in the traders' behavior: (i) the strategic discount, (ii) the learning discount. The strategic discount arises in the first order condition that characterizes the optimal order:

$$q_i^* = \frac{E_i[v] - p - \frac{\rho}{\tau} e_i}{\lambda + \frac{\rho}{\tau}}, \text{ where } \lambda \equiv \frac{1}{n\beta_p}. \tag{10}$$

For a given belief $(E_i[v], \tau)$, traders discount the *monetary* impact of their own orders through λ.[12] Thus, the higher λ lowers the absolute size of the order.

The learning discount arises in a new signal contingent on prices:

$$h_i \equiv \frac{n\beta_p p - q_i}{n\beta_s} = v_1 + \overline{\varepsilon}_{-i} - \frac{\beta_e}{\beta_s} \left(\sqrt{1-u} x_0 + \sqrt{u} \overline{x}_{-i} \right). \tag{11}$$

To construct (11), traders discount the *informational* impact of their own orders q_i. Using (9), the informativeness of (11) is measured by $Var\left[h_i | v_1\right] = \frac{1}{n\tau_\varepsilon} \frac{1}{\varphi}$. By the Bayes rule,

$$E_i[v_1] = \frac{\tau_\varepsilon}{\tau_1} \frac{s_i - \varphi \left\{ s_i - w \frac{\beta_e}{\beta_s} \left(1 + (1-u)n\right) e_i - w \frac{\beta_p}{\beta_s} (n+1) p \right\}}{1 + (1-w)(nw-1)\varphi}. \tag{12}$$

Given $t > 0$, $\sqrt{t} E_i[v_1]$ enters in $E_i[v]$ and in (10). Thus, (12) shows that positive φ results

[12] The price impact of q_i is internalized by $p = p_i + \lambda q_i$, the relationship derived from the market-clearing condition. See Lemma A3 in the Appendix for the detail. λ measures the extent to which each trader moves prices by increasing his demand. This is used as a measure of market illiquidity in the literature.

in discounted weights on (s_i, e_i, p) due to the internalization of the informational impact on prices. Also, the learning discount of β_e is amplified by a large n through the aggregate shock $1-u$. Importantly, the learning discount and the strategic discount are distinct,[13] but affect each other. In a trade equilibrium, it can be shown that β_p is proportional to

$$\frac{n-1}{n}\left\{1-\left(1+w\frac{n+1}{n-1}\right)\varphi\right\}. \tag{13}$$

The first part $\frac{n-1}{n}$ is the result of the strategic discount, that arises in a trade equilibrium regardless of whether information is symmetric or not. The second part is the result of the learning discount, which arises only when information is asymmetric. In a price-taking equilibrium, (13) is replaced by $1-\varphi$. The learning discount is amplified in a trade equilibrium, because the price impact increases when the learning discount reduces trading volume. As a result, when $\varphi \in (0,1)$ approaches one, (13) approaches zero, making $\lambda \equiv \frac{1}{n\beta_p}$ too large to support trading, i.e., illiquid market. Therefore, we identified a potential channel through which information aggregation hurts welfare: when $\varphi \to 1$, the market shuts down due to the learning discount amplified by the strategic discount.

Endogenous trading motives. The two-way interaction between information aggregation and the balance of trading motives can be seen in (8) and (9). First, recall a conjectured equilibrium strategy: $q_i(p; H_i) = \beta_s s_i - \beta_e e_i - \beta_p p$. The market clearing price averages $\beta_s s_i - \beta_e e_i$, which is a noisy version of s_i because of $\frac{\beta_e}{\beta_s} e_i$. Hence, trader i's choice of $\frac{\beta_e}{\beta_s}$ determines how much information content of s_i is aggregated by prices. Naturally, (9) shows that a higher $\frac{\beta_e}{\beta_s}$ leads to less informative prices. Second, (8) shows the impact of information aggregation on beliefs. For a fixed n, a higher φ raises τ_1 relative to τ_v. Therefore, more information aggregation results in the less risky value of endowments. The impact of higher φ on the balance of trading motives $\frac{\beta_e}{\beta_s}$ is ambiguous. On the one hand, the

[13] For example, in Diamond and Verrecchia (1981) a price-taking behavior is assumed and only the learning discount exists. In Pagano (1989) symmetric information is assumed and only the strategic discount exists. In Kyle (1989), both types of discounts exist.

higher τ makes the unit value of endowments less risky, which reduces the hedging motive. On the other hand, higher φ increases the informational content of prices relative to that of s_i, which reduces the speculation motive. The net effect on $\frac{\beta_e}{\beta_s}$ is ambiguous.[14]

Finally, (8) and (9) also show how market size affects this two-way interaction. For a given $\varphi > 0$, (8) shows that τ_1 increases in n. This is natural because for a *fixed* per trader information aggregation φ, the total information aggregation $n\varphi$ increases in n. On the other hand, (9) shows that the presence of an aggregate shock in endowments ($u < 1$) makes the extensive margin (n) and the intensive margin ($\frac{\beta_e}{\beta_s}$) complementary in their joint impact on φ. Adding more traders whose endowments are subject to an *identical* shock increases variance of aggregate endowment, and the higher $\frac{\beta_e}{\beta_s}$ makes prices more sensitive to the increased variance. The overall effect of n on φ and $\frac{\beta_e}{\beta_s}$ depends on parameters $(t, u, w, \rho, \tau_\varepsilon, \tau_x, \tau_v)$.

The next result shows the importance of market size for the viability of trading.

Proposition 1

(a) *Trade equilibrium exists if and only if* $1 < n$ *and*

$$\frac{n+1}{n-1} < \frac{1-\varphi}{\varphi} + 1 - w. \tag{14}$$

(b) $\lim_{n\to\infty} \varphi = 0$ *if and only if* $tw < 1$ *or* $\frac{4\tau_\varepsilon \tau_x}{\rho^2} \le u < tw = 1$.

In this case, (14) *is satisfied for sufficiently large* n.

(c) $\varphi = \left(1 + \frac{\rho^2}{\tau_\varepsilon \tau_x}\right)^{-1}$ *if and only if* $twu = 1$.

In this case, (14) *is satisfied if and only if* $\frac{n+1}{n-1} < \frac{\rho^2}{\tau_\varepsilon \tau_x}$.

(d) $\lim_{n\to\infty} \varphi = 1$ *if and only if* $tw = 1$ *and* $u < \min\left\{\frac{4\tau_\varepsilon \tau_x}{\rho^2}, 1\right\}$.

In this case, (14) *is violated for sufficiently large* n.

(e) $n\varphi$ *is increasing in* n.

Proposition 1(a) shows that small φ is necessary for the existence of trade equilibrium.

[14] In Section 3.2.1, we show that parameter values (t, u) can be classified into three groups for which $\frac{\beta_e}{\beta_s}$ is increasing, decreasing or constant in n.

In the Appendix, we show that the second order condition for the traders' optimization problem is satisfied if and only if (14) holds. The intuition for why small φ is necessary follows from the interaction of the strategic discount and the learning discount. As n increases, the left-hand side of (14) decreases to one, so a trade equilibrium exists for sufficiently large n if φ converges to zero. On the other hand, if φ approaches one, (14) cannot be satisfied.

Proposition 1(b)-(d) show parameter values of (t, u, w) for which $\lim_{n\to\infty} \varphi = 0$, $\lim_{n\to\infty} \varphi \in (0, 1)$, and $\lim_{n\to\infty} \varphi = 1$ respectively. Part **(b)** provides a sufficient condition for a trade equilibrium to exist for sufficiently large n. If $t < 1$, a part of the payoff v_0 is never revealed by prices. Therefore, traders can always trade to share the endowment risk associated with v_0 as long as there are sufficiently many traders. Similarly, if $w < 1$, a part of noise ϵ_0 in signal s_i remains in prices no matter how large the market becomes, which prevents v_1 from being fully revealed by prices. If $t = w = 1$, there is neither residual risk v_0 in the payoff nor residual noise ϵ_0 in signals. Thus, an additional trader must bring enough noise to the market to make sure $\lim_{n\to\infty} \varphi = 0$. The condition $\frac{4\tau_\varepsilon \tau_x}{\rho^2} \leq u < 1$ in part **(b)** consists of two parts: $\frac{\rho^2}{\tau_\varepsilon \tau_x} > 4$ and $u \in \left[\frac{4\tau_\varepsilon \tau_x}{\rho^2}, 1\right)$. The first part states that the fundamental hedging needs (measured by $\frac{\rho^2}{\tau_x}$) is sufficiently strong relative to the speculation needs (measured by τ_ε). The second part states that the aggregate shock to endowments exists ($u < 1$), but not too large. These conditions imply that the relative hedging motive does not disappear, i.e., $\lim_{n\to\infty} \frac{\beta_e}{\beta_s} > 0$. This makes prices sufficiently noisy in a large market.[15]

Part **(c)** is a knife-edge case where both $\frac{\beta_e}{\beta_s}$ and φ are independent of n. In this case, with the additional condition $\frac{\rho^2}{\tau_\varepsilon \tau_x} > 1$, a trade equilibrium exists for sufficiently large n. Part **(d)** shows that if $\frac{\rho^2}{\tau_\varepsilon \tau_x}$ is sufficiently small in the presence of the aggregate shock to endowments, (14) is violated for large n. For this case, we need to characterize what happens in equilibrium as n approaches a finite value for which (14) holds with equality. We discuss this case in Section 3.2.2. Finally, part **(e)** shows that $n\varphi$, the total amount of information revealed by prices, increases in n even if φ converges to zero.

[15] With $u < 1$ and $\frac{\beta_e}{\beta_s}$ bounded away from zero, (9) implies $\lim_{n\to\infty} \varphi = 0$.

3.2 Multiplicity and extreme illiquidity

First, we identify conditions under which the multiplicity arises and show that multiple equilibria are Pareto-ranked. Second, we study cases where the market liquidity dries up.

3.2.1 Multiplicity

The way $\frac{\beta_e}{\beta_s}$ is determined in equilibrium as a self-fulfilling conjecture opens a door to multiple equilibria. For example, a large (small) conjectured value of $\frac{\beta_e}{\beta_s}$ and rationally anticipated small (large) φ from (9) may justify large (small) $\frac{\beta_e}{\beta_s}$ as a best response. To identify conditions under which this occurs, we classify a parameter space of (t, u) for a given $w \in (0, 1]$ using

$$X(t, u, w) \equiv \frac{1 - w + c_s (w - t)}{(1 - u)(1 - tc_s)}, \text{ where } c_s \equiv \frac{\tau_\varepsilon}{\tau_v + \tau_\varepsilon}.$$

For a fixed $w \in (0, 1]$, define for each $m \in [0, \infty]$

$$\begin{aligned} \chi_w (m) &\equiv \{(t, u) | X(t, u, w) = m\} \text{ for } m < \infty, \\ \chi_w (\infty) &\equiv \{(t, u) | t \in (0, 1], u = 1\}. \end{aligned}$$

Note that $\chi_w (m)$ defines a positive relationship between t and u for any $m < \infty$. **Figure 3** illustrates this classification.

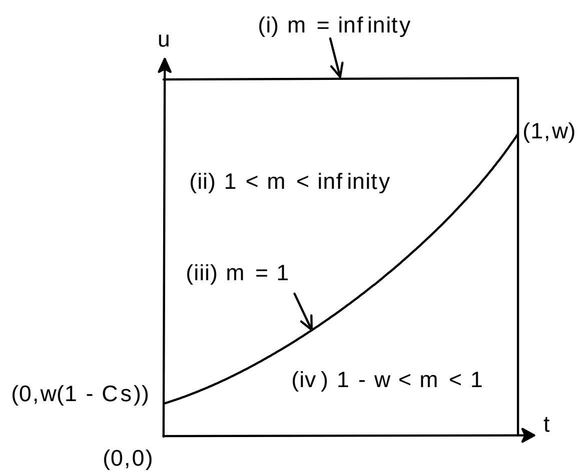


Figure 3. Classification of (t, u) for a fixed $w \in (0, 1)$.

We argued that the effect of n on $\frac{\beta_e}{\beta_s}$ was ambiguous. In fact, we can show that $\frac{\beta_e}{\beta_s}$ increases (decreases, constant) in n if and only if (t, u) is above (below, on) the line defined by $\chi_w(1)$.[16] To interpret this result, consider low t, high u, low w so that we are in the parameter region above the line $\chi_w(1)$. These parameters imply that the learning incentive is low and risk sharing gains are high. With this (t, u, w), the hedging balance $\frac{\beta_e}{\beta_s}$ increases in n. If we either raise t or lower u or raise w, the economy will move toward the region below the line $\chi_w(1)$. Thus, the higher learning incentive and/or less risk sharing gains weaken the positive effect of n on $\frac{\beta_e}{\beta_s}$. We show that multiple equilibria arise only in the parameter region *well below* $\chi_w(1)$. Intuitively, in the north-west region of **Figure 3**, risk sharing is a sufficiently dominant trading motive that self-fulfilling motives can not play a decisive role.

Proposition 2

(a) *Given that* (14) *is satisfied, multiple equilibria exist if only if*

$$\text{(i) } w > \frac{8}{9},\ (t,u) \in \bigcup_{m\in\left[0,\frac{1}{9}\right)} \chi_w(m)\,,\ \text{(ii) } \frac{8}{1-9m} < n(1-u),\ \textit{and}\ \text{(iii) } \frac{\rho^2}{\tau_\varepsilon \tau_x} \in \left[\alpha^-_{n,m}, \alpha^+_{n,m}\right],$$

where $\alpha^{\pm}_{n,m}$ *are defined in the proof.*

(b) *An equilibrium with a larger* $\frac{\beta_e}{\beta_s}$ *is Pareto-superior.*

The multiplicity arises with (i) the high level of learning incentive (large t and w), (ii) a large aggregate shock to prices (large $n(1-u)$), and (iii) balanced *fundamental* motives for trading (not too extreme $\frac{\rho^2}{\tau_\varepsilon \tau_x}$). To understand the economic meaning of (i)-(iii) in part **(a)**, recall that the ratio $\frac{\beta_e}{\beta_s}$ captures the *conjectured* balance of two trading motives, and start with a conjecture of a typical trader that other traders choose large $\frac{\beta_e}{\beta_s}$. From

$$\varphi = \left\{1 + \frac{\tau_\varepsilon}{\tau_x} u\,(1 + (1-u)\,n) \left(\frac{\beta_e}{\beta_s}\right)^2\right\}^{-1},$$

[16] See Lemma A6 in the Appendix.

this trader expects that prices become less informative. This creates more hedging needs for this trader, justifying a larger β_e as a best response. However, β_s would also be higher because this trader would put a larger weight on his signal given less informative prices. If the net effect on $\frac{\beta_e}{\beta_s}$ is small, the initial conjecture cannot be sustained in a symmetric equilibrium. The aggregate shock ($u < 1$) weakens this offsetting force by making φ more sensitive to $\frac{\beta_e}{\beta_s}$ through $n(1-u)$ that appears in the expression of φ. This channel is created by the presence of the aggregate shock x_0 to endowments. Because x_0 does not cancel out when averaged, the negative impact of the individual $\frac{\beta_e}{\beta_s}$ on φ is multiplied by n. For large n and a large conjectured value of $\frac{\beta_e}{\beta_s}$, prices are expected to be so noisy that, as a best response, β_e can increase much more than β_s does. This rationalizes large $\frac{\beta_e}{\beta_s}$ as the best response. Thus, condition (ii) highlights the role of market size n: with an aggregate shock, the conjectured $\frac{\beta_e}{\beta_s}$ has a bigger effect on the informativeness of prices in a larger market.

To support both small and large conjectured $\frac{\beta_e}{\beta_s}$ in equilibrium, the *fundamental* trading motives cannot be extreme. For example, sufficiently small (large) $\frac{\rho^2}{\tau_\varepsilon \tau_x}$ would allow only very small (large) $\frac{\beta_e}{\beta_s}$ (consider traders being close to risk-neutral or highly risk averse). When one motive does not dominate the other, fundamental motives $\frac{\rho^2}{\tau_\varepsilon \tau_x}$ do not imply extreme values of $\frac{\beta_e}{\beta_s}$, and a conjectured $\frac{\beta_e}{\beta_s}$ may justify itself in equilibrium. Thus, $\frac{\rho^2}{\tau_\varepsilon \tau_x}$ should not be too extreme (condition (iii)) to support multiple values of $\frac{\beta_e}{\beta_s}$. Similarly, too small t or w would make the speculation motive trivial, making it difficult for the low conjecture of $\frac{\beta_e}{\beta_s}$ to be self-fulfilling. This explains condition (i).

Given our interest in the optimal market size, it is natural to select an equilibrium that is not Pareto-dominated. Part (**b**) shows that equilibria are Pareto-ranked by the size of $\frac{\beta_e}{\beta_s}$, and hence by φ. This ranking is intuitive. The higher $\frac{\beta_e}{\beta_s}$ (and hence lower φ) is consistent with trading generated by risk sharing, not by speculation. This results in a higher ex-ante welfare. Accordingly, whenever multiple equilibria exist, we select the one with the smallest φ: the equilibrium with the most risk sharing and the least information sharing.

3.2.2 Extreme illiquidity

For the case covered by **Proposition 1(d)**, a trade equilibrium fails to exist for a large n. As we discussed in Section 3.1, there is a particularly strong negative externality for this case. It is not only that a particular type of equilibria disappear, but also that volume converges to zero *as n increases to a finite value*. We denote equilibrium volume by $q_i(p^*; H_i)$.

Proposition 3

(a) *Consider $t = w = 1$ and $u < \min\left\{\frac{4\tau_\varepsilon\tau_x}{\rho^2}, 1\right\}$.*

If there is $\widetilde{n} \in (1, \infty)$ such that $\frac{\widetilde{n}+1}{\widetilde{n}-1} = \frac{1-\varphi}{\varphi}$, then $\lim_{n \nearrow \widetilde{n}} \lambda = \infty$ and $\lim_{n \nearrow \widetilde{n}} q_i(p^*; H_i) = 0$.

(b) *For any fixed n for which a trade equilibrium exists, there exists $\underline{u}(n) \in (0, 1)$ s.t.*

(i) (14) *is violated for all $u \leq \underline{u}(n)$,*

(ii) $\lim_{u \searrow \underline{u}(n)} q_i(p^*; H_i) > 0$ *if $w < 1$, while* $\lim_{u \searrow \underline{u}(n)} q_i(p^*; H_i) = 0$ *if $w = 1$.*

For the parameter values stated in part **(a)**, a trade equilibrium approaches a no-trade outcome as n increases to some finite value $\widetilde{n}$. Therefore, the GFT must approach zero. This happens when more traders result in more speculation, which increases φ until (14) is violated at $n = \widetilde{n}$. As n increases to $\widetilde{n}$, $\lambda \equiv \frac{1}{n\beta_p}$ increases to infinity, i.e., the market becomes infinitely illiquid. Part **(b)** shows another comparative statics which also exhibits market illiquidity. For fixed market size, equilibrium disappears as u drops to a *strictly positive* lower bound. This means that if traders believe there is a sufficiently large aggregate shock to endowments, the market shut down even though there are potential gains from trade. Conditions on (t, u, w) in **Proposition 3** may appear rather special. However, in Section 4.3.2 we show that the extreme illiquidity arises more generally in a generalized model.[17]

[17] For example, assume $t = u = w = 1$ but model an individual endowment variance $\widetilde{\tau}_x$ by $\widetilde{\tau}_x = n\tau_x$ for a fixed τ_x. Define $\widetilde{n}^{\pm}$ by the two solutions to $\widetilde{n} = \frac{\frac{\rho^2}{\tau_\varepsilon\tau_x}+\widetilde{n}}{\frac{\rho^2}{\tau_\varepsilon\tau_x}-\widetilde{n}}$ given $3 + 2\sqrt{2} < \frac{\rho^2}{\tau_\varepsilon\tau_x}$. It follows that $1 < \widetilde{n}^- < \widetilde{n}^+ < \frac{\rho^2}{\tau_\varepsilon\tau_x}$ and (14) is satisfied for all $n \in (\widetilde{n}^-, \widetilde{n}^+)$. As $n \to \widetilde{n}^+$, the extreme illiquidity occurs.

4 Ex Ante Gains From Trade and Market Size

We show first that $n^* = \infty$ in a symmetric information benchmark (**Lemma 1**). Second, **Proposition 4** shows that the optimal market size is finite whenever equilibria exist with $t > 0$ and $w > 0$. Third, we decompose an expression for the GFT (**Lemma 2**) to show four distinct economic forces at work. Finally, we study the robustness of the main result.

4.1 Optimal market size

First, we study a symmetric information benchmark.

Lemma 1 *Assume $t = 0$ or $w = 0$.*

(a) *A trade equilibrium exists for $n > 1$ and volume is $q_i(p^*; H_i) = -\frac{n-1}{n}(e_i - \overline{e})$.*

(b) *The GFT shown below increases in n and decreases in τ_v and τ_ε.*

$$G^s = \frac{1}{2\rho} \log \left(1 + \frac{\rho^2}{\tau_v \tau_x} u \frac{n-1}{n} \frac{\tau_v + (1-t)\,\tau_\varepsilon}{\tau_v + \tau_\varepsilon} \right). \tag{15}$$

Setting $t = 0$ in (15) yields the result for $t = 0$ with any $w \in [0, 1]$. When $t = 0$ or $w = 0$, the information about the asset is symmetric and the equilibrium price aggregates no information. The two-way interaction is lost and *beliefs are independent of* n. As long as there is a diversifiable risk ($u > 0$), the GFT increase in n and in the limit $n \to \infty$, everyone holds the average position.[18] Importantly, G decreases in τ_v and also in τ_ε if $t > 0$. This shows that *exogenous* changes in beliefs such that traders know more about v decrease gains from trade. This is the Hirshleifer effect, but market size plays no role here. With asymmetric information, however, beliefs are endogenously determined and depend on market size n. Since endogenous beliefs τ increase in n, there are now two opposing forces: (i) risk sharing gains increase in n for fixed beliefs, (ii) equilibrium beliefs change in a way that discourages the hedging motive. The next proposition is the main result of this paper.

[18] $q_i(p^*; H_i) + e_i = \frac{1}{n} e_i + \frac{n-1}{n} \overline{e}$ and the strategic discount $\frac{n-1}{n}$ disappears as $n \to \infty$.

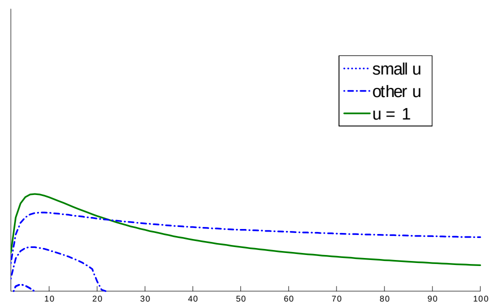
Proposition 4 *Assume $t > 0$ and $w > 0$. In a trade equilibrium,*

(a) *The optimal market size n^* is finite and* $\lim\limits_{t \to 0 \text{ or } w \to 0} n^* = \infty$.

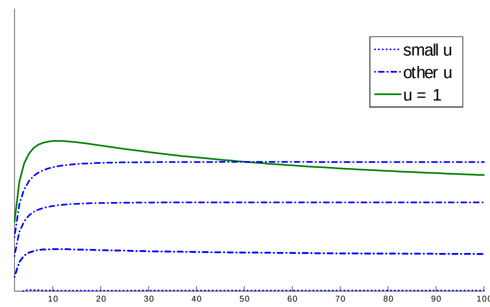
(b) *If $t = w = u = 1$, then* $\lim\limits_{n \to \infty} G = 0$.

(c) *If $t = w = 1$ and $u < \min\left\{\frac{4\tau_\varepsilon \tau_x}{\rho^2}, 1\right\}$, then there is $\widetilde{n} < \infty$ s.t.* $\lim\limits_{n \to \widetilde{n}} G = 0$.

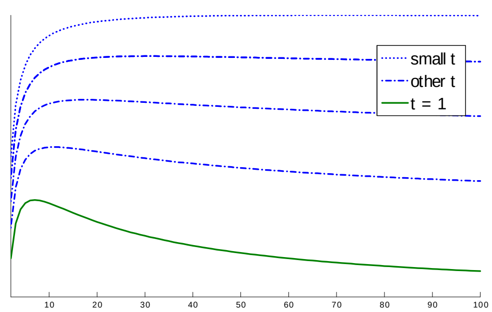
Figure 4 shows G as a function of n for $w = 1$ and different (t, u).



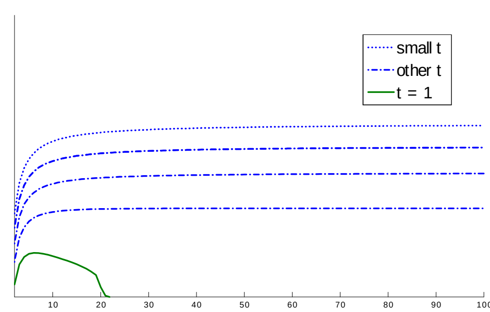
(a) $t = 1$.



(b) $t < 1$.



(c) $u = 1$.



(d) $u < 1$.

Figure 4. Gains from trade for $w = 1$ with different (t, u). The ex ante gains from trade $G(n)$ is measured on the vertical axis as a function of the number of traders n. The case with $t = 1$ has no residual risk v_0. The case with $u = 1$ has no aggregate shock x_0.

Part **(a)** shows that the presence of asymmetric information, no matter how small, makes the optimal market size finite. While this may sound counter-intuitive, notice that as t or w

approaches zero, n^* approaches ∞ as in the symmetric information benchmark case. Hence, for sufficiently small t or w, the optimal market size is so large that positive externalities should arise for a *practical* range of n. Part **(b)** and **(c)** show that if $t = w = 1$ and u is either 1 or take values that cause the extreme illiquidity, then the GFT converge to zero and hence exhibit a stronger negative externality.[19] Thus, markets are more likely to be fragmented in the presence of the high learning incentive captured by $t = w = 1$. The residual uncertainty v_0 $(t < 1)$ or a common shock in signal noise ϵ_0 $(w < 1)$ works against market fragmentation, because they weaken the learning incentive relative to the hedging motive.

4.2 Hirshleifer discount and aggregate shock discount

We use the following result to study how market size affects the GFT.

Lemma 2 *Assume $t > 0$ and $w > 0$.*

$$G = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x}u\frac{n-1}{n}\left(1 - \frac{n+1}{n-1}\frac{w\varphi}{1-\varphi}\right)\frac{\tau_v}{\tau}\frac{\frac{1-\varphi}{\varphi} + w}{\frac{1-\varphi}{\varphi} + w\left\{1 + (1-u)\,n\right\}}\right). \tag{16}$$

Table below lists the expression inside the log function in (16) and the symmetric information case (15). We also compare the same expression under the price-taking assumption.[20]

Symmetric information ($t = 0$ or $w = 0$)	$1 + \frac{\rho^2}{\tau_v\tau_x}u \times \frac{n-1}{n} \times 1 \times \frac{\tau_v + (1-t)\tau_\varepsilon}{\tau_v + \tau_\varepsilon} \times 1$
Asymmetric information ($tw > 0$)	$1 + \frac{\rho^2}{\tau_v\tau_x}u \times \frac{n-1}{n} \times \left(1 - \frac{n+1}{n-1}\frac{w\varphi}{1-\varphi}\right) \times \frac{\tau_v}{\tau} \times \frac{\frac{1-\varphi}{\varphi}+w}{\frac{1-\varphi}{\varphi}+w+w(1-u)n}$
Price-taking ($tw > 0$)	$1 + \frac{\rho^2}{\tau_v\tau_x}u \times \frac{n}{n+1} \times \left(1 - \frac{w\varphi}{1-\varphi(1-w)}\right) \times \frac{\tau_v}{\tau} \times \frac{\frac{1-\varphi}{\varphi}+w}{\frac{1-\varphi}{\varphi}+w+w(1-u)n}$

We label terms separated by multiplication signs as follows:

(i) fundamental gains from trade, (ii) strategic discount, (iii) learning discount,

(iv) Hirshleifer discount, (v) aggregate shock discount.

Market size affects the GFT through the last four channels. First, in a trade equilibrium, there is a strategic discount relative to a price-taking equilibrium ($\frac{n-1}{n} < \frac{n}{n+1}$) due to the

[19] With $t = u = w = 1$, n^* can be explicitly solved. See **Lemma A2** in the Appendix.

[20] See **Lemma 3** below for details.

internalization of price impacts. However, the expression $\frac{n-1}{n}$ itself is increasing in n. Hence, this is a source of positive externalities. Second, with asymmetric information, there is a learning discount. As we discussed in Section 3.2, this is amplified by the strategic discount ($\frac{n+1}{n-1}\frac{w\varphi}{1-\varphi} > \frac{w\varphi}{1-\varphi(1-w)}$). The learning discount can increase or decrease in n depending on whether φ increases in n or not. If φ does not increase in n, then the learning discount is mitigated by a large n. If $\lim_{n\to\infty}\varphi = 1$, the learning discount destroys the GFT for a finite n. Thus, the extreme illiquidity is a distinct phenomenon from the Hirshleifer effect: it is not driven by the loss of risk sharing opportunity, but by the learning discount amplified by the strategic discount. Third, information aggregation changes beliefs relative to prior beliefs, and this shows up as the Hirshleifer discount. In the symmetric information case, the Hirshleifer discount exists if $t > 0$, but *it does not depend on* n. With asymmetric information, the Hirshleifer discount $\frac{\tau_v}{\tau}$ depends on market size n. Finally, the aggregate shock discount arises only with asymmetric information and the aggregate shock to endowments (i.e. $w(1-u) > 0$). This is because the learning discount of the hedging motive β_e is amplified by the aggregate shock $(1-u)n$, which can be seen in (12). In the proof of **Proposition 4**, we show that the combined negative effect of the Hirshleifer discount and the aggregate shock discount on welfare become dominant as market size increases.

In sum, with symmetric information, only the strategic discount exists. Because larger n weakens the strategic discount, the optimal market size is infinite. Asymmetric information creates three other channels. Market size negatively affects the GFT through the Hirshleifer discount and the aggregate shock discount, while the learning discount can work in either way depending on parameter values. However, the combined negative effect is strong enough to make the optimal market size always finite. For example, setting $u = 1$ shuts down the aggregate shock channel and makes the learning discount a source of positive externalities. Nevertheless, the Hirshleifer discount alone leads to the finite optimal market size.

4.3 Robustness of the main result

We study two extensions[21] which improve our understanding of the baseline case: (i) price-taking equilibrium, (ii) the individual parameters $(\tau_v, \tau_\varepsilon, \tau_x)$ as functions of n.

4.3.1 Price-taking equilibrium

Lemma 3

(a) *For all $n \geq 1$, a price-taking equilibrium exists and*

$$G < G^{pt} = \frac{1}{2\rho} \log \left(1 + \frac{\rho^2}{\tau_v \tau_x} u \frac{n}{n+1} \left(1 - \frac{w\varphi}{1 - \varphi (1-w)} \right) \frac{\tau_v}{\tau} \frac{\frac{1-\varphi}{\varphi} + w}{\frac{1-\varphi}{\varphi} + w + w(1-u)n} \right).$$

(b) *The optimal market size n^* is finite.*

(c) *n^* is larger in a price-taking equilibrium than in a trade equilibrium if and only if extreme illiquidity occurs in a trade equilibrium.*

First, a price-taking equilibrium always exists. This is because there is no strategic discount that amplifies the learning discount. While the learning discount $1 - \frac{w\varphi}{1-\varphi(1-w)}$ exists, it is strictly positive for all $\varphi \in (0,1)$. It is important to note that in a linear equilibrium, the price impact λ in (10) affects three coefficients $(\beta_s, \beta_e, \beta_p)$ *proportionally.* As a result, the condition that determines $\frac{\beta_e}{\beta_s}$ is independent of whether price impacts are internalized or not. Because the same amount of information is revealed under both assumptions, the Hirshleifer discount and the aggregate shock discount are the same in both equilibria. However, less risk is shared in trade equilibrium due to the strategic discount and the amplified learning discount. This results in the lower welfare in the trade equilibrium.

[21] The Appendix contains two more extensions (transaction fees and sequentially opening markets).

Intuition for part **(c)** follows from the comparison of the two discounts:

$$\text{Trade equilibrium:} \qquad \frac{n-1}{n} \times \left(1 - \frac{n+1}{n-1}\frac{w\varphi}{1-\varphi}\right).$$
$$\text{Price-taking equilibrium:} \qquad \frac{n}{n+1} \times \left(1 - \frac{w\varphi}{1-\varphi\,(1-w)}\right).$$

If φ does not increase in n, a larger n increases the GFT through both channels in both equilibria. The two expressions above show that the welfare gap between two equilibria decreases in n. Because the trade equilibrium needs *more* traders to catch up with the price-taking GFT, the optimal market size must be larger in the trade equilibrium. If φ increases in n, the welfare gap between two equilibria *increases* in n, because the GFT go to zero for a finite n in a trade equilibrium, while they go to zero only in the limit $n \to \infty$ in a price-taking equilibrium. Because the trade equilibrium needs to reduce traders to catch up with the price-taking GFT, the optimal market size must be smaller in the trade equilibrium.

4.3.2 Individual parameters dependent on n

We replace constant $(\tau_v, \tau_\varepsilon, \tau_x)$ with $(\widetilde{\tau}_v, \widetilde{\tau}_\varepsilon, \widetilde{\tau}_x)$ that depend on n as follows:

$$\widetilde{\tau}_v = \tau_v n^{\delta_v},\ \widetilde{\tau}_\varepsilon = \tau_\varepsilon n^{\delta_\varepsilon},\ \widetilde{\tau}_x = \tau_x n^{\delta_x}. \tag{17}$$

Our baseline model has $(\delta_v, \delta_\varepsilon, \delta_x) = (0, 0, 0)$. These parameters describe how the *exogenous* information structure changes as n increases. For example, if $\delta_x < 0$ (> 0), then traders have more (less) volatile endowment shocks x_0 and x_1 as n increases.[22] To restrict our attention to $(\delta_\varepsilon, \delta_x)$, we impose the following restriction:

$$\delta_v = \delta_\varepsilon. \tag{18}$$

[22] If $\delta_v > 0$ (< 0), then traders have more (less) precise prior knowledge about v_0 and v_1 as n increases. If $\delta_\varepsilon > 0$ (< 0), then traders have more (less) precise signals about v_1 as n increases.

This means that precision of prior knowledge about v and that of private signals move in the same direction at the same rate as n increases.[23] We show in the Appendix that

$$G=\frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}u\frac{n-1}{n}\left(1-\frac{n+1}{n-1}\frac{w\varphi}{1-\varphi}\right)\frac{\tau_v}{\tau}\frac{\frac{1-\varphi}{\varphi}+w}{\frac{1-\varphi}{\varphi}+w\left\{1+\left(1-u\right)n\right\}}n^{-(\delta_\varepsilon+\delta_x)}\right),$$

$$\varphi=\left[1+\frac{\tau_\varepsilon}{\tau_x}u\left\{1+(1-u)n\right\}\left(\frac{\beta_e}{\beta_s}\right)^2 n^{-(\delta_\varepsilon+\delta_x)}\right]^{-1}.$$

Because only the sum $\delta_\varepsilon+\delta_x$ matters, define $\delta\equiv\delta_\varepsilon+\delta_x$.

Lemma 4 *Assume* (17) *and* (18).

(a) $\lim_{n\to\infty}\varphi=0$, $\lim_{n\to\infty}n\varphi=\infty$, $\lim_{n\to\infty}G=0$ *if and only if either:*

(i) $0<tw<u=1$ *and* $\delta\in(0,2)$, *or*

(ii) $0<tw<1$, $u<1$ *and* $\delta\in(0,1)$, *or*

(iii) $twu=1$ *and* $\delta\in\left(-\frac{1}{2},0\right)$, or

(iv) $u<tw=1$, $\delta\in(-1,0)$ *and* $\delta_k=\delta$.

For each case above,

(b) *If* δ *is above the upper bound,* $\lim_{n\to\infty}\varphi=1$ *and the extreme illiquidity arises.*

(c) *If* δ *is below the lower bound, then* $\lim_{n\to\infty}G=\infty$.

The environment with $\delta=0$ is a special (and boundary) case of the more general environment studied here. Part **(a)** shows that, as long as information is asymmetric (i.e. $tw>0$), for each parameter values of (t,u,w) there exists *a non-empty set* of δ for which $\lim_{n\to\infty}G=0$ and $\lim_{n\to\infty}\varphi=0$. Part **(b)** shows that the extreme illiquidity *always* arises for sufficiently high value of δ. With $\delta=0$, we showed that gains from trade decrease, but not necessarily to zero. Thus, **Lemma 4** strengthens and generalizes the main finding of the paper.[24]

[23] This can occur when both public signal and private signals are generated by the same information source, whose quality monotonically depends on the number of customers n.

[24] In Lemma4 (a)(iv), δ_k is an order of $k\equiv\frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}\sim n^{\delta_k}$. In this case, the multiplicity in the *order* of $\frac{\beta_e}{\beta_s}$ can occur. Recall that the multiplicity in the *level* of $\frac{\beta_e}{\beta_s}$ occured with $u<tw=1$ in the baseline case. See Lemma A14 in the Appendix for the detail.

Part **(c)** shows that for sufficiently small δ, the GFT increase without a bound. This is because with negative δ an increase in n exogenously creates more *per trader* risk. However, we view $\delta \geq 0$ as more reasonable than $\delta < 0$ for the following reason. In the model, and also in reality to some extent, the GFT are larger for traders facing higher idiosyncratic risk. Hence, if traders are heterogenous in $(\tau_\varepsilon, \tau_x)$, those with low values of $(\tau_\varepsilon, \tau_x)$ have the higher incentive to join the market than those with high values of $(\tau_\varepsilon, \tau_x)$. When the participation cost goes down, the entry of new traders should *raise* the average value of $(\tau_\varepsilon, \tau_x)$ in the market. Thus, once we consider the ex ante incentive of traders to join the market, it is likely that the average precision $(\tau_\varepsilon, \tau_x)$ in the market *increases* in n. This argument supports $\delta \geq 0$ and our focus on the case with negative externalities.

5 Application: Endogenous Market Structure

We present a model of endogenous market structure based on the rational expectations model of trading studied in previous sections. We study a two-stage game played by traders and intermediaries at the ex ante stage. As we believe that market formation would take much more time compared to trading, we assume that traders and intermediaries can commit to a market structure determined before the realization of private information to avoid the cost of redesigning it in every information state before trading.

Let $N = \{1, .., \overline{n}\}$ be the set of potential traders and let $J = \{0, 1, .., \overline{j}\}$ denote the set of markets. A market structure for a given N is a partition of N such that $N = \underset{j \in J}{\cup} N_j$ and $N_j \cap N_k = \emptyset$ for any $j \neq k$. We use N_0 for the set of traders who do not participate in any markets and N_j with $j > 0$ for a set of traders in a market j. A lowercase letter n_j denotes the number of traders in N_j. The GFT for each trader in N_j are given by $G(n_j)$, while $C(n_j)$ is the total operational cost faced by an intermediary who runs a market j with n_j traders. To focus on the implication of $G(\cdot)$, we assume $C(n) = cn$. This assures that market fragmentation is not due to ad hoc assumptions on the cost function.[25] We assume that traders decide whether

[25] Other things equal, a rapidly increasing cost function would trivially make each market small. Qualitative

or not to participate in any market based on the GFT G and entry fees set by intermediaries. The market formation game proceeds in two steps. First, intermediaries simultaneously set the fees $\{\phi_j\}_{j \in J}$, where $\phi_0 \equiv 0$.[26] Second, traders simultaneously decide which market to participate in, taking the fees as given. After we analyze traders' participation decision for given fees, we let intermediaries compete to determine fees.

Because traders' participation decisions are complementary in their nature, there are many Nash equilibria. As an equilibrium selection criterion, we use a cooperative notion of stability in that no subset of traders in one market has an incentive to deviate.

Definition 3 (stable participation)

A Nash equilibrium is not stable if the resulting market structure satisfies either

(i) $\exists j, k \in J \backslash \{0\}$ *and* $n \in \{1, .., n_j\}$ *s.t.* $G(n_j) - \phi_j < G(n + n_k) - \phi_k$, *or*

(ii) $\exists j \in J \backslash \{0\}$ *and* $n \in \{1, .., n_0\}$ *s.t.* $0 < G(n + n_j) - \phi_j$.

Definition 3 stipulates that any subsets of traders in each market are free to move if there is a better alternative. By focusing on stable equilibria, we exclude a trivial equilibrium where no trader participates in any markets, as well as a mixed-strategy equilibrium where traders randomly choose a market. Let $0 \leq \underline{\phi} \equiv \lim_{n \to \infty} G(n) < \overline{\phi} \equiv G(n^*)$, where $n^* \equiv \arg\max_n G(n)$. By setting $\phi_j > \overline{\phi}$, intermediary j does not attract any traders, while setting $\phi_j < c$ would only cause a loss. Hence, we can focus on $\phi_j \in [c, \overline{\phi}]$. **Lemma 5** is a necessary condition for the existence of the stable equilibrium.

Lemma 5 *In a stable equilibrium given* $\{\phi_j\}_{j \in J}$, *for all* $j \in J \backslash \{0\}$,

$$G(n_j) - \phi_j \geq \max\{G(n_k + 1) - \phi_k, 0\} \text{ for all } k \in J \backslash \{0, j\}. \tag{19}$$

results hold for more general cost functions, as long as $G(n) - \frac{C(n)}{n}$ has a single peak.

[26] There is no room for horizontal differentiation because all traders are ex ante identical. This contrasts our analysis to that of Economides and Siow (1988), which is a model of horizontal differentiation.

Additionally, either one of the following two conditions holds:

$$\text{(i) } \textit{For all } j \in J \backslash \{0\},\ G(n_j) \geq G(n_j+1). \tag{20}$$

$$\text{(ii) } n_0 = 0 \textit{ and there exists one } j \in J\backslash\{0\} \textit{ s.t. } G(n_j) < G(n_j+1) \textit{ and}$$

$$G(n_k+1) - \phi_k \leq G(n_j) - \phi_j < \overline{\phi} - \phi_j \leq G(n_k) - \phi_k \text{ for all } k \in J\backslash\{0, j\}. \tag{21}$$

Lemma 5 shows that there are two types of stable equilibria. The first type satisfies (20), i.e., all markets operate on the decreasing part of G. The second type satisfies (21), i.e., all but one market operates on the decreasing part of G. The second type of equilibrium is asymmetric in that the smallest market offers the smallest net benefits to traders compared to all other markets.[27] While the possibility of the asymmetric equilibrium is interesting, we focus on a symmetric equilibrium characterized by (19) and (20) for two reasons. First, given symmetric intermediaries, it seems natural to expect they operate in a symmetric way, if there is such an equilibrium. Second, the asymmetric equilibrium requires that G should be quickly decreasing so that it is not worth switching to larger markets. The numerical evaluation of G shows that it is unlikely to be satisfied (see **Figure 4**).[28]

We write trader i's problem as:

$$\max_{r_i = \{r_{i,j}\}_{j\in J}} \sum_{j\in J\backslash\{0\}} r_{i,j} \left\{ G\left(\sum_{i\in N} r_{i,j}\right) - \phi_j \right\} \tag{22}$$

$$\text{s.t. } r_{i,j} \in \{0,1\} \text{ for all } j \in J \text{ and } \sum_{j\in J} r_{i,j} = 1. \tag{23}$$

A participation equilibrium for given $\{\phi_j\}_{j\in J}$ is $\{r_i^*\}_{i\in N}$ such that for all $i \in N$, r_i^* solves (22) subject to (23) given r_{-i}^*, and it is stable. Note that (23) constrains each trader to

[27] Traders in the smallest market do not want to move to other markets because if they do, the larger markets will be too "congested".

[28] A possible exception is the case of the extreme illiquidity studied in Section 3.

participate in at most one market. The equilibrium $\{r_i^*\}_{i\in N}$ determines a market size

$$n_j(\phi_j, \phi_{-j}) = \sum_{i\in N} r_{i,j}^* \tag{24}$$

for each j. From intermediary j's perspective, (24) is a demand function. Given (24), intermediaries compete by setting fees $\left\{\phi_j\right\}_{j\in J}$. Intermediary j's problem is

$$\max_{\phi_j} \ (\phi_j - c)n_j, \text{ s.t. } n_j = n_j(\phi_j, \phi_{-j}). \tag{25}$$

A fee-setting equilibrium is $\{\phi_j^*\}_{j\in J}$ such that for all $j \in J\backslash\{0\}$, ϕ_j^* solves (25) given ϕ_{-j}^*.

For the rest of our analysis, we use an approximation to avoid issues associated with the integer restriction.[29] The conditions (19) and (20) are approximated by

$$G\left(n_j\right) - \phi_j = G\left(n_k\right) - \phi_k \geq 0 \text{ for all } j, k \in J\backslash\{0\}, \tag{26}$$

$$n^* \leq \ n_j \text{ for all } j \in J\backslash\{0\}. \tag{27}$$

In a participation equilibrium characterized by (26) and (27), all markets operate at the decreasing part of G and provide the same net benefit to traders. The indifference condition (26) implicitly defines demand functions for intermediaries.

Monopoly. When there is one intermediary ($\bar{j} = 1$), the nature of the demand function depends on the relative size of $\overline{n}$ and n^*. If $\overline{n} \leq n^*$, the market size is always $\overline{n}$ as long as $G(\overline{n}) > c$. The demand is completely elastic: setting $\phi \leq G(\overline{n})$ attracts all traders $\overline{n}$ while setting $\phi > G(\overline{n})$ attracts no trader. If $n^* < \overline{n}$, the demand function has an inelastic part: setting $\phi \in (G(\overline{n}), \overline{\phi}]$ would attract only some traders $n(\phi) = G^{-1}(\phi) \in [n^*, \overline{n})$. In this part of the demand function, raising a fee only slightly reduces participating traders.

Lemma 6 *If $\overline{n} > n^*$, then the monopoly market size n^m is larger than n^*.*

[29] This approximation was also used by Economides and Siow (1988).

To see why the monopoly market is larger than n^*, define a marginal surplus function

$$H(n) \equiv \frac{\partial}{\partial n} \{(G(n) - c)\, n\} = G(n) - c + G'(n)n.$$

This is an increase in social surplus by adding a marginal trader in a market of size n. The term $G(n) - c$ captures the net surplus for the marginal trader, while $G'(n)n$ represents the externality imposed on the others. At n^*, there is no externality ($G'(n^*) = 0$), but the profit is still increasing in the market size because $H(n^*) = G(n^*) - c > 0$. Therefore, the intermediary would like to have more than n^* traders.

Competition. We characterize a fee-setting equilibrium assuming that $\bar{j} \geq 2$ competing intermediaries rationally anticipate that a stable participation equilibrium will be played once they set fees. The indifference condition (26) implies that each intermediary offers the same level of net benefit to traders in a participation equilibrium. We assume that intermediaries treat this value as a parameter denoted by U.[30]

$$G(n_j) - \phi_j = U \geq 0. \tag{28}$$

Because G is invertible at n_j by (27), (28) determines a demand function $n_j(\phi_j) = G^{-1}(U + \phi_j)$. By substituting (28) into (25), intermediary j solves

$$\max_{\phi_j} \{G\left(n_j(\phi_j)\right) - c - U\}n_j(\phi_j). \tag{29}$$

Because $G - c$ is the surplus per trader, U determines the share of the surplus left for each trader. There are two possibilities. First, if $U = 0$, then some traders must be excluded from any markets ($n_0 > 0$). Second, if $U > 0$, then all traders must be participating in markets ($n_0 = 0$). **Lemma 7** summarizes competition between intermediaries.

[30] This is the market utility approach commonly used in the competitive search literature (Galenianos and Kircher 2012). Alternatively, we could analyze each intermediary's strategic influence on the value of U. A previous version of the paper considers this and obtains a qualitatively similar result.

Lemma 7 *In a fee-setting equilibrium, either*

(i) $U = 0$ *and for all* $j \in J \backslash \{0\}$, $(G(n_j) - c)\, n_j$ *is maximized at* $n_j = n^m \in \left(n^*, \frac{\overline{n}}{j}\right)$ *and* $\phi_j^* = G(n^m)$, *or*

(ii) $U = H\left(\frac{\overline{n}}{j}\right) > 0$ *and for all* $j \in J \backslash \{0\}$, $(G(n_j) - c)\, n_j$ *is increasing at* $n_j = \frac{\overline{n}}{j}$ *and* $\phi_j^* = c - G'\left(\frac{\overline{n}}{j}\right) \frac{\overline{n}}{j} \in \left(c, G\left(\frac{\overline{n}}{j}\right)\right)$.

Lemma 7 shows that competing intermediaries either (i) behave as if each of them is a monopoly, leaving some traders excluded from markets and extracting all the surplus from participating traders, or (ii) accommodate all traders and extract some, but not all, surplus from traders. In both cases, intermediaries' negotiation power comes from the negative externality among traders. However, for the latter case, each trader receives a surplus that is equal to *his marginal contribution to the surplus generated by one market.* **Figure 5** illustrates the latter case with two intermediaries.

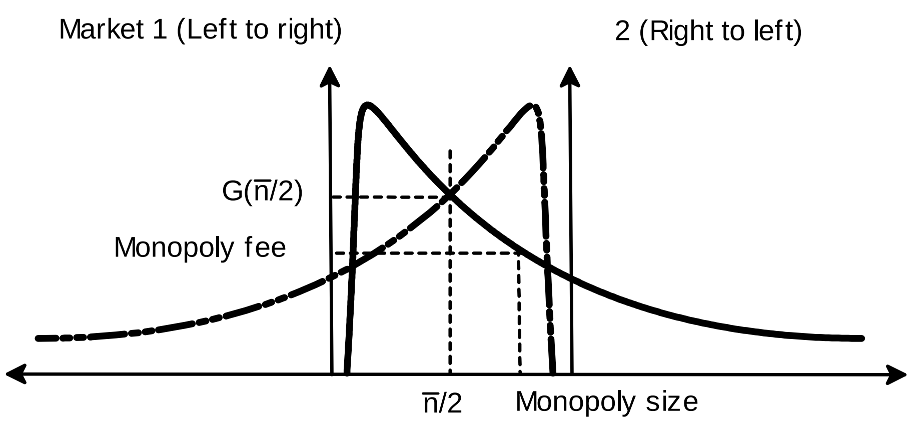


Figure 5. Two markets. The length between the two vertical axes represents the number of potential traders $\overline{n}$. The size of market 1 is measured from the left axis to the right, while the size of market 2 is measured in the opposite direction.

In **Figure 5**, the shared market size is smaller than the monopoly size (i.e., $n^* < \frac{\overline{n}}{2} < n^m$). The equilibrium fee is $\phi^* = c - G'\left(\frac{\overline{n}}{2}\right) \frac{\overline{n}}{2} < G\left(\frac{\overline{n}}{2}\right)$. At the shared market size, the total profit is increasing in the market size, which creates an incentive to set the fee lower than $G\left(\frac{\overline{n}}{2}\right)$. However, such an incentive is not strong enough to compete away profits. Because the slope of G at the shared market size is negative, i.e., the demand is inelastic, the equilibrium fee is higher than marginal costs, and intermediaries earn positive profits.[31] We discuss the implication of free entry into intermediation as well as other extensions in the Appendix.

6 Conclusion

Large and small markets are different because they aggregate different amount of information. As long as traders rationally anticipate this difference, their ex ante welfare is also affected. We showed that negative participation externalities can arise in a standard trading model subject to information asymmetries. This result implies that negative externalities in financial markets may be more relevant than usually believed. Also, the model predicts that (i) multiplicity in traders' motives arise only in a large market subject to an aggregate endowment shock, (ii) a large aggregate shock can create a new type of market illiquidity that can not be resolved by increasing the number of traders. In related works, Spiegel and Subrahmanyam (1992) show that when traders are limited to using market orders, negative externalities can arise due to price volatility. Foucault and Menkveld (2008) argue that order fragmentation enhances liquidity by reducing limit order congestion. Identifying and quantifying different sources of negative externalities are important tasks for a better understanding of trading behavior and the design of market structures.

Empirical studies suggest that some traders have different needs from others, which causes sorting of traders into different markets (Ready 2009; Cantillon and Yin 2010). The current model leaves no room for such horizontal differentiation, because traders are identical when markets are formed. However, as Economides and Siow (1988) showed, incorporating horizontal differentiation is likely to amplify the force for market fragmentation.

Finally, in this paper we focused on the implications of information aggregation on

[31] This result is related to Kreps and Scheinkman (1983), who show that Bertrand competition combined with capacity constraints yields Cournot outcomes. Intermediaries in our model are not subject to physical capacity constraints, but they use *a trading rule*, which creates negative externalities among traders. This works as an *endogenous* capacity constraint.

traders. If the equilibrium market prices provide useful information outside financial markets (e.g. for guiding real investment decisions), then the socially optimal market structure should reflect the benefit of price discovery outside the financial markets. If endogenous market structure fails to internalize this additional informational externality, then it is likely to be suboptimal. Thus, our analysis of market structure should be taken only as a starting point for more general analysis of the socially optimal market structure.

References

[1] 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.

[2] Brown, W. O., Mulherin J. H., and M. D. Weidenmier (2008): "Competing with the New York Stock Exchange." *Quarterly Journal of Economics*, 1679-1719.

[3] Cantillon, E. and P. Yin (2010): "Competition between Exchanges: Lessons from the Battle of the Bund." *Working paper.*

[4] Cantillon, E. and P. Yin (2011): "Competition between Exchanges: A Research Agenda." *International Journal of Industrial Organization*, 29, 329-336.

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

[6] Economides, N., and A. Siow (1988): "The Division of Markets Is Limited by the Extent of Liquidity (Spatial Competition with Externalities)." *American Economic Review*, 78, 1, 108-121.

[7] Foucault, T., and A. J. Menkveld (2008): "Competition for Order Flow and Smart Order Routing System." *Journal of Finance*, 63, 1, 119-158.

[8] Foucault, T., and C. A. Parlour (2004): "Competition for Listings," *Rand Journal of Economics*, 35, 2, 329-355.

[9] Galenianos M., and P. Kircher (2012): "On the Game-Theoretic Foundations of Competitive Search Equilibrium." *International Economic Review*, 53, 1, 1-21.

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

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

[12] Kreps, D.M., and J. A. Scheinkman (1983): "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics*, 14, 2, 326-337.

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

[14] Madhavan, A. (1992): "Trading Mechanisms in Securities Markets." *Journal of Finance*, 47, 2, 607-641.

[15] Madhavan, A. (2000): "Market Microstructure: A Survey." *Journal of Financial Markets*, 3, 205-258.

[16] Manzano, C., and X. Vives (2011): "Public and Private Learning from Prices, Strategic Substitutability and Complementarity, and Equilibrium Multiplicity." *Journal of Mathematical Economics*, 47, 346–369.

[17] Marin, J. M., and R. Rahi (2000): "Information Revelation and Market Incompleteness." *Review of Economic Studies*, 67, 563-579.

[18] Naik, N. Y., Neuberger, A., and S. Viswanathan (1999): "Trade Disclosure Regulation in Markets with Negotiated Trades." *The Review of Financial Studies*, 12, 4, 873-900.

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

[20] Pagano, M. (1989): "Trading Volume and Asset Liquidity." *Quarterly Journal of Economics*, 104, 2, 255-274.

[21] Pagnotta, E. (2013): "Asset Pricing Frictions in Fragmented Markets." *Working paper.*

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

[23] Ready, M. J. (2009): "Determinants of Volume in Dark Pools." *Working paper.*

[24] Santos, T. and J. A. Scheinkman (2001): "Competition among Exchanges." *Quarterly Journal of Economics*, 116, 3, 1027-1061.

[25] Schlee, E. E. (2001): "The Value of Information in Efficient Risk-Sharing Arrangements." *American Economic Review*, 91, 3, 509-524.

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

[27] Vives, X. (2008): "Information and Learning in Markets." *Princeton University Press.*

[28] Weretka, M. (2011): "Endogenous Market Power." *Journal of Economic Theory*, 146, 2281-2306.

7 Appendix

The appendix has four parts. Section 7.1 contains all the proofs for the results presented in the main body of the paper. Section 7.2 describes the market-clearing rule. Section 7.3 offers discussion of various extensions to the market game studied in Section 5. Section 7.4 contains background results for Sections 3 and 4 (Lemma A1-A16) and their proofs.

To save spaces, we use the following notations throughout the Appendix:

$$\alpha \equiv \frac{\rho^2}{\tau_\varepsilon \tau_x},\ c_s \equiv \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_v},\ N_u \equiv 1 + (1-u)\, n.$$

7.1 Proofs for the main text

Proof of Proposition 1

(a) See Lemma A3(a).
(b)(e) See Lemma A7 and A8.
(c)(d) See Lemma A8. ■

Proof of Proposition 2

See Lemma A6. ■

Proof of Proposition 3

(a) See Lemma A9.
(b) See Lemma A13(a) ■

Proof of Lemma 1

See Lemma A1. ■

Proof of Proposition 4

(a) See Lemma A10, A11 and A13(c).
(b) See Lemma A11(c).
(c) See Lemma A11(d) and its proof. ■

Proof of Lemma 2

See Lemma A5(b) and note that $\frac{\tau_v}{\tau} = A\,(T; t, w)$ and $u \frac{\frac{1-\varphi}{\varphi}+w}{\frac{1-\varphi}{\varphi}+w\{1+(1-u)n\}} = B\,(T; u)$. ■

Proof of Lemma 3

(a) See Lemma A5(a)(c).
(b)(c) See Lemma A12. ■

Proof of Lemma 4

See Lemma A14. ■

Proof of Lemma 5

First, (19) is the optimality condition for traders in N_j. Suppose neither (i) nor (ii) holds. Then one of the following two must be true: (a) $G\,(n_j) < G\,(n_j + 1)$ for two markets

$j = 1, 2$; or (b) $G(n_j) < G(n_j + 1)$ for only one market and either $n_0 < 0$ or (21) is violated. For (a), if $G(n_1) - \phi_1 \leq G(n_2) - \phi_2$, any trader in N_1 has an incentive to move to market 2, and vice versa. Contradiction. For (b) If $n_0 > 0$, a trader in N_0 wants to participate in N_j. Note that $G(n_j) < \overline{\phi}$ and $n_j < n^* < n_k$ for all $k \neq j$. If there is $k \neq j$ such that $G(n_k + 1) - \phi_k > G(n_j) - \phi_j$, any trader in N_j wants to move to N_k. If there is $k \neq j$ such that $\overline{\phi} - \phi_j > G(n_k) - \phi_k$, $n^* - n_j$ traders in N_k want to move to N_j. Contradiction. ■

Proof of Lemma 6

Let $n(\phi)$ be the market size determined in a stable participation equilibrium for given $\phi \in [c, \overline{\phi}]$. From **Lemma 5**, either (i) the market operates at the decreasing part of G so that $n^* \leq n(\phi) < \overline{n}$ and $G(n(\phi)) = \phi$, or (ii) the market operates at the increasing part of G so that $n(\phi) = \overline{n} < n^*$ and $\phi \leq G(\overline{n})$. Therefore, the intermediary faces the demand

$$n(\phi) = \begin{cases} 0 & if \quad \phi > \overline{\phi} \\ 0 & if \quad \phi \in (G(\overline{n}), \overline{\phi}] \text{ and } \overline{n} \leq n^* \\ G^{-1}(\phi) & if \quad \phi \in (G(\overline{n}), \overline{\phi}] \text{ and } n^* < \overline{n} \\ \overline{n} & if \quad \phi \leq G(\overline{n}) \end{cases}.$$

First define a marginal profit function $H(n) \equiv \frac{\partial\{(G(n)-c)n\}}{\partial n} = G(n) + G'(n)n - c$. This is the derivative of the total surplus created by a single market. At the optimal market size, $G'(n^*) = 0$ and $H(n^*) = G(n^*) - c > 0$. Because the total surplus is still increasing in the market size, the intermediary maximizes the profit by having more than n^* traders.

If $c \in \left(\underline{\phi}, \overline{\phi}\right)$, the GFT curve intersects with the marginal cost curve twice, so $n^c > n^*$ exists. By setting $\phi = \overline{\phi}$, the profit is $(\overline{\phi} - c)n^* > 0$. By setting $\phi = c$, the profit is 0. Because the demand changes continuously with ϕ in $[c, \overline{\phi}]$, there is an optimal fee in $[c, \overline{\phi}]$ that maximizes the profit. Because the demand function is invertible, the intermediary's problem can be written as $\max_{n \in [n^*, n^c]} (G(n) - c)n$. The first-order condition is $H(n) = G'(n)n + G(n) - c = 0$. Note that $H(n^*) = \overline{\phi} - c > 0$ and $H(n^c) = G'(n^c)n^c < 0$. Therefore, $n^m \in (n^*, n^c)$ maximizes the total surplus. ■

Proof of Lemma 7

The first-order condition to (29) is $\{H(n_j) - U\} \frac{\partial n_j(\phi_j)}{\partial \phi_j} = 0$, where $\frac{\partial n_j(\phi_j)}{\partial \phi_j} = \frac{1}{G'(n_j)} < 0$ by the implicit function theorem. Therefore, the optimality of ϕ_j^* requires $H(n_j) = U$ in equilibrium and $\phi_j^* = G(n_j) - U = G(n_j) - H(n_j) = c - G'(n_j)\, n_j$. First, for $U = H(n_j) = 0$ to hold in equilibrium, the total surplus must be maximized at n_j and some traders must be excluded from markets. Therefore, $n_j = n^m < \frac{\overline{n}}{j_{\max}}$. Second, for $U = H(n_j) > 0$ to hold in equilibrium, the total surplus must be increasing at n_j and all traders must participate in markets. Therefore, $n_j = \frac{\overline{n}}{j_{\max}}$. Hence, in a symmetric equilibrium

$$\phi_j^* = \begin{cases} G(n^m) & \text{if } n^m < \frac{\overline{n}}{j_{\max}} \text{ solves } H(n) = 0, \\ c - G'\left(\frac{\overline{n}}{j_{\max}}\right)\frac{\overline{n}}{j_{\max}} & \text{otherwise.} \end{cases}$$ ■

7.2 Market-clearing rule

We use the market-clearing rule described on page 321 in Kyle (1989). A limit order $q_i(p; H_i)$ is allowed to be any convex-valued, upper-hemicontinuous correspondence that maps prices p into non-empty subsets of the closed infinite interval $[-\infty, \infty]$. An intermediary calculates the set of market-clearing prices and quantity allocation. 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 intermediary 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. This guarantees that infinite prices and quantities do not occur in equilibrium.

7.3 Discussion of market formation game

7.3.1 Free entry

If positive profits invite a new entry, it makes a shared market size smaller and closer to n^*. Because total surplus is increasing near $n = n^*$, the case (ii) in **Lemma 7** applies, and the equilibrium fee approaches $\phi_j^* = c - G'(n^*)\, n^* = c$. Therefore, entry continues until $j^* \equiv \frac{\overline{n}}{n^*}$ markets are established and all traders participate in one of the markets. When the optimal market size n^* is small, j^* is larger and free entry leads to a more fragmented market structure. This is typically the case if the trading environment is subject to a high level of information asymmetry.

Because j^* is not necessarily an integer number, the analysis above is only an approximation. Let j' be the maximum integer that satisfies $\overline{n} > n^* j'$. When j' markets compete, they all make small profits in equilibrium because each market size is close to but larger than n^*. If $j' + 1$ markets compete, however, there is no stable participation equilibrium characterized by (26) and (27), because at least one market must be operating on an increasing part of its GFT curve. This non-existence problem can be avoided by the following "authorization" procedure: j' intermediaries are randomly chosen from a pool of intermediaries who *applied* to open a market. Selected j' intermediaries will make small profits in equilibrium, but a large number of applications drive the expected profits down to zero.

7.3.2 Fixed costs

If intermediaries incur a fixed cost to set up a market, the lower level of fixed cost leads to the larger number of markets in our model. Suppose that the fixed setup cost is greater than the monopoly profit. Then no market would open. If the fixed setup cost is smaller than the monopoly profit, but greater than the level of profit that each of two competing intermediaries would make, then only one market would open and runs monopoly profit. If the fixed setup cost becomes smaller so that each of two competing intermediaries can cover it by its profit, but not small enough to be covered by the profit of each of three competing intermediaries makes, then two markets would open. Thus, the model predicts that a decrease in the fixed cost would lead to more market fragmentation.

7.3.3 Competition by transaction fees

Suppose that on top of the payment of pq_i traders face the quadratic transaction cost:

$$-\frac{c_q}{2}q_i^2,\ c_q > 0$$

Because the form of transaction cost maintains the linearity of equilibrium, all the informational property of equilibrium is not affected by the value of c_q.

Lemma

(a) *The equilibrium existence condition is same as* (14)*.*

(b) *The ex ante GFT are*

$$G^c = \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}u\frac{n-1}{n}\left(1-\frac{n+1}{n-1}\frac{w\varphi}{1-\varphi}\right)\frac{\tau_v}{\tau}\frac{\frac{1-\varphi}{\varphi}+w}{\frac{1-\varphi}{\varphi}+w\left\{1+(1-u)\,n\right\}}\left(1-\frac{c_q\tau}{\rho+c_q\tau}\right)\right).$$

(c) *Fee revenue is maximized by setting* $c_q(n)=\frac{\rho}{\tau}$*, which implies* $1-\frac{c_q\tau}{\rho+c_q\tau}=\frac{1}{2}$.

Proof. See proof of **Lemma A15** below.

The results above show that $c_q > 0$ introduces an additional term $1-\frac{c_q\tau}{\rho+c_q\tau}$, which decreases in τ. Therefore, for a fixed $c_q > 0$, intermediaries profiting from volume would strengthen the negative externalities. However, if traders anticipate that the fee level will be chosen to maximize profits for a given n, then the fee discount is $\frac{1}{2}$ independent of n.

It can be shown that trade equilibrium exists with subsidies in the form of volume discount $c_q < 0$, and that there exists a level of c_q which implements a price-taking equilibrium. If intermediaries can commit to the level of c_q at the ex ante stage, then they would commit to implementing a price-taking equilibrium, and try to maximize the entry fee revenues. From **Lemma 3**, the relevant GFT curve will be higher than the baseline case. On the other hand, if they cannot commit to the level of c_q, then traders rationally anticipate that c_q will adjust to n to maximize transaction fee revenues. From part **(b)** and **(c)**, the relevant GFT curve will be lower. Thus, depending on the commitment assumption, the relevant GFT curves shift either up or down, but the economic force arising from negative externalities would be similar.

7.3.4 Sequentially opening markets

Suppose there are two markets and they open sequentially. Traders in the second market observe a market-clearing price in the first market before they trade. We characterize equilibrium and gains from trade in the second market assuming $t = u = w = 1$. Suppose that the first market has $n_1 + 1$ traders and the second market has $n_2 + 1$ traders. Each trader in the second market has the information set $H_{i2} = \{s_i, e_i, p_1, p\}$, where p_1 is the publicly observed market-clearing price in the first market and p is the equilibrium price in the second market to be determined. Ex ante gains from trade for the second market $G_2(n_2; n_1)$ depends on n_1. We define the optimal second market size for given n_1 by

$$n_2^*(n_1) \equiv \arg\max_{n_2} G_2(n_2; n_1).$$

With $t = u = w = 1$, the equilibrium existence condition imposes the explicit lower bound on the size of the first market:

$$\underline{n} \equiv \frac{1}{1 - 2\varphi} < n_1.$$

Using $\underline{n}$, the optimal size for the first market (see **Lemma A2** below) is

$$n_1^* = \underline{n}\left(1 + \sqrt{1 + \frac{1}{\underline{n} c_s \varphi}}\right), \text{ where } c_s \equiv \frac{\tau_\varepsilon}{\tau_v + \tau_\varepsilon}.$$

Lemma
Assume $t = u = w = 1$. In market 2, ex ante GFT and the optimal market size are

$$\begin{aligned} G_2(n_2; n_1) &= \frac{1}{2\rho} \log\left(1 + \frac{\rho^2}{\tau_v \tau_x} \frac{\frac{n_2 - 1}{n_2} - 2\varphi}{1 - \varphi} \frac{1 - c_s}{1 + c_s(1 + n_1 + n_2)\varphi}\right), \\ n_2^*(n_1) &= \underline{n}\left(1 + \sqrt{1 + \frac{1}{\underline{n} c_s \varphi} + \frac{1 + n_1}{\underline{n}}}\right) > n_1^*. \end{aligned}$$

Proof. See proof of **Lemma A16** below.

The results above show that the optimal market size for the second market is larger than that for the first market, but it is still finite for any given finite n_1. Because the first market imposes informational externalities on the second market, otherwise identical two markets have very different gains from trade curves. The relationship between $G_2(n_2; n_1)$ and $G_1(n_1)$ is studied in **Lemma A16**.

With sufficiently large number of potential traders, multiple markets can still coexist when they open sequentially. To see this, define

$$n_2^m(n_1) \equiv \arg\max_{n_2} n_2 \left(G_2(n_2; n_1) - c\right).$$

This is the maximum surplus the second market can extract taking n_1 as given. Thus, if there are more than $n^m + n_2^m(n^m)$ potential traders, there is no reason for the two markets to lower fees from $\phi_1 = G_1(n^m)$ and $\phi_2 = G_2(n_2^m(n^m); n^m)$ respectively. An extension with three or more markets is straightforward.

7.3.5 Arbitrage across markets

The analysis of arbitrage would require the following. Consider two markets 1 and 2 that open sequentially, and $n_1 + 1$ traders who trade only in market 1, $n_2 + 1$ traders who trade only in market 2, and $n_a + 1$ traders who trade in both markets. In market 1, $n_1 + 1$ local traders and $n_a + 1$ arbitrageurs interact. Let $G_1(n_1, n_a)$ be gains from trade for a local trader in market 1, and $G_{a1}(n_1, n_a)$ for an arbitrageur in market 1. In market 2, $n_2 + 1$ local traders and $n_a + 1$ arbitrageurs interact. Because they all observe a market-clearing price in market 1, their gains from trade depend on n_1 and n_a. Let $G_2(n_2, n_a; n_1, n_a)$ and $G_{a2}(n_2, n_a; n_1, n_a)$ be gains from trade for a local trader and an arbitrageur in market 2. Suppose that the participation fee cannot discriminate different traders. Let ϕ_1 be the fee in market 1 and ϕ_2

be the fee in market 2. Then, participation equilibrium must satisfy:

$$0 \leq G_1(n_1, n_a) - \phi_1 = G_2(n_2, n_a; n_1, n_a) - \phi_2 = G_{a1}(n_1, n_a) + G_{a2}(n_2, n_a; n_1, n_a) - (\phi_1 + \phi_2).$$

This illustrates the complication that arises for any analysis with ex ante heterogeneity. There must be a gains from trade function for each group of traders with different ex ante characteristics, and each function must have as its arguments the numbers of traders for each characteristics group in a given market. These functions can be seen as endogenous matching functions, one for each trader. In the example above, there are three groups (one local group in each market and a group of arbitrageurs) and the size of each group will be endogenously determined by participation decisions. Unfortunately, even for a simple setup (such as $n_a = 1$) the characterization of ex ante gains from trade is extremely complicated. We leave the analysis of arbitrage, and more generally the impact of ex ante heterogeneity on market structure, for future works.

7.4 Results for Sections 3 and 4.

We present a series of results for the general case with $t > 0$ and $w > 0$. This subsection is organized as follows:

1. Two benchmark cases (Lemma A1 and A2)
2. Trade equilibrium (Lemma A3).
3. Gains from trade (Lemma A4 and A5).
4. Multiplicity (Lemma A6).
5. Information aggregation (Lemma A7 for $w < 1$ and Lemma A8 and A9 for $w = 1$).
6. Negative externalities (Lemma A10 for $w < 1$ and Lemma A11 for $w = 1$).
7. Price-taking equilibrium (Lemma A12).
8. Limits $t \to 0$, $w \to 0$, and $u \to 0$ (Lemma A13).
9. τ_ε and τ_x that depends on n (Lemma A14).
10. Equilibrium with the transaction fee (Lemma A15).
11. Sequentially opening markets (Lemma A16).

These results are used to prove the results presented in the main text.

7.4.1 Two benchmark cases

We compare two benchmark cases: (i) $t = 0$ or $w = 0$, and (ii) $t = u = w = 1$ and discuss their contrasting implications for market structures.[32] When $t = 0$, $v = v_0$ and any signal $s_i = v_1 + \varepsilon_i$ is irrelevant. When $w = 0$, all traders have an identical signal $s = v_1 + \epsilon_0$ and price reveals no new information. In both cases, there is no negative externality due to information revelation and the optimal market size is infinite.

Lemma A1 (symmetric information: $t = 0$ or $w = 0$**)**

[32] Pagano (1989) studies the case with $t = 0$ and $u = 1$, while Diamond and Verrecchia (1981) and Madhavan (1992) study the case with $t = u = w = 1$. However, radically different implications for market structures were never explored.

(a) *A trade equilibrium exists for* $n > 1$ *and*

$$\begin{aligned} q_i(p; H_i) &= \frac{n-1}{n}\left\{\frac{\tau_v}{\rho}\frac{\sqrt{t}c_s}{1-tc_s}s - e_i - \frac{\tau_v}{\rho}\frac{1}{1-tc_s}p\right\}, \\ p^* &= \sqrt{t}c_s s - \frac{\rho}{\tau_v}\frac{1}{1-tc_s}\overline{e}, \\ q_i(p^*; H_i) &= -\frac{n-1}{n}(e_i - \overline{e}), \end{aligned}$$

where $s = v_1 + \epsilon_0$ *is an identical signal.*

(b) $G = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v \tau_x}\frac{n-1}{n}(1 - tc_s)\,u\right)$ *is increasing in* n *and decreasing in* τ_v.

Setting $t = 0$ in the above yields the results for $t = 0$ with any $w \in [0, 1]$. When $t = 0$ or $w = 0$, the information about the payoff v_0 is symmetric and the equilibrium price reveals no information about the payoff. As long as there is a diversifiable risk ($u > 0$), the ex ante GFT *per trader* increase in n. This positive externality exists because each trader creates more risk sharing opportunity for the other traders due to the idiosyncratic component of his endowment. In the limit as n approaches infinity, everyone trades to hold the average position ($q_i(p^*; H_i) + e_i = \frac{1}{n}e_i + \frac{n-1}{n}\overline{e}$).

On the other hand, when $t = u = w = 1$, the information about the payoff v_1 is asymmetric. Also, both endowments and noise in signals are i.i.d. across traders. In this case, φ is independent of n, and optimal market size n^* has an explicit solution.

Lemma A2 ($t = u = w = 1$)

(a) *A trade equilibrium exists if and only if* $1 < n$ and $\frac{n+1}{n-1} < \frac{\rho^2}{\tau_\varepsilon \tau_x}$.

In a trade equilibrium, $\varphi = \left(1 + \frac{\rho^2}{\tau_\varepsilon \tau_x}\right)^{-1}$ *and*

$$\begin{aligned} q_i(p; H_i) &= \left(\frac{n-1}{n} - 2\varphi\right)\left\{\frac{\tau_\varepsilon}{\rho}s_i - e_i - \frac{\tau_\varepsilon}{\rho}\frac{1 + c_s n\varphi}{c_s(1 + n\varphi)}p\right\}, \\ p^* &= \frac{c_s(1 + n\varphi)}{1 + c_s n\varphi}\left(\overline{s} - \frac{\rho}{\tau_\varepsilon}\overline{e}\right), \\ q_i(p^*; H_i) &= \left(\frac{n-1}{n} - 2\varphi\right)\left\{\frac{\tau_\varepsilon}{\rho}(s_i - \overline{s}) - (e_i - \overline{e})\right\}, \\ G &= \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v \tau_x}\frac{\frac{n-1}{n} - 2\varphi}{1 - \varphi}\frac{1}{1 + \frac{\tau_\varepsilon}{\tau_v}(1 + n\varphi)}\right), \\ n^* &= \frac{1}{1 - 2\varphi} + \sqrt{\frac{1}{1 - 2\varphi}\left(\frac{1}{c_s\varphi} + \frac{1}{1 - 2\varphi}\right)}. \end{aligned}$$

(c) *A price-taking equilibrium always exists.*

In equilibrium, φ and p^* are same as above and

$$\begin{aligned} q_i(p; H_i) &= (1-\varphi)\left\{\frac{\tau_\varepsilon}{\rho}s_i - e_i - \frac{\tau_\varepsilon}{\rho}\frac{1+c_s n\varphi}{c_s(1+n\varphi)}p\right\}, \\ q_i(p^*; H_i) &= (1-\varphi)\left\{\frac{\tau_\varepsilon}{\rho}(s_i - \overline{s}) - (e_i - \overline{e})\right\}, \\ G^{pt} &= \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x}\frac{n}{1+n}\frac{1-\varphi}{1+\frac{\tau_\varepsilon}{\tau_v}(1+n\varphi)}\right), \\ n^*_{pt} &= \sqrt{\frac{1}{c_s\varphi}} < n^*. \end{aligned}$$

(c) $\lim\limits_{n\to\infty} G = 0$ and $\lim\limits_{n\to\infty} G^{pt} = 0$.

From part (**a**), if $1 < \frac{\rho^2}{\tau_\varepsilon\tau_x}$, a trade equilibrium exists for all $n > \frac{1}{1-2\varphi}$. In the limit as n goes to infinity, the optimal order is well defined (β_s, β_e, β_p all have positive and finite limits). Therefore, the limit market is perfectly liquid ($\lim\limits_{n\to\infty}\frac{1}{n\beta_p} = 0$) and it is tempting to argue that "the thick market is informationally efficient" because p^* converges to v almost surely as n increases to infinity. However, the optimal market size is in fact finite, and as part (**c**) shows, the limit market offers vanishing incentive for each trader to participate.

Lemma A1 and **A2** present contrasting results about the nature of externalities. Economides and Siow (1988) and Pagano (1989) study endogenous market participation without asymmetric information. In their models, the optimal market size is infinite as shown in **Lemma A1**, but exogenous transaction costs or search costs limit the market size. Therefore, their models predict the rise of a single market as these costs decrease. We show that with asymmetric information, the optimal market size is finite because of endogenous negative externalities among traders. Therefore, even if transaction and search costs decrease, market fragmentation may survive for assets subject to information asymmetries.

Proof of Lemma A1

The proof follows the same step with that of **Lemma A3**, **A4** and **A5** except that traders' beliefs are fixed by their prior beliefs ($E_i[v] = 0$ and $Var_i[v] = \frac{1}{\tau_v}$) for $t = 0$, or by homogeneous beliefs based on $s = v_1 + \epsilon_0$ ($E_i[v] = \frac{\tau_\varepsilon s}{\tau_v + \tau_\varepsilon}$ and $Var_i[v] = \frac{1}{\tau_v}\frac{\tau_v + (1-t)\tau_\varepsilon}{\tau_v+\tau_\varepsilon}$) for $w = 0$. There is no need to characterize φ. Because G is a monotonic transformation of $\frac{n-1}{n}$, it is monotonically increasing in n. If $t = 0$ or 1, G is clearly decreasing in τ_v. To show that

G is decreasing in τ_v for $t \in (0,1)$, take a derivative of $\frac{1}{\tau_v}\left(1-\frac{t\tau_\varepsilon}{\tau_v+\tau_\varepsilon}\right)$ to obtain

$$\begin{aligned}
&-\frac{1}{\tau_v^2}\left(1-\frac{t\tau_\varepsilon}{\tau_v+\tau_\varepsilon}\right)+\frac{1}{\tau_v}\frac{t\tau_\varepsilon}{(\tau_v+\tau_\varepsilon)^2} \\
=\ & \frac{-1}{\tau_v^2(\tau_v+\tau_\varepsilon)^2}\left\{(\tau_v+\tau_\varepsilon)(\tau_v+(1-t)\tau_\varepsilon)-t\tau_\varepsilon\tau_v\right\} \\
=\ & \frac{-1}{\tau_v^2(\tau_v+\tau_\varepsilon)^2}\left\{\tau_v^2-2(1-t)\tau_\varepsilon\tau_v+(1-t)\tau_\varepsilon^2\right\}.
\end{aligned}$$

This is negative because $\{(1-t)\tau_\varepsilon\}^2-(1-t)\tau_\varepsilon^2=(1-t)\tau_\varepsilon^2(1-t-1)<0$ implies $\tau_v^2-2(1-t)\tau_v\tau_\varepsilon+(1-t)\tau_\varepsilon^2>0$ for all τ_v. ■

Proof of Lemma A2

First, set $t=u=1$ in (60) to obtain

$$F(k;t=u=w=1)=\left(\frac{\rho^2}{\tau_\varepsilon\tau_x}k^2+1\right)(k-1)=0,$$

for which $k=1$ is the unique solution. Use $k=1$ in (9) and (14) to obtain $\varphi=\left(1+\frac{\rho^2}{\tau_\varepsilon\tau_x}\right)^{-1}$ and verify the second-order condition. The optimal order, the equilibrium price, the quantity traded, and the ex ante GFT are obtained by plugging $t=u=w=k=1$ in those in **Lemma A3** and **Lemma A5**. Because φ does not depend on n, G^{pt} is maximized when $\frac{n}{1+n}\frac{1}{1+c_sn\varphi}$ is maximized, while G is maximized when $\frac{\frac{n-1}{n}-2\varphi}{1+c_sn\varphi}$ is maximized. The first derivative of $\frac{n}{1+n}\frac{1}{1+c_sn\varphi}$ is $\frac{1-c_s\varphi n^2}{(1+n)^2(1+c_sn\varphi)^2}$. Hence, G^{pt} is increasing for $n<\sqrt{\frac{1}{c_s\varphi}}$ and decreasing for $n>\sqrt{\frac{1}{c_s\varphi}}$. The first derivative of $\frac{\frac{n-1}{n}-2\varphi}{1+c_sn\varphi}$ is $\frac{1+2c_s\varphi n-(1-2\varphi)c_s\varphi n^2}{(1+n)^2(1+c_sn\varphi)^2}$. Two solutions for

$$1+2c_s\varphi n-(1-2\varphi)c_s\varphi n^2=0$$

are

$$n^\pm\equiv\frac{c_s\varphi\pm\sqrt{(c_s\varphi)^2+(1-2\varphi)c_s\varphi}}{(1-2\varphi)c_s\varphi}.$$

Because the smaller solution $n^-<0$, G is increasing for

$$\begin{aligned}
n &< n^+ \\
&= \frac{c_s\varphi+\sqrt{(c_s\varphi)^2+(1-2\varphi)c_s\varphi}}{(1-2\varphi)c_s\varphi} \\
&= \frac{1}{1-2\varphi}+\sqrt{\frac{1}{1-2\varphi}\left(\frac{1}{c_s\varphi}+\frac{1}{1-2\varphi}\right)}
\end{aligned}$$

and it is decreasing for $n>n^+$. Because $\frac{1}{1-2\varphi}>1$, $n^+>\sqrt{\frac{1}{c_s\varphi}}$. That $\lim_{n\to\infty}G=\lim_{n\to\infty}G^{pt}=0$ is trivial from the expressions of G and G^{pt}. ■

7.4.2 Trade equilibrium

Lemma A3 (trade equilibrium for $t > 0$ and $w > 0$)

(a) *The second-order condition of the traders' problem is satisfied if and only if*

$$1 < n \text{ and } \frac{n+1}{n-1} < \frac{1-\varphi}{\varphi} + 1 - w. \tag{30}$$

(b) *The optimal order has coefficients*

$$\begin{aligned}
\beta_s &= \frac{\sqrt{t}\left\{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi\right\}}{1 + (1-w)(nw-1)\varphi + (1-t)\frac{\tau_\varepsilon}{\tau_v}\left\{1 + \varphi\left(nw - (1-w)\right)\right\}}\frac{\tau_\varepsilon}{\rho}, \\
\beta_e &= \frac{\sqrt{t}\left\{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi\right\}}{1 + (1-w)(nw-1)\varphi + (1-t)\frac{\tau_\varepsilon}{\tau_v}\left\{1 + \varphi\left(nw - (1-w)\right)\right\}}k, \\
\beta_p &= \frac{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi}{1 + (1-w)(nw-1)\varphi}\frac{\tau}{\rho},
\end{aligned}$$

and the quantity traded at the equilibrium price p^* is $q_i(p^*; H_i) =$

$$\frac{\sqrt{t}\left\{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi\right\}}{1 + \varphi(1-w)(nw-1) + (1-t)\frac{\tau_\varepsilon}{\tau_v}\left\{1 + \varphi\left(nw - (1-w)\right)\right\}}\left\{\frac{\tau_\varepsilon}{\rho}(s_i - \overline{s}) - k(e_i - \overline{e})\right\},$$

where $\tau \equiv (Var_i[v])^{-1}$, $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$ and $\varphi = (1 + \alpha u N_u k^2)^{-1}$ are characterized in the proof.

Proof.
We proceed in four steps:

1) Characterize beliefs $E_i[v_1]$ and $\tau_1 \equiv (Var_i[v_1])^{-1}$.
2) Derive the optimal order $q_i(p; H_i)$.
3) Characterize $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$.
4) Check the second order condition.

[Step 1] Characterize $E_i[v_1]$ and τ_1.
First, from the conjectured order

$$q_i(p; H_i) = \beta_s s_i - \beta_e e_i - \beta_p p \tag{31}$$

and the market-clearing condition, information in p from trader i's perspective is summarized by

$$h_i \equiv \frac{n\beta_p p - q_i}{n\beta_s} = v_1 + \overline{\varepsilon}_{-i} - \frac{\beta_e}{\beta_s}\left(\sqrt{1-u}x_0 + \sqrt{u}\overline{x}_{-i}\right),$$

where $\overline{\varepsilon}_{-i} = \sqrt{1-w}\epsilon_0 + \sqrt{w}\overline{\epsilon}_{-i}$. Hence, $[v_1, s_i, e_i, h_i]^\top$ is jointly normal with mean zero and a variance-covariance matrix

$$\begin{bmatrix} \frac{1}{\tau_v} & \frac{1}{\tau_v} & 0 & \frac{1}{\tau_v} \\ & \frac{1}{\tau_v} + \frac{1}{\tau_\varepsilon} & 0 & \frac{1}{\tau_v} + (1-w)\frac{1}{\tau_\varepsilon} \\ & & \frac{1}{\tau_x} & -\frac{\beta_e}{\beta_s}(1-u)\frac{1}{\tau_x} \\ & & & \frac{1}{\tau_v} + \frac{1}{n\tau_\varepsilon}\left\{w + n(1-w) + \left(\frac{\beta_e}{\beta_s}\right)^2 \frac{\tau_\varepsilon}{\tau_x}(u + n(1-u))\right\} \end{bmatrix}.$$

Let Σ be the variance-covariance matrix of $[s_i, e_i, h_i]^\top$. By Bayes' rule,

$$\begin{aligned} E_i[v_1] &= \left[\frac{1}{\tau_v}, 0, \frac{1}{\tau_v}\right]\Sigma^{-1}[s_i, e_i, h_i]^\top, \\ \tau_1^{-1} &= \tau_v^{-1} - \left[\frac{1}{\tau_v}, 0, \frac{1}{\tau_v}\right]\Sigma^{-1}\left[\frac{1}{\tau_v}, 0, \frac{1}{\tau_v}\right]^\top. \end{aligned}$$

Define

$$\varphi \equiv \left\{1 + \left(\frac{\beta_e}{\beta_s}\right)^2 \frac{\tau_\varepsilon}{\tau_x} u N_u\right\}^{-1}.$$

Matrix algebra shows

$$\begin{aligned} E_i[v_1] &= \frac{\tau_\varepsilon}{\tau_1}\frac{(1-\varphi)s_i + w\varphi\left\{\frac{\beta_e}{\beta_s}N_u e_i + \frac{\beta_p}{\beta_s}(n+1)p\right\}}{1 + (1-w)(nw-1)\varphi}, \\ \tau_1 &= \tau_v + \tau_\varepsilon \frac{1 + (nw - (1-w))\varphi}{1 + (1-w)(nw-1)\varphi} \\ &= \tau_v \frac{1 + (1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1 + (nw-(1-w))\varphi\}}{1 + (1-w)(nw-1)\varphi}. \end{aligned} \quad (32)$$

Also,

$$
\begin{aligned}
\tau^{-1} &\equiv Var_i[v] \\
&= (1-t)\frac{1}{\tau_v} + t\frac{1}{\tau_1} \\
&= \frac{t\tau_v + (1-t)\tau_1}{\tau_v\tau_1} \\
&= \frac{1}{\tau_v}\frac{1+(1-w)(nw-1)\varphi}{1+(1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}} \times \\
&\quad \left(t + (1-t)\frac{1+(1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}{1+(1-w)(nw-1)\varphi}\right) \\
&= \frac{1}{\tau_v}\frac{t\{1+(1-w)(nw-1)\varphi\} + (1-t)\left\{1+(1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}\right\}}{1+(1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}} \\
&= \frac{1}{\tau_v}\frac{1+(1-w)(nw-1)\varphi + (1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}{1+(1-w)(nw-1)\varphi + \frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}.
\end{aligned}
$$

Use $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$, $\alpha \equiv \frac{\rho^2}{\tau_\varepsilon\tau_x}$ and $N_u \equiv 1+(1-u)n$ to write

$$\varphi = \left(1+\alpha u N_u k^2\right)^{-1}. \tag{33}$$

[Step 2] Derive $q_i(p; H_i)$.

We derive the optimal order given the belief $E_i[v_1]$ and τ_1 derived above. From the conjecture (31) and the market-clearing condition $\sum_{j\neq i} q_j + q_i = 0$,

$$
\begin{aligned}
\sum_{j\neq i} q_j &= \beta_s \sum_{j\neq i} s_j - \beta_e \sum_{j\neq i} e_j - n\beta_p p \\
&= -q_i.
\end{aligned}
$$

Solving for the price, we obtain

$$p = p_i + \lambda q_i, \tag{34}$$

where $p_i \equiv \frac{\beta_s}{\beta_p}\overline{s}_{-i} - \frac{\beta_e}{\beta_p}\overline{e}_{-i}$ and $\lambda \equiv \frac{1}{n\beta_p}$. Trader i maximizes $E_i[-\exp(-\rho\pi_i)]$. Because of the normality of v conditional on each trader's information, the objective becomes

$$E_i[v](q_i+e_i) - \frac{\rho}{2}Var_i[v](q_i+e_i)^2 - pq_i \tag{35}$$

subject to (34). The first-order condition and the second-order condition are

$$E_i[v] - \frac{\rho}{\tau}(q_i+e_i) = p_i + 2\lambda q_i = p + \lambda q_i, \tag{36}$$

$$2\lambda + \frac{\rho}{\tau} > 0, \tag{37}$$

From (36), we obtain

$$q_i^* = \frac{E_i[v] - p - \frac{\rho}{\tau} e_i}{\lambda + \frac{\rho}{\tau}}. \tag{38}$$

By substituting $E_i[v] = \sqrt{t} E_i[v_1] = \sqrt{t}\left(k_s s_i + k_e e_i + k_p p\right)$, where k_s, k_e, k_p are given in (32), into (38), obtain

$$q_i^* = \frac{\sqrt{t} k_s s_i - \left(\frac{\rho}{\tau} - \sqrt{t} k_e\right) e_i - \left(1 - \sqrt{t} k_p\right) p}{\lambda + \frac{\rho}{\tau}}.$$

By equating coefficients, we have three equations

$$\beta_s = \frac{\tau_\varepsilon}{\lambda\tau + \rho} \frac{1 - \varphi}{1 + (1 - w)(nw - 1)\varphi} \frac{\tau}{\tau_1} \sqrt{t}, \tag{39}$$

$$\beta_e = \frac{\rho}{\lambda\tau + \rho} \left(1 - \frac{w\varphi}{1 + (1 - w)(nw - 1)\varphi} \frac{\tau_\varepsilon}{\rho} N_u \frac{\beta_e}{\beta_s} \frac{\tau}{\tau_1} \sqrt{t}\right), \tag{40}$$

$$\beta_p = \frac{\tau}{\lambda\tau + \rho} \left(1 - \frac{w\varphi}{1 + (1 - w)(nw - 1)\varphi} \frac{(n+1)\tau_\varepsilon}{\tau_1} \frac{\beta_p}{\beta_s} \sqrt{t}\right). \tag{41}$$

Solving for $\frac{\beta_p}{\beta_s}$ shows that $\frac{\beta_p}{\beta_s} = \frac{\tau_1}{\sqrt{t}\tau_\varepsilon} \frac{1+(1-w)(nw-1)\varphi}{1+\{wn-(1-w)\}\varphi}$. Using (41) and

$$\lambda = \frac{1}{n\beta_p} = \frac{\tau_\varepsilon \sqrt{t}}{n\beta_s \tau_1} \frac{1 + \{wn - (1 - w)\}\varphi}{1 + (1 - w)(nw - 1)\varphi},$$

obtain

$$\beta_s = \frac{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi}{1 + (1 - w)(nw - 1)\varphi} \sqrt{t} \frac{\tau_\varepsilon}{\rho} \frac{\tau}{\tau_1}.$$

Similarly, solving for $\frac{\beta_e}{\beta_s}$ from (39) and (40) shows that

$$\sqrt{t} \frac{\tau_\varepsilon}{\rho} \frac{\beta_e}{\beta_s} = \frac{\tau_1}{\tau} \frac{1 + (1 - w)(nw - 1)\varphi}{1 + \{w(1 - u)n - (1 - w)\}\varphi}. \tag{42}$$

Using

$$\begin{aligned} \frac{\tau}{\tau_1} &= \frac{1}{\tau_1} \left\{(1 - t)\frac{1}{\tau_v} + t\frac{1}{\tau_1}\right\}^{-1} \\ &= \frac{1}{1 + (1 - t)\frac{\tau_\varepsilon}{\tau_v} \frac{1+\varphi(nw-(1-w))}{1+(1-w)(nw-1)\varphi}} \\ &= \frac{1 + (1 - w)(nw - 1)\varphi}{1 + (1 - w)(nw - 1)\varphi + (1 - t)\frac{\tau_\varepsilon}{\tau_v}\left\{1 + \varphi\left(nw - (1 - w)\right)\right\}}, \end{aligned} \tag{43}$$

the optimal order has coefficients

$$\begin{aligned}
\beta_s &= \frac{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}\sqrt{t}\frac{\tau_\varepsilon}{\rho},\\
\beta_e &= \frac{\rho}{\tau_\varepsilon}k\beta_s,\\
\beta_p &= \frac{\tau_1}{\sqrt{t}\tau_\varepsilon}\frac{1+(1-w)(nw-1)\varphi}{1+\{nw-(1-w)\}\varphi}\beta_s,
\end{aligned}$$

and

$$\begin{aligned}
q_i(p;H_i) &= \frac{\sqrt{t}\left\{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi\right\}}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}\\
&\quad\times\left\{\frac{\tau_\varepsilon}{\rho}s_i-ke_i-\frac{\tau_1}{\sqrt{t}\rho}\frac{1+(1-w)(nw-1)\varphi}{1+\{nw-(1-w)\}\varphi}p\right\},
\end{aligned} \tag{44}$$

where τ_1 is given in (32). Coefficients can be written as

$$\begin{aligned}
\beta_s &= \frac{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}\sqrt{t}\frac{\tau_\varepsilon}{\rho},\\
\beta_e &= \frac{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}\sqrt{t}k,\\
\beta_p &= \frac{1+(1-w)(nw-1)\varphi+\frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}\frac{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi}{1+(1-w)(nw-1)\varphi}\frac{\tau_v}{\rho}.
\end{aligned}$$

[Step 3] Characterize $k\equiv\frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$.
From (42),

$$k\equiv\frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}=\frac{\tau_1}{\sqrt{t}\tau}\frac{1+(1-w)(nw-1)\varphi}{1+\{w(1-u)n-(1-w)\}\varphi}. \tag{45}$$

Plug $\varphi=(1+\alpha uN_uk^2)^{-1}$ into the above expression. After simplification, this becomes a cubic equation in k:

$$\begin{aligned}
F(k) &\equiv \left(\alpha uN_uk^2+w\right)\left\{\sqrt{t}k-\left(1+(1-t)\frac{\tau_\varepsilon}{\tau_v}\right)\right\}\\
&\quad+wn\left\{(1-u)\sqrt{t}k-\left(1-w+(1-t)\frac{\tau_\varepsilon}{\tau_v}\right)\right\}=0.
\end{aligned}$$

Below we use $1+(1-t)\frac{\tau_\varepsilon}{\tau_v}=\frac{1-tc_s}{1-c_s}$ and $1-w+(1-t)\frac{\tau_\varepsilon}{\tau_v}=\frac{1-tc_s}{1-c_s}-w$ to write

$$F(k)\equiv\left(\alpha uN_uk^2+w\right)\left(\sqrt{t}k-\frac{1-tc_s}{1-c_s}\right)+\ wn\left\{(1-u)\sqrt{t}k-\left(\frac{1-tc_s}{1-c_s}-w\right)\right\}=0. \tag{46}$$

Because $\lim_{k\to-\infty}F(k)=-\infty$, $\lim_{k\to\infty}F(k)=\infty$ and $F(0)<0$, (46) has at least one and at most three positive solutions.

[Step 4] Check the second order condition.
Plug $\lambda = \frac{1}{n\beta_p}$ into (37) to obtain $\frac{2}{n\beta_p} + \frac{\rho}{\tau} > 0 \Leftrightarrow 0 < 1 + \frac{\tau}{\rho}\frac{2}{n\beta_p}$. Then,

$$\begin{aligned} & 1 + \frac{\tau}{\rho}\frac{2}{n\beta_p} \\ = \; & 1 + \frac{2}{n}\frac{1 + (1-w)(nw-1)\varphi + (1-t)\frac{\tau_\varepsilon}{\tau_v}\{1 + \varphi(nw - (1-w))\}}{\left\{\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi\right\}}\frac{\tau}{\tau_1}\frac{1 + \varphi(nw - (1-w))}{1 + (1-w)(nw-1)\varphi} \\ = \; & 1 + \frac{2\{1 + (nw - (1-w))\varphi\}}{n - 1 - \{n(1+w) - (1-w)\}\varphi} \\ = \; & \frac{(n+1)\{1 - (1-w)\varphi\}}{n - 1 - \{n(1+w) - (1-w)\}\varphi} \\ = \; & \frac{(n+1)\{1 - (1-w)\varphi\}}{n\{1 - (1+w)\varphi\} - \{1 + (1-w)\varphi\}}. \end{aligned}$$

Because $1 - (1-w)\varphi > 0$,

$$\begin{aligned} (37) \; & \Leftrightarrow \; n\{1 - (1+w)\varphi\} > \{1 + (1-w)\varphi\} \\ & \Leftrightarrow \; 1 < n \text{ and } \frac{n+1}{n-1} < \frac{1-\varphi}{\varphi} + 1 - w. \end{aligned} \quad \blacksquare$$

7.4.3 Gains from trade

Next, we derive interim and ex ante gains from trade both for a trade equilibrium and for a price-taking equilibrium. Denote the interim profit, the interim GFT, and the ex ante GFT in a price-taking equilibrium by Π_i^{pt}, G_i^{pt}, and G^{pt}. First, we derive G_i^{pt} and relate it to G_i taking $E_i[v]$ and τ as given (**Lemma A4**). Second, we derive G^{pt} and G using the expression of $E_i[v]$ and τ (**Lemma A5**).

Lemma A4 (interim gains from trade: G_i^{pt} and G_i)

(a) *In a price-taking equilibrium, the interim GFT are* $G_i^{pt} = \frac{\tau}{2\rho}(E_i[v] - p - \frac{\rho}{\tau}e_i)^2$.

(b) *In a trade equilibrium, the interim GFT are* $G_i = \left(1 - \widetilde{\lambda}\right) G_i^{pt}$,

where $\widetilde{\lambda} \equiv \left(\frac{\left(w - \frac{1-w}{n}\right)\varphi + \frac{1}{n}}{1-\varphi}\right)^2 \in [0,1]$.

Proof. By plugging the optimal order (38) into the interim profit (35) and simplifying, obtain

$$\Pi_i = \left(1 - \left(\frac{\lambda\tau}{\rho + \lambda\tau}\right)^2\right)\left\{\frac{\tau}{2\rho}(E_i[v] - p)^2 + pe_i\right\} + \left(\frac{\lambda\tau}{\rho + \lambda\tau}\right)^2\left(E_i[v]e_i - \frac{\rho}{2\tau}e_i^2\right). \quad (47)$$

The interim no-trade profit is $\Pi_i^{nt} = E_i[v]e_i - \frac{\rho}{2\tau}e_i^2$. By setting, $\lambda = 0$ in (47), obtain

$$\Pi_i^{pt} = \frac{\tau}{2\rho}(E_i[v] - p)^2 + pe_i.$$

It is straightforward to show

$$\begin{aligned} G_i^{pt} &\equiv \Pi_i^{pt} - \Pi_i^{nt} \\ &= \frac{\tau}{2\rho}\left(E_i[v] - p - \frac{\rho}{\tau}e_i\right)^2, \end{aligned}$$

which proves (**a**). Because $\Pi_i = \left(1 - \left(\frac{\lambda\tau}{\rho+\lambda\tau}\right)^2\right)\Pi_i^{pt} + \left(\frac{\lambda\tau}{\rho+\lambda\tau}\right)^2\Pi_i^{nt}$,

$$\begin{aligned} G_i &\equiv \Pi_i - \Pi_i^{nt} \\ &= \left(1 - \left(\frac{\lambda\tau}{\rho+\lambda\tau}\right)^2\right)\left(\Pi_i^{pt} - \Pi_i^{nt}\right) \\ &= \left(1 - \left(\frac{\lambda\tau}{\rho+\lambda\tau}\right)^2\right)G_i^{pt}. \end{aligned}$$

Next, we show $\frac{\lambda\tau}{\rho+\lambda\tau} = \frac{\left(w - \frac{1-w}{n}\right)\varphi + \frac{1}{n}}{1-\varphi}$. Because $\lambda = \frac{1}{n\beta_p}$, $\frac{\lambda\tau}{\rho+\lambda\tau} = \frac{1}{\rho\beta_p\frac{n}{\tau}+1}$. First,

$$\begin{aligned} \rho\beta_p\frac{n}{\tau} &= \frac{\tau_1}{\tau}\frac{n-1-(n(1+w)-(1-w))\varphi}{1+(1-w)(nw-1)\varphi+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+(nw-(1-w))\varphi\}}\frac{1+(1-w)(nw-1)\varphi}{1+(nw-(1-w))\varphi} \\ &= \frac{n-1-(n(1+w)-(1-w))\varphi}{1+(nw-(1-w))\varphi}, \end{aligned}$$

where we used (43). Hence,

$$\begin{aligned} \rho\beta_p\frac{n}{\tau}+1 &= \frac{n-1-(n(1+w)-(1-w))\varphi+1+(nw-(1-w))\varphi}{1+(nw-(1-w))\varphi} \\ &= \frac{n(1-\varphi)}{1+(nw-(1-w))\varphi}. \end{aligned}$$

Therefore, $\frac{\lambda\tau}{\rho+\lambda\tau} = \frac{\left(w-\frac{1-w}{n}\right)\varphi+\frac{1}{n}}{1-\varphi}$. Finally, $\frac{\lambda\tau}{\rho+\lambda\tau} < 1 \Leftrightarrow \frac{1-\varphi}{\varphi} > w\frac{n+1}{n-1}$. Because $w\frac{n+1}{n-1} \leq w + \frac{2}{n-1} \Leftrightarrow w \leq 1$ for $n > 1$, we conclude $\frac{\lambda\tau}{\rho+\lambda\tau} < 1$ whenever the second-order condition (30) is satisfied. ■

To derive ex ante gains from trade, define

$$T(n, \varphi; w) \equiv \frac{wn\varphi}{1-(1-w)\varphi}.$$

Also define $A(T;t,w)$ and $B(T;u)$ by

$$A(T;t,w) \equiv \frac{1 - tc_s + (1 - w + (w-t)\, c_s)\, T\,(n,\varphi;w)}{1 + (1 - w + wc_s)\, T\,(n,\varphi;w)}, \tag{48}$$

$$B(T;u) \equiv \frac{u}{1 + (1-u)\, T\,(n,\varphi;w)}. \tag{49}$$

When $w = 1$, $T\,(n,\varphi;1) = n\varphi$, and (48) and (49) become

$$A(n\varphi;t,1) = \frac{1 - tc_s + (1-t)\, c_s n\varphi}{1 + c_s n\varphi},$$

$$B(n\varphi;u) = \frac{u}{1 + (1-u)\, n\varphi}.$$

Lemma A5 (ex ante gains from trade: G^{pt} **and** G**)**

(a) $G^{pt} = \frac{1}{2\rho} \log\left(1 + \frac{\rho^2}{\tau_v \tau_x} \frac{n}{1+n} \frac{1-\varphi}{1-(1-w)\varphi} A(T;t,w) B(T;u)\right)$.

(b) $G = \frac{1}{2\rho} \log\left(1 + \frac{\rho^2}{\tau_v \tau_x} \left(\frac{n-1}{n} - \frac{n+1}{n} \frac{w\varphi}{1-\varphi}\right) A(T;t,w) B(T;u)\right)$.

(c) $G < G^{pt}$ for all $n \geq 1$.

Proof. Write $E_i[v] = \sqrt{t} E_i[v_1] = \gamma_s s_i + \gamma_e e_i + \gamma_p p$, where

$$\gamma_s = \sqrt{t} k_s,\ \gamma_e = \sqrt{t} k_e,\ \gamma_p = \sqrt{t} k_p \text{ with } k_s,\ k_e,\ k_p \text{ in (32).}$$

Then

$$G_i^{pt} = [s_i, e_i, p]^\top C [s_i, e_i, p] \text{ and } G_i = \left(1 - \widetilde{\lambda}\right) G_i^{pt}, \tag{50}$$

where

$$C \equiv \frac{\tau}{2\rho} \left[\gamma_s, -\left(\frac{\rho}{\tau} - \gamma_e\right), -(1-\gamma_p)\right]^\top \left[\gamma_s, -\left(\frac{\rho}{\tau} - \gamma_e\right), -(1-\gamma_p)\right]$$

is a 3-by-3 matrix. We use the following fact:

Fact 1. *Given the n-dimensional random vector z that is normally distributed with mean zero and variance-covariance matrix Σ,*

$$E[-\exp(-\rho(z^\top C z))] = -\left[\det\left(I_n + 2\rho \Sigma C\right)\right]^{-\frac{1}{2}},$$

where I_n is the n-dimensional identity matrix.

Apply this to (50) to obtain

$$G^{pt} = \frac{1}{2\rho}\log\left[\det\left(I_3 + 2\rho\Sigma C\right)\right] \text{ and } G = \frac{1}{2\rho}\log\left[\det\left(I_3 + 2\rho\left(1-\widetilde{\lambda}\right)\Sigma C\right)\right],$$

where $\Sigma \equiv Var[[s_i, e_i, p]] = \begin{bmatrix} V_s & 0 & V_{sp} \\ & V_e & V_{ep} \\ & & V_p \end{bmatrix}$.

Next, we use the following fact:

Fact 2. *Given the n-by-1 column vector A and the 1-by-n row vector B,*

$$\det\left(I_n + AB\right) = 1 + BA.$$

We apply this fact to $\det\left(I_3 + 2\rho\Sigma C\right)$ and $\det\left(I_3 + 2\rho\left(1-\widetilde{\lambda}\right)\Sigma C\right)$, where

$$\Sigma C = \frac{\tau}{2\rho}\Sigma\left[\gamma_s, -\left(\frac{\rho}{\tau}-\gamma_e\right), -(1-\gamma_p)\right]^\top \times \left[\gamma_s, -\left(\frac{\rho}{\tau}-\gamma_e\right), -(1-\gamma_p)\right].$$

First,

$$\Sigma\left[\gamma_s, -\left(\frac{\rho}{\tau}-\gamma_e\right), -(1-\gamma_p)\right]^\top = \left[C_1, -C_2, C_3\right]^\top,$$

where

$$\begin{aligned} C_1 &\equiv \gamma_s V_s - (1-\gamma_p)V_{sp}, \\ C_2 &\equiv \left(\frac{\rho}{\tau}-\gamma_e\right)V_e + (1-\gamma_p)V_{ep}, \\ C_3 &\equiv \gamma_s V_{sp} - \left(\frac{\rho}{\tau}-\gamma_e\right)V_{ep} - (1-\gamma_p)V_p. \end{aligned}$$

Therefore,

$$\left[\gamma_s, -\left(\frac{\rho}{\tau}-\gamma_e\right), -(1-\gamma_p)\right] \times \left[C_1, -C_2, C_3\right]^\top = \gamma_s C_1 + \left(\frac{\rho}{\tau}-\gamma_e\right)C_2 - (1-\gamma_p)C_3.$$

Using $p^* = \frac{\beta_s}{\beta_p}\bar{s} - \frac{\beta_e}{\beta_p}\bar{e}$, it can be shown that

$$\begin{aligned} V_{sp} &= \frac{\gamma_s}{\gamma_p}\frac{\varphi}{1-\varphi}\left\{(1+c_s n)V_s + (1-w)\frac{n}{\tau_\varepsilon}\right\}, \\ V_{ep} &= -\frac{\gamma_e}{\gamma_p}V_e, \\ V_p &= (1+n)\left(\frac{\gamma_s}{\gamma_p}\frac{w\varphi}{1-\varphi}V_{sp} - \frac{\gamma_e}{\gamma_p}\frac{1}{N_u}V_{ep}\right), \end{aligned}$$

where $V_s = \frac{1}{c_s \tau_v}$ and $V_e = \frac{1}{\tau_x}$. Using these, algebra shows that

$$\begin{aligned}
C_1 &= w \frac{\gamma_s}{\tau_\varepsilon} \frac{n}{1+n}, \\
C_2 &= \frac{\rho}{\tau_x \tau} \left\{ 1 - \frac{N_u}{1+n} \frac{1+\varphi\left(wn-(1-w)\right)}{1+\varphi\left(wN_u - 1\right)} \right\}, \\
C_3 &= 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
\det\left(I_3 + 2\rho\Sigma C\right) &= 1 + \tau\left\{ \gamma_s C_1 + \left(\frac{\rho}{\tau} - \gamma_e\right) C_2 \right\}, \\
\det\left(I_3 + 2\rho\left(1-\widetilde{\lambda}\right)\Sigma C\right) &= 1 + \left(1-\widetilde{\lambda}\right)\tau\left\{ \gamma_s C_1 + \left(\frac{\rho}{\tau} - \gamma_e\right) C_2 \right\}.
\end{aligned}$$

This proves (**c**).

Next, calculate

$$\begin{aligned}
\tau\gamma_s C_1 &= \frac{\tau}{\tau_\varepsilon}\left(\frac{\sqrt{t}\tau_\varepsilon}{\tau}\frac{1-\varphi}{1+\varphi\left(1-w\right)\left(nw-1\right)}\right)^2 \frac{n}{n+1} w, \\
\tau\left(\frac{\rho}{\tau} - \gamma_e\right) C_2 &= \frac{\rho\left(1-\varphi\right)}{1+\varphi\left(wN_u-1\right)} \frac{\rho}{\tau_x\tau_v} \frac{\tau_v}{\tau} \frac{n\left(1-\varphi\right)}{1+n} \frac{u}{1+\varphi\left(wN_u-1\right)}.
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\tau\left\{ \gamma_s C_1 + \left(\frac{\rho}{\tau} - \gamma_e\right) C_2 \right\} &= (1-\varphi)^2 \frac{n}{1+n} \frac{\rho^2}{\tau_x\tau_v} \frac{\tau_v}{\tau} \\
&\quad \times \left\{ \frac{\tau_x\tau_\varepsilon}{\rho^2}\left(\frac{\sqrt{t}\tau}{\tau_1}\frac{1}{1+\varphi\left(1-w\right)\left(nw-1\right)}\right)^2 w + \frac{u}{\left\{1+\varphi\left(wN_u-1\right)\right\}^2} \right\}.
\end{aligned}$$

Using (45),

$$\tau\left\{ \gamma_s C_1 + \left(\frac{\rho}{\tau} - \gamma_e\right) C_2 \right\} = \left(\frac{1-\varphi}{1+\varphi\left(wN_u-1\right)}\right)^2 \frac{n}{1+n} \frac{\rho^2}{\tau_v\tau_x} \frac{\tau_v}{\tau} \left(\frac{w}{\alpha k^2} + u\right).$$

From (33), we have $\frac{1}{\alpha k^2} = \frac{\varphi u N_u}{1-\varphi}$ and

$$\begin{aligned}
\frac{w}{\alpha k^2} + u &= \frac{w\varphi u N_u + u\left(1-\varphi\right)}{1-\varphi} \\
&= \frac{u\left\{1+\varphi\left(wN_u-1\right)\right\}}{1-\varphi}.
\end{aligned}$$

Hence,

$$\tau\left\{ \gamma_s C_1 + \left(\frac{\rho}{\tau} - \gamma_e\right) C_2 \right\} = \frac{\rho^2}{\tau_v\tau_x} \frac{n(1-\varphi)}{1+n} \frac{\tau_v}{\tau} \frac{u}{1+\varphi\left(wN_u-1\right)}.$$

Using $wN_u - 1 = wn(1-u) - (1-w)$,

$$\begin{aligned}
\frac{u}{1+\varphi(wN_u-1)} &= \frac{u}{1-\varphi(1-w)+wn\varphi(1-u)} \\
&= \frac{1}{1-\varphi(1-w)}\frac{u}{1+\frac{wn\varphi}{1-\varphi(1-w)}(1-u)} \\
&= \frac{B(T;u)}{1-\varphi(1-w)}.
\end{aligned}$$

Also,

$$\begin{aligned}
\frac{\tau_v}{\tau} &= \tau_v\left((1-t)\frac{1}{\tau_v}+t\frac{1}{\tau_1}\right) \\
&= 1-t+t\frac{\tau_v}{\tau_v+\tau_v\frac{1+\varphi(nw-(1-w))}{1+\varphi(1-w)(nw-1)}} \\
&= \frac{1+\varphi(1-w)(nw-1)+(1-t)\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}{1+\varphi(1-w)(nw-1)+\frac{\tau_\varepsilon}{\tau_v}\{1+\varphi(nw-(1-w))\}}.
\end{aligned}$$

Using $1+(1-t)\frac{\tau_\varepsilon}{\tau_v} = \frac{1-tc_s}{1-c_s}$ and $1+\frac{\tau_\varepsilon}{\tau_v} = \frac{1}{1-c_s}$,

$$\begin{aligned}
\frac{\tau_v}{\tau} &= \frac{\frac{1-tc_s}{1-c_s}(1-(1-w)\varphi)+n\varphi w\left(\frac{1-tc_s}{1-c_s}-w\right)}{\frac{1}{1-c_s}(1-(1-w)\varphi)+n\varphi w\left(\frac{1}{1-c_s}-w\right)} \\
&= \frac{1-tc_s+\frac{wn\varphi}{1-(1-w)\varphi}(1-tc_s-w(1-c_s))}{1+\frac{wn\varphi}{1-(1-w)\varphi}(1-w(1-c_s))} \\
&= A(T;t,w).
\end{aligned}$$

Finally,

$$\begin{aligned}
1-\widetilde{\lambda} &= \left(1-\left(\frac{\left(w-\frac{1-w}{n}\right)\varphi+\frac{1}{n}}{1-\varphi}\right)^2\right) \\
&= \frac{1}{(1-\varphi)^2}\left(1-\varphi-\left(w-\frac{1-w}{n}\right)\varphi-\frac{1}{n}\right)\left(1-\varphi+\left(w-\frac{1-w}{n}\right)\varphi+\frac{1}{n}\right) \\
&= \frac{1}{(1-\varphi)^2}\left(\frac{n-1}{n}-\left(\frac{n-1}{n}+\frac{n+1}{n}w\right)\varphi\right)\left(\frac{n+1}{n}-(1-w)\frac{n+1}{n}\varphi\right) \\
&= \frac{1}{(1-\varphi)^2}\left\{\frac{n-1}{n}(1-\varphi)-\frac{n+1}{n}w\varphi\right\}(1-(1-w)\varphi)\frac{n+1}{n} \\
&= \frac{1-(1-w)\varphi}{1-\varphi}\frac{n+1}{n}\left\{\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right\}.
\end{aligned}$$

This proves (**a**) and (**b**). ■

7.4.4 Multiplicity due to self-fulfilling trading motives

Next, we study multiplicity. To characterize $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$ that solves the cubic equation (46) and associated φ, define $X(t,u,w)$ and $X_n(t,u,w)$ by

$$X(t,u,w) \equiv \frac{1 - w + c_s(w-t)}{(1-u)(1-tc_s)}, \tag{51}$$

$$X_n(t,u,w) \equiv \frac{\frac{1}{n} + \frac{1-w+(w-t)c_s}{1-tc_s}}{\frac{1}{n} + 1 - u}. \tag{52}$$

We suppress the notation (t,u,w) and write them X and X_n below. Note that only one of the three conditions below holds:

$$\begin{aligned}
\forall n &\geq 1,\ 1 < X_n < X \text{ and } X_n \text{ monotonically increases in } n \text{ and } \lim_{n\to\infty} X_n = X. \\
\forall n &\geq 1,\ 1 > X_n > X \text{ and } X_n \text{ monotonically decreases in } n \text{ and } \lim_{n\to\infty} X_n = X. \\
\forall n &\geq 1,\ 1 = X_n = X.
\end{aligned}$$

Using X, we classify a parameter space of (t,u) for a given w. For a fixed $w \in (0,1]$, define for each $m \in [0,\infty]$

$$\begin{aligned}
\chi_w(m) &\equiv \{(t,u) | X(t,u,w) = m\} \text{ for } m < \infty, \\
\chi_w(\infty) &\equiv \{(t,u) | t \in (0,1], u = 1\}.
\end{aligned}$$

Note that for $m > 0$,

$$X(t,u,w) = m \Leftrightarrow u = \frac{m-1}{m} + \frac{w}{m}\frac{1-c_s}{1-tc_s}. \tag{53}$$

Hence, $\chi_w(m)$ defines a positive relationship between t and u for given values of w and m. In particular, $\chi_w(1)$ has two end points $(t,u) = (0, w(1-c_s))$ and $(t,u) = (1,w)$, because $X(0,u,w) = 1 \Leftrightarrow u = w(1-c_s)$ and $X(1,u,w) = 1 \Leftrightarrow u = w$. Also note that for $w < 1$,

$$\chi_w(1-w) = \{(1,0)\} \text{ and } \chi_w(m) = \emptyset \text{ for } m < 1-w,$$

See **Figure A1** before **Lemma A7** for the illustration of $\chi_w(m)$ on a (t,u)-plane for different values of m.

Lemma A6 (Multiplicity)

(a) *Multiple solutions exist for* (46) *if and only if* $w > \frac{8}{9}$ *and*

$$\text{(i) } (t,u) \in \bigcup_{m\in\left[0,\frac{1}{9}\right)} \chi_w(m), \text{ (ii) } \frac{8}{1-9m} < n(1-u), \text{ and (iii) } \frac{\rho^2}{\tau_\varepsilon\tau_x} \in \left[\alpha^-_{n,m}, \alpha^+_{n,m}\right],$$

where $\alpha^\pm_{n,m}$ *are defined in the proof.*

(**b**) *If larger k does not satisfy* (30), *neither does smaller k.*

If multiple solutions satisfy (30), *an equilibrium with a larger k is Pareto-superior.*

Proof. We use the following fact:

Fact 3. *$a_3x^3 + a_2x^2 + a_1x + a_0 = 0$ has three distinct real solutions if and only if*

$$a_3 \neq 0 \text{ and } 0 < \Delta \equiv -4a_1^3a_3 + (a_1a_2)^2 - 4a_0a_2^3 + 18a_0a_1a_2a_3 - 27(a_0a_3)^2.$$

Recall $N_u \equiv 1 + (1-u)n$ and $\alpha \equiv \frac{\rho^2}{\tau_\varepsilon\tau_x}$. Apply the fact for (46) to obtain

$$\begin{aligned}
\Delta \; = \; & \alpha u N_u^2 \times \left[tw^2 \left\{ \alpha u \left(\frac{1-tc_s}{1-c_s} \right)^2 - 4tw \right\} N_u^2 \right. \\
& -2\alpha u w \frac{1-tc_s}{1-c_s} \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\} \left\{ 2\alpha u \left(\frac{1-tc_s}{1-c_s} \right)^2 - 9tw \right\} N_u \\
& \left. -27\alpha t u w^2 \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\}^2 \right].
\end{aligned} \tag{54}$$

Write terms in a square bracket as a quadratic function of α:

$$\begin{aligned}
\widetilde{\Delta}(\alpha) \; \equiv \; & -4t^2w^3N_u^2 - 4u^2w \left(\frac{1-tc_s}{1-c_s} \right)^3 \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\} N_u \alpha^2 \\
& +tuw^2 \left\{ \begin{array}{c} \left(\frac{1-tc_s}{1-c_s} \right)^2 N_u^2 + 18\frac{1-tc_s}{1-c_s} \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\} N_u \\ -27 \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\}^2 \end{array} \right\} \alpha.
\end{aligned}$$

$\widetilde{\Delta}(\alpha) = 0$ has two real solutions if and only if

$$\begin{aligned}
0 \; < \; & \left[\left(\frac{1-tc_s}{1-c_s} \right)^2 N_u^2 + 18\frac{1-tc_s}{1-c_s} \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\} N_u \right. \\
& \left. -27 \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\}^2 \right]^2 \\
& -64 \left(\frac{1-tc_s}{1-c_s} \right)^3 \left\{ \frac{1-tc_s}{1-c_s} + n\left(1 - w + \frac{c_s(1-t)}{1-c_s}\right) \right\} N_u^3.
\end{aligned} \tag{55}$$

Otherwise, $\widetilde{\Delta}(\alpha) \leq 0$ for all α.

Note that

$$\begin{aligned}\frac{\frac{1-tc_s}{1-c_s}+n\left(1-w+\frac{c_s(1-t)}{1-c_s}\right)}{\frac{1-tc_s}{1-c_s}N_u} &= \frac{1+\frac{(1-w)(1-c_s)+c_s(1-t)}{1-tc_s}n}{1+(1-u)n} \\ &= \frac{\frac{1}{n}+\frac{1-w+(w-t)c_s}{1-tc_s}}{\frac{1}{n}+1-u} \\ &= X_n.\end{aligned}$$

Thus, divide (55) by $\left(\frac{1-tc_s}{1-c_s}N_u\right)^4$ to obtain:

$$\begin{aligned}H(X_n) &\equiv \left(1+18X_n-27X_n^2\right)^2-64X_n \\ &= \left(1+18X_n-27X_n^2+8\sqrt{X_n}\right)\left(1+18X_n-27X_n^2-8\sqrt{X_n}\right) \\ &= \left(1+18X_n-27X_n^2+8\sqrt{X_n}\right)\left(\frac{1}{3}-\sqrt{X_n}\right)\left\{\left(3\sqrt{X_n}+1\right)(9X_n-5)+8\right\}.\end{aligned} \tag{56}$$

A necessary condition for (46) to have three solutions is $H(X_n)>0$. Given $H(X_n)>0$, two real solutions of $\widetilde{\Delta}(\alpha)=0$ are

$$\alpha_{n,m}^{\pm}=\frac{(1-m(1-u))(1-w+wc_s-m(1-u))}{c_s uw}\frac{1+18X_n-27X_n^2\pm\sqrt{H(X_n)}}{8X_n}, \tag{57}$$

where $t=\frac{1}{c_s}\left\{1-\frac{w(1-c_s)}{1-m(1-u)}\right\}$ and $\frac{1-tc_s}{1-c_s}=\frac{w}{1-m(1-u)}$ were used to eliminate t. Thus, if $\alpha\in(\alpha_{n,m}^-,\alpha_{n,m}^+)$, then $\Delta>0$ and (46) has three solutions. $\alpha\in\{\alpha_{n,m}^-,\alpha_{n,m}^+\}$ is a knife-edge case where (46) has two real solutions. If $\alpha\notin\left[\alpha_{n,m}^-,\alpha_{n,m}^+\right]$, then $\Delta<0$ and (46) has a unique solution. This establishes the necessity and sufficiency of (iii).

To see why (ii) is necessary, recall $\alpha\equiv\frac{\rho^2}{\tau_\varepsilon\tau_x}>0$. Hence, at least one of $\alpha_{n,m}^{\pm}$ must be positive to have $\alpha\in\left[\alpha_{n,m}^-,\alpha_{n,m}^+\right]$. This requires $1+18X_n-27X_n^2>0\Leftrightarrow X_n\in\left[0,\frac{1}{3}\left(1+\sqrt{\frac{4}{3}}\right)\right)$. In this range, from (56), $H(X_n)>0$ if and only if $X_n<\frac{1}{9}$.

To see why (i) is necessary, note that if $\frac{1}{9}\le X$, then $X_n>X\ge\frac{1}{9}$ for all n. Therefore $H(X_n)>0$ cannot be satisfied. If $X=m<\frac{1}{9}$, then X_n approaches m from above as n increases, and for sufficiently large n, $X_n<\frac{1}{9}$. This establishes the necessity of (i).

Note that

$$\begin{aligned}
X_n &< \frac{1}{9} \\
&\Leftrightarrow 1+\frac{(1-w)(1-c_s)+c_s(1-t)}{1-tc_s}n < \frac{1}{9}(1+(1-u)n) \\
&\Leftrightarrow \left\{\frac{1-u}{9}-\frac{(1-w)(1-c_s)+c_s(1-t)}{1-tc_s}\right\}n > \frac{8}{9} \\
&\Leftrightarrow \{1-u-9(1-u)m\}\,n > 8 \\
&\Leftrightarrow (1-u)(1-9m)n > 8 \\
&\Leftrightarrow n > \frac{8}{(1-u)(1-9m)}.
\end{aligned}$$

We used $X = m \Leftrightarrow m = \frac{(1-w)(1-c_s)+c_s(1-t)}{1-u}$, and $m < X_n < \frac{1}{9}$ for the last step. This establishes the necessity of (ii).

Finally, the condition $w > \frac{8}{9}$ is necessary because if $w \leq \frac{8}{9}$, then $\underset{m\in\left[0,\frac{1}{9}\right)}{\cup}\chi_w(m)$ is empty. From the above discussion, the sufficiency of the joint conditions $w > \frac{8}{9}$ and (i) through (iii) is obvious. This proves (**a**).

From (33), the larger solution k implies lower φ for a fixed parameters n, t, u, w. Therefore, if the larger solution does not satisfy (30), neither does the smaller solution. From (48) and (49), it is easy to verify $A(T;t,w)$ and $B(T;u)$ are decreasing in T. From $T(n,\varphi;w) \equiv \frac{wn\varphi}{1-(1-w)\varphi}$, T is increasing in φ for a fixed n and w. Finally, from the expression of G^{pt} and G in **Lemma A5**, it is clear that lower φ raises G^{pt} and G for a fixed parameters n, t, u, w. This proves (**b**). ■

7.4.5 Information aggregation with $w < 1$

Next, we characterize information aggregation with $w < 1$. A case with $w = 1$ needs a separate treatment because extreme negative externalities may arise. Recall that $\chi_w(m) \equiv \{(t,u)|X(t,u,w)=m\}$ for $m < \infty$, where

$$X(t,u,w) \equiv \frac{1-w+c_s(w-t)}{(1-u)(1-tc_s)},$$

and $\chi_w(\infty) \equiv \{(t,u)|t\in(0,1], u=1\}$. We partition the parameter space $(0,1]^2$ of (t,u) into the following four groups.

(i) $(t,u) \in \chi_w(\infty) \equiv \{(t,u)|t\in(0,1], u=1\}$ (assets without x_0).

(ii) $(t,u) \in \underset{m\in(1,\infty)}{\cup}\chi_w(m)$.

(iii) $(t,u) \in \chi_w(1) = \{(t,u)|X(t,u,w)=1\}$.

(iv) $(t,u) \in \underset{m\in(1-w,1)}{\cup}\chi_w(m)$.

Figure A1 illustrates this classification.

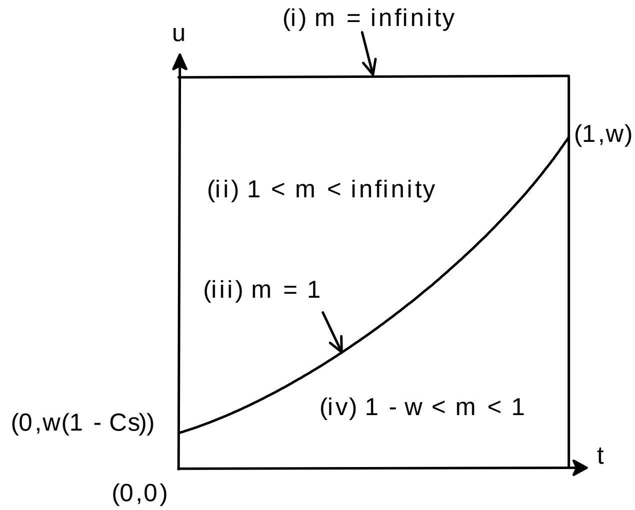


Figure A1. Classification of (t, u) for a fixed $w \in (0, 1)$.

Lemma A7 (information aggregation with $w \in (0, 1)$**)**

(a) *For* (i), *$n\varphi$ and k increases in n at the rate $n^{\frac{1}{3}}$.*

(b) *For* (ii), *k increases in n with an upper bound $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$.*

φ decreases in n at the rate $(k^2 n)^{-1}$.

(c) *For* (iii), *$k = \frac{1-tc_s}{\sqrt{t}(1-c_s)}$. φ decreases in n at the rate n^{-1}.*

(d) *For* (iv), *the largest solution k decreases in n with a lower bound $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$.*

φ decreases in n at the rate $(k^2 n)^{-1}$.

(e) *$n\varphi$ is increasing in n for all groups* (i) *to* (iv)*.*

It converges to a finite value for groups (ii)(iii)(iv)*.*

Proof. We proceed in three steps:
1) characterize a solution k^* to the cubic equation (46), where the largest solution is selected if there are multiple solutions,
2) characterize φ,
3) characterize $n\varphi$.

[Step 1] Characterize k^*.
Because (46) is linear in n, it can be written as

$$F(k) = \frac{\partial F}{\partial n} n + \left(\alpha u k^2 + w\right) \left(\sqrt{t} k - \frac{1 - tc_s}{1 - c_s}\right), \tag{58}$$

where $\frac{\partial F}{\partial n} =$

$$\begin{cases} -w\left(\frac{1-tc_s}{1-c_s} - w\right) \text{ if } u = 1, \\ (1-u)\left\{(\alpha u k^2 + w)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + w\left(\frac{1-tc_s}{1-c_s} - \frac{\frac{1-tc_s}{1-c_s} - w}{1-u}\right)\right\} \text{ otherwise.} \end{cases}$$

First, we show that k^* is increasing in n if $1 < m$ ((i) and (ii)). Note that

$$1 < m \Leftrightarrow \frac{1-tc_s}{1-c_s} < \frac{\frac{1-w+c_s(w-t)}{1-c_s}}{1-u} = \frac{\frac{1-tc_s}{1-c_s} - w}{1-u}.$$

From (46), the solution k^* must satisfy $\sqrt{t}k^* \in \left(\frac{1-tc_s}{1-c_s}, \frac{\frac{1-tc_s}{1-c_s}-w}{1-u}\right)$ for $m \in (1,\infty)$ and $\sqrt{t}k^* > \frac{1-tc_s}{1-c_s}$ for $m = \infty$. Let $\left.\frac{\partial F}{\partial n}\right|_{k^*}$ denote $\frac{\partial F}{\partial n}$ evaluated at k^*. From (58), $\left.\frac{\partial F}{\partial n}\right|_{k^*} < 0$ because $F(k^*) = 0$ and the second term is positive. By **Lemma A6** the solution k^* is unique for $m > 1$, and therefore $F'(k^*) > 0$. By the implicit function theorem, k^* is increasing in n. The upper bound for k^* for a fixed $m \in (1,\infty)$ is $\frac{\frac{1-tc_s}{1-c_s}-w}{\sqrt{t}(1-u)} = \frac{1-tc_s-w(1-c_s)}{\sqrt{t}(1-u)(1-c_s)} = \frac{m}{\sqrt{t}}\frac{1-tc_s}{1-c_s}$.

Next, consider (iii), where

$$m = 1 \Leftrightarrow \frac{1-tc_s}{1-c_s}(1-u) = \frac{1-tc_s}{1-c_s} - w.$$

From (46), the solution k^* must satisfy $\sqrt{t}k^* = \frac{1-tc_s}{1-c_s} = \frac{\frac{1-tc_s}{1-c_s}-w}{1-u}$. Thus, k^* is independent of n.

Finally, consider (iv), where

$$m < 1 \Leftrightarrow \frac{1-tc_s}{1-c_s}(1-u) > \frac{1-tc_s}{1-c_s} - w.$$

From (46), the solution k^* must satisfy $\sqrt{t}k^* \in \left(\frac{\frac{1-tc_s}{1-c_s}-w}{1-u}, \frac{1-tc_s}{1-c_s}\right)$. If the solution k^* is unique, it satisfies $F'(k^*) > 0$ and k^* is decreasing in n by the implicit function theorem. If there are three solutions, the largest solution is chosen and it also satisfies $F'(k^*) > 0$. In the knife-edge case where there are two solutions, the larger solution satisfies either $F'(k^*) > 0$ or $F'(k^*) = 0$. In the latter case, because $\left.\frac{\partial F}{\partial n}\right|_{k^*} > 0$ from (58), the solution disappears with the small increase in n and the selected solution discontinuously drops to the lower value. Hence, the selected k^* is decreasing in n. The lower bound for k^* for a fixed $m \in (1-w, 1)$ is $\frac{\frac{1-tc_s}{1-c_s}-w}{\sqrt{t}(1-u)} = \frac{1-tc_s-w(1-c_s)}{\sqrt{t}(1-u)(1-c_s)} = \frac{m}{\sqrt{t}}\frac{1-tc_s}{1-c_s}$.

[Step 2] Study behavior of φ.

For (ii), (iii), (iv), the results are obvious from (9) and the results in Step 1. This proves (**b**), (**c**), (**d**). For (i), the unique k^* solves

$$F(k; u=1) = (\alpha k^2 + w)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) - wn\left(\frac{1-tc_s}{1-c_s} - w\right) = 0.$$

Therefore, $\sqrt{t}k^* > \frac{1-tc_s}{1-c_s}$ and k^* increases in n without a bound at the rate $n^{\frac{1}{3}}$. From (33), φ decreases in n at the rate $n^{-\frac{2}{3}}$. This proves (**a**).

[Step 3] Study behavior of $n\varphi$.
From (33),

$$\frac{1}{n\varphi} = \frac{1}{n}\left(1 + \alpha u N_u k^2\right) = \frac{1}{n}\left(1 + \alpha u k^2\right) + \alpha u\left(1-u\right)k^2. \tag{59}$$

For (iii) and (iv), k^* is not increasing in n and hence $n\varphi$ is increasing in n.
For (i) and (ii), $F(k) = 0$ implies

$$\frac{1}{nw}\left(w + \alpha u N_u k^2\right) = \frac{\frac{1-tc_s}{1-c_s} - w - (1-u)\sqrt{t}k}{\sqrt{t}k - \frac{1-tc_s}{1-c_s}}.$$

Because k^* is increasing in n for (i) and (ii),

$$\begin{aligned}\frac{1}{n}\left(1 + \alpha u N_u k^2\right) &= \frac{1}{n}\left(w + \alpha u N_u k^2\right) + \frac{1-w}{n} \\ &= \frac{1}{w}\frac{\frac{1-tc_s}{1-c_s} - w - (1-u)\sqrt{t}k}{\sqrt{t}k - \frac{1-tc_s}{1-c_s}} + \frac{1-w}{n}\end{aligned}$$

is decreasing in n. Hence $n\varphi$ is increasing in n. For (ii)(iii)(iv), k has a finite positive limit as $n \to \infty$. Then (59) shows that $n\varphi$ has a limit $\frac{1}{\alpha u(1-u)k^2}$. This proves (**e**). ■

Information aggregation with $w = 1$.

Next, we study information aggregation with $w = 1$. In this case, (46) simplifies to:

$$F\left(k; w=1\right) = \left(\alpha u N_u k^2 + 1\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + \ n\left\{(1-u)\sqrt{t}k - (1-t)\frac{\tau_\varepsilon}{\tau_v}\right\} = 0. \tag{60}$$

For $m \in [0, \infty]$, we have

$$\chi_1\left(m\right) \equiv \left\{(t,u) | X(t,u,1) = m\right\}, \text{ where } X(t,u,1) = \frac{1-t}{1-u}\frac{c_s}{1-tc_s}, \tag{61}$$

$\chi_1\left(\infty\right) \equiv \{(t,u) | t \in (0,1), u = 1\}$ and $\chi_1\left(0\right) \equiv \{(t,u) | t = 1, u \in (0,1)\}$. We treat $(t,u) = (1,1)$ separately. Thus, we partition the parameter space $(0,1]^2$ of (t,u) into the following six groups.

(i) $(t,u) \in \chi_1\left(\infty\right) \equiv \{(t,u) | u = 1, t \in (0,1)\}$ (assets without x_0).

(ii) $(t,u) \in \underset{m\in(1,\infty)}{\cup} \chi_1\left(m\right)$.

(iii) $(t,u) \in \chi_1\left(1\right) = \{(t,u) | X(t,u,1) = 1\}$.

(iv) $(t,u) \in \underset{m\in(0,1)}{\cup} \chi(m)$.

(v) $(t,u) \in \chi_1(0) \equiv \{(t,u)|t=1, u \in (0,1)\}$ (assets without v_0).

(vi) $(t,u) = (1,1)$.

Figure A2 illustrates this classification.

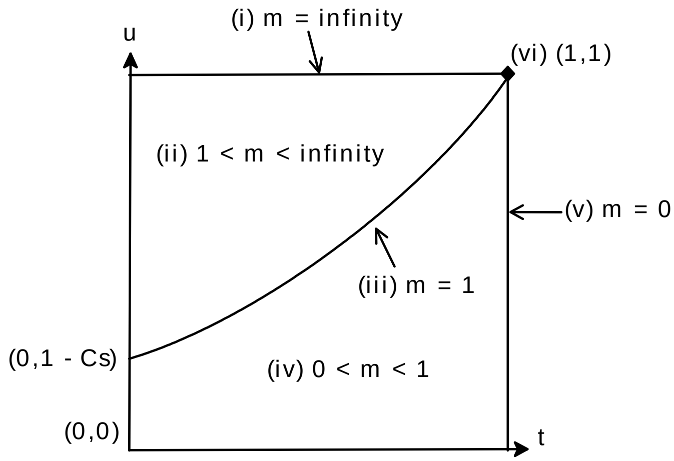


Figure A2. Classification of (t,u) with $w=1$.

Lemma A8 (information aggregation with $w=1$)

(a) *For* (i), *$n\varphi$ and k increases in n at the rate $n^{\frac{1}{3}}$.*

(b) *For* (ii), *k increases in n with an upper bound $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$.*

φ decreases in n at the rate $(k^2n)^{-1}$.

(c) *For* (iii), *$k = \frac{1-tc_s}{\sqrt{t}(1-c_s)}$. φ decreases in n at the rate n^{-1}.*

(d) *For* (iv), *the largest solution k decreases in n with a lower bound $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$.*

φ decreases in n at the rate $(k^2n)^{-1}$.

(e) *For* (v), *the largest solution k decreases in n. If $\frac{\rho^2}{\tau_\varepsilon\tau_x} \geq \frac{4}{u}$, k has a positive lower bound and φ decreases in n at the rate $(k^2n)^{-1}$. If $\frac{\rho^2}{\tau_\varepsilon\tau_x} < \frac{4}{u}$, k converges to zero at the rate n^{-1} and φ converges to one.*

(f) *For* (vi), *there is a unique solution $k=1$ and $\varphi = \left(1+\frac{\rho^2}{\tau_\varepsilon\tau_x}\right)^{-1}$.*

(g) *$n\varphi$ is increasing in n for all groups* (i) *to* (vi).

It converges to a finite value for groups (ii)(iii)(iv) *and* (v) *with $\frac{\rho^2}{\tau_\varepsilon\tau_x} \geq \frac{4}{u}$.*

Proof. For (vi), it is immediate that $k = 1$ is the unique solution to (60) and $\varphi = \left(1 + \frac{\rho^2}{\tau_\varepsilon \tau_x}\right)^{-1}$ from (33). Hence $n\varphi$ is increasing in n and unbounded. This proves (**f**).

For the other cases, we proceed in three steps:

1) characterize a solution k^* to the cubic equation (60), where the largest solution is selected if there are multiple solutions,
2) characterize φ,
3) characterize $n\varphi$.

[Step 1] Characterize k^*.

Because (60) is linear in n, it can be written as

$$F(k) = \frac{\partial F}{\partial n} n + \left(\alpha u k^2 + 1\right) \left(\sqrt{t}k - \frac{1 - tc_s}{1 - c_s}\right), \tag{62}$$

where $\frac{\partial F}{\partial n} =$

$$\begin{cases} -(1-t)\frac{\tau_\varepsilon}{\tau_v} \text{ if } u = 1, \\ (1-u)\left\{\left(\alpha u k^2 + 1\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + \frac{1-tc_s}{1-c_s} - \frac{1-t}{1-u}\frac{\tau_\varepsilon}{\tau_v}\right\} \text{ otherwise.} \end{cases}$$

First, we show that k^* is increasing in n if $1 < m$ ((i) and (ii)). Note that $1 < m \Leftrightarrow \frac{1-tc_s}{1-c_s}(1-u) < (1-t)\frac{\tau_\varepsilon}{\tau_v}$. From (60), the solution k^* must satisfy

$$\begin{aligned} \sqrt{t}k^* &\in \left(\frac{1 - tc_s}{1 - c_s}, \frac{(1-t)\frac{\tau_\varepsilon}{\tau_v}}{1 - u}\right) \text{ for } m \in (1, \infty), \\ \sqrt{t}k^* &> \frac{1 - tc_s}{1 - c_s} \quad \text{for } m = \infty. \end{aligned}$$

Let $\left.\frac{\partial F}{\partial n}\right|_{k^*}$ denote $\frac{\partial F}{\partial n}$ evaluated at k^*. From (62), $\left.\frac{\partial F}{\partial n}\right|_{k^*} < 0$ because $F(k^*) = 0$ and the second term is positive. By **Lemma A6** the solution k^* is unique for $m > 1$, and therefore $F'(k^*) > 0$. By the implicit function theorem, k^* is increasing in n. The upper bound for k^* for a fixed $m \in (1, \infty)$ is $\frac{(1-t)\frac{\tau_\varepsilon}{\tau_v}}{\sqrt{t}(1-u)} = \frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$.

Next, consider (iii), where

$$m = 1 \Leftrightarrow \frac{1 - tc_s}{1 - c_s}(1 - u) = (1 - t)\frac{\tau_\varepsilon}{\tau_v}.$$

From (60), the solution k^* must satisfy $\sqrt{t}k^* = \frac{(1-t)\frac{\tau_\varepsilon}{\tau_v}}{1-u} = \frac{1-tc_s}{1-c_s}$. Thus, k^* is independent of n.

Finally, consider (iv) and (v), where

$$m < 1 \Leftrightarrow \frac{1 - tc_s}{1 - c_s}(1 - u) > (1 - t)\frac{\tau_\varepsilon}{\tau_v}.$$

From (60), the solution k^* must satisfy $\sqrt{t}k^* \in \left(\frac{(1-t)\frac{\tau_\varepsilon}{\tau_v}}{1-u}, \frac{1-tc_s}{1-c_s}\right)$. From (62), $\left.\frac{\partial F}{\partial n}\right|_{k^*} > 0$. If the solution k^* is unique, it satisfies $F'(k^*) > 0$ and k^* is decreasing in n by the implicit function theorem. If there are three solutions, the largest solution is chosen and it also satisfies $F'(k^*) > 0$. In the knife-edge case where there are two solutions, the larger solution satisfies either $F'(k^*) > 0$ or $F'(k^*) = 0$. In the latter case, the solution disappears with the small increase in n and the selected solution discontinuously drops to the lower value. Hence, the selected k^* is decreasing in n. The lower bound for k^* for a fixed $m < 1$ is $\frac{1}{\sqrt{t}}\frac{(1-t)\frac{\tau_\varepsilon}{\tau_v}}{1-u} = \frac{m}{\sqrt{t}}\frac{1-tc_s}{1-c_s}$.

[Step 2] Study behavior of φ.

For (ii), (iii), (iv), the results are obvious from (33) and the results in Step 1. This proves (**b**), (**c**), (**d**).

For (i), the unique k^* solves

$$F(k; u = 1) = \left(\alpha k^2 + 1\right)\left(\sqrt{t}k - \frac{1 - tc_s}{1 - c_s}\right) - \frac{1 - tc_s}{1 - c_s}n = 0.$$

Therefore, $\sqrt{t}k^* > 1 + (1-t)\frac{\tau_\varepsilon}{\tau_v}$ and k^* increases in n without a bound at the rate $n^{\frac{1}{3}}$. From (33), φ decreases in n at the rate $n^{-\frac{2}{3}}$. This proves (**a**).

For (v), there may be multiple $k^* \in (0, 1)$ that solve

$$F(k; t = 1) = \left(\alpha uk^2 + 1\right)(k - 1) + (1 - u)\left\{\left(\alpha uk^2 + 1\right)(k - 1) + 1\right\} n = 0. \tag{63}$$

Note that $(\alpha uk^2 + 1)(k - 1) + 1 = k(\alpha uk^2 - \alpha uk + 1)$. Firs, $\lim_{n\to\infty} k = 1$ is impossible, because the second term in (63) would go to infinity. Therefore, as n increases, k^* must be approaching one of the solutions to $k(\alpha uk^2 - \alpha uk + 1) = 0$. To consider multiplicity in the limit, note that $\alpha^-_{n,0}$ approaches $\frac{4}{u}$ from below as $n \to \infty$ while $\lim_{n\to\infty} \alpha^+_{n,0} = \infty$. This is so because

$$\alpha^\pm_{n,0} = \frac{1}{8u}\left\{N_u + 18 - 27N_u^{-1} \pm \sqrt{(N_u + 18 - 27N_u^{-1})^2 - 64N_u}\right\},$$

and

$$\begin{aligned}
& N_u + 18 - 27N_u^{-1} - \sqrt{(N_u + 18 - 27N_u^{-1})^2 - 64N_u} \\
= \; & N_u\left\{1 + 18N_u^{-1} - 27N_u^{-2} - \sqrt{(1 + 18N_u^{-1} - 27N_u^{-2})^2 - 64N_u^{-1}}\right\} \\
= \; & N_u\left\{1 + 18N_u^{-1} - 27N_u^{-2} - \sqrt{(1 + 18N_u^{-1} - 27N_u^{-2})^2 - 64N_u^{-1}}\right\} \\
& \times \frac{1 + 18N_u^{-1} - 27N_u^{-2} + \sqrt{(1 + 18N_u^{-1} - 27N_u^{-2})^2 - 64N_u^{-1}}}{1 + 18N_u^{-1} - 27N_u^{-2} + \sqrt{(1 + 18N_u^{-1} - 27N_u^{-2})^2 - 64N_u^{-1}}} \\
= \; & \frac{64}{1 + 18N_u^{-1} - 27N_u^{-2} + \sqrt{(1 + 18N_u^{-1} - 27N_u^{-2})^2 - 64N_u^{-1}}}.
\end{aligned}$$

Here, $X_n = \frac{\frac{1}{n}}{\frac{1}{n}+1-u} = N_u^{-1}$ and $X_n < \frac{1}{9} \Leftrightarrow N_u > 9$. The denominator of the expression above is $\frac{8}{3}$ when $N_u = 9$ (hence $\alpha_{n,0}^{\pm} = \frac{3}{u}$) and decreases to 2 as $n \to \infty$. Hence, $\alpha_{n,0}^{-}$ approaches $\frac{4}{u}$ from below as $n \to \infty$.

If $(\alpha u)^2 - 4\alpha u \geq 0 \Leftrightarrow \alpha \geq \frac{4}{u}$, then for sufficiently large n, $\alpha_{n,0}^{-} < \frac{4}{u} \leq \alpha < \alpha_{n,0}^{+}$. Therefore, there are three solutions to (63) and by the selection of the largest solution, k^* approaches $\frac{1}{2}\left(1+\sqrt{1-\frac{4}{\alpha u}}\right) \in \left[\frac{1}{2}, 1\right)$.[33] If $(\alpha u)^2 - 4\alpha u < 0 \Leftrightarrow \alpha < \frac{4}{u}$, then $\alpha u k^2 - \alpha u k + 1 > 0$ for all k. For sufficiently large n, $\alpha < \alpha_{n,0}^{-} < \frac{4}{u}$. Thus, the unique k^* must be approaching zero at the rate $\frac{1}{n}$. From (33), φ converges to one. This proves (**e**).

[Step 3] Study behavior of $n\varphi$.

Recall $N_u \equiv 1 + (1-u)n$. From (33),

$$\frac{1}{n\varphi} = \frac{1}{n}\left(1 + \alpha u N_u k^2\right) = \frac{1}{n}\left(1 + \alpha u k^2\right) + \alpha u (1-u) k^2. \tag{64}$$

For (iii), (iv), (v), k^* is not increasing in n and hence $n\varphi$ is increasing in n.

For (i) and (ii), $F(k) = 0$ implies

$$\frac{1}{n}\left(1 + \alpha u N_u k^2\right) = \frac{(1-t)\frac{\tau_\varepsilon}{\tau_v} - (1-u)\sqrt{t}k}{\sqrt{t}k - \left(1 + (1-t)\frac{\tau_\varepsilon}{\tau_v}\right)}.$$

This is decreasing in n because k^* is increasing in n for (i) and (ii). Hence $n\varphi$ is increasing in n. For all cases except (i), (v) with $\frac{\rho^2}{\tau_\varepsilon \tau_x} < \frac{4}{u}$, and (vi), $(1-u)k^2$ has a finite limit bounded away from zero as $n \to \infty$. From (64) $n\varphi$ has a limit $\frac{1}{\alpha u (1-u) k^2}$. This proves (**g**). ■

Remark on information aggregation.

For both $w < 1$ and $w = 1$, the value of m determines how $k = \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$ depends on n. More precisely, **Lemma A7** and **A8** show that

$$m \gtreqless 1 \Leftrightarrow k \text{ is } \begin{cases} \text{increasing in} \\ \text{independent of} \\ \text{decreasing in} \end{cases} n.$$

In the parameter region with $m = 1$, $(t, u) \in \chi_w(1) = \{(t,u) | X(t,u,w) = 1\}$. Note that $\chi_w(1)$ defines a positive relationship between t and u such that the relative importance of two trading motives does not depend on n. If we increase u, the iid component of endowments increases. Thus, with (t,u) above $\chi_w(1)$ (i.e., $m > 1$) the risk-sharing motive becomes dominant as n increases. Contrary, with (t,u) below $\chi_w(1)$ (i.e., $m < 1$), the speculation motive becomes dominant as n increases. In terms of information aggregation, the latter case leads to faster information aggregation than the former.

The difference between $w < 1$ and $w = 1$ lies in the limiting behavior of φ: φ can be non-decreasing in n only if $w = 1$. From (33), φ is not decreasing in n if and only if

[33] If $\alpha = \frac{4}{u}$, the largest k^* approaches a double solution $\frac{1}{2}$ in the limit.

$\{1+(1-u)\,n\}\,k^2$ is not increasing in n. There are two cases where this can happen. First, that $u=1$ and k is independent of n occurs if and only if $w=t=1$, because with $w<1$, $\chi_w(1)$ lies below $u=1$ for all t and $(t,1)\in\chi_1(1)$ implies $t=1$. Intuitively, because $w<1$ reduces the incentive for speculation, the incentive for risk-sharing must be reduced by lowering u to counteract keep k independent of n for a given t. Second, that $u<1$ and nk^2 does not increase in n occurs if and only if $w=t=1$. For this to happen, k must be approaching zero. For $m<1$, the lower bound on k is $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}$. Recall that $m>1-w$, because $\chi_w(1-w)=\{(1,0)\}$. Therefore, when $w<1$, $\frac{m(1-tc_s)}{\sqrt{t}(1-c_s)}>\frac{(1-w)(1-tc_s)}{\sqrt{t}(1-c_s)}>0$ and the lower bound on k is bounded away from zero. Intuitively, with correlation in signal noise, traders always discount the signal and the speculation motive cannot be infinitely more dominant relative to the risk-sharing motive as $n\to\infty$.

Lemma A9 (extreme illiquidity with $u<t=w=1$)

Assume $u<\min\left\{1,\frac{4\tau_\varepsilon\tau_x}{\rho^2}\right\}$. *If there is* $\widetilde{n}\in(1,\infty)$ *such that* $\frac{\widetilde{n}+1}{\widetilde{n}-1}=\frac{1-\varphi}{\varphi}$*,*

then for all n sufficiently close to but smaller than $\widetilde{n}$, (30) *is satisfied.*

Also, $\lim_{n\nearrow\widetilde{n}}\beta_s=\lim_{n\nearrow\widetilde{n}}\beta_e=\lim_{n\nearrow\widetilde{n}}\beta_p=0$ *and* $\lim_{n\nearrow\widetilde{n}}\frac{1}{n\beta_p}=\infty$.

Proof.

From **Lemma A8**, φ is increasing in n for this case. Thus, if there is $\widetilde{n}\in(1,\infty)$ which solves $\frac{\widetilde{n}+1}{\widetilde{n}-1}=\frac{1-\varphi}{\varphi}$, then for all $n\in(1,\widetilde{n})$, $\frac{n+1}{n-1}<\frac{1-\varphi}{\varphi}$ and (30) is satisfied. From **Lemma A3**, β_s, β_e, β_p are $\frac{n-1}{n}-2\varphi$ times some positive and bounded constant. Because $\frac{\widetilde{n}+1}{\widetilde{n}-1}=\frac{1-\varphi}{\varphi}\Leftrightarrow\frac{\widetilde{n}-1}{\widetilde{n}}=2\varphi$, $\lim_{n\nearrow\widetilde{n}}\beta_s=\lim_{n\nearrow\widetilde{n}}\beta_e=\lim_{n\nearrow\widetilde{n}}\beta_p=0$. From **Lemma A3**,

$$
\begin{aligned}
n\beta_p &= \frac{n\left(\frac{n-1}{n}-2\varphi\right)}{1+(1-t)\frac{\tau_\varepsilon}{\tau_v}(1+n\varphi)}\frac{\tau_1}{\rho(1+n\varphi)} \\
&= \frac{n\left(\frac{n-1}{n}-2\varphi\right)(1-c_s)}{1-tc_s+(1-t)\,c_s n\varphi}\frac{\tau_1}{\rho(1+n\varphi)}.
\end{aligned}
$$

This converges to zero as $n\nearrow\widetilde{n}$. Therefore $\lim_{n\nearrow\widetilde{n}}\frac{1}{n\beta_p}=\infty$. ■

Remark on strategic foundation for equilibrium multiplicity.

Our analysis provides a strategic foundation for equilibrium multiplicity that the literature identified in a competitive environment with a continuum of traders (Ganguli and Yang 2009 and Manzano and Vives 2011). We also provide an equilibrium selection criterion based on ex ante welfare.

Ganguli and Yang (2009) assume $t=w=1$ and $u<1$. This corresponds to $m=0$ in **Figure A2**. In this case, with a finite number of traders, the extreme illiquidity is possible. So to give a strategic foundation (i.e., multiplicity with $n\to\infty$), the conditions that avoid the extreme illiquidity must be identified. For $m=0$, $X_n=\frac{\frac{1}{n}}{\frac{1}{n}+1-u}=N_u^{-1}$ and

$X_n < \frac{1}{9} \Leftrightarrow N_u > 9 \Leftrightarrow n > \frac{8}{1-u}$. Also,

$$\alpha_{n,0}^{\pm} = \frac{1}{8u}\left\{N_u + 18 - 27N_u^{-1} \pm \sqrt{\left(N_u + 18 - 27N_u^{-1}\right)^2 - 64N_u}\right\},$$

and $\lim_{n\to\infty} \alpha_{n,0}^{\pm} = \left(\frac{4}{u}, \infty\right)$. Recall that $\alpha_{n,0}^{-}$ approaches $\frac{4}{u}$ from below. Therefore, if $\alpha < \frac{4}{u}$, there is a unique solution to (60) for sufficiently large n. However, this is when the extreme negative externalities arise and the limit value is not a trade equilibrium. If $\alpha \geq \frac{4}{u}$, there are three solutions to (60) for sufficiently large n, and the larger two solutions are bounded away from zero. These two solutions form trade equilibria in the limit. The smallest solution approaches zero and the extreme negative externalities arise. Hence, we conclude that the case $m = 0$ with $\alpha \geq \frac{4}{u}$ yields two equilibria in the limit $n \to \infty$. This is indeed the case studied by Ganguli and Yang (2009). For $m \in \left(0, \frac{1}{9}\right)$, $\lim_{n\to\infty} \alpha_{n,m}^{\pm}$ is well-defined, with X_n replaced with its limit $X = m$ in (57). Therefore three solutions for (60) persist in the limit $n \to \infty$ if $\alpha \in \left(\lim_{n\to\infty} \alpha_{n,m}^{-}, \lim_{n\to\infty} \alpha_{n,m}^{+}\right)$. Because $\lim_{n\to\infty} \varphi = 0$ for all three solutions, they are all trade equilibria in the limit. Therefore, Ganguli and Yang (2009)'s analysis misses the case with three equilibria.

Manzano and Vives (2011) assume $t = 1$, $u \in (0, 1]$, $w \in [0, 1]$. Therefore, the extreme illiquidity possible with $u < w = 1$ as discussed above. For $w \in \left(\frac{8}{9}, 1\right)$, $\lim_{n\to\infty} \alpha_{n,m}^{\pm}$ is well-defined for $m \in \left(1 - w, \frac{1}{9}\right)$ and three solutions for (60) persist in the limit $n \to \infty$ if $\alpha \in \left(\lim_{n\to\infty} \alpha_{n,m}^{-}, \lim_{n\to\infty} \alpha_{n,m}^{+}\right)$. They are all trade equilibria in the limit.

All in all, we established the strategic foundation for equilibrium multiplicity for the cases $w = t = 1$ with $\alpha \geq \frac{4}{u}$ (two equilibria), and $w \in \left(\frac{8}{9}, 1\right]$, $m \in (1 - w, \frac{1}{9})$ with $\alpha \in \left(\lim_{n\to\infty} \alpha_{n,m}^{-}, \lim_{n\to\infty} \alpha_{n,m}^{+}\right)$ (three equilibria). For all these cases, our analysis shows that equilibrium with less speculation relative to hedging (larger $\frac{\beta_e}{\beta_s}$) and less informative prices (smaller φ) is better in terms of ex ante welfare.

7.4.6 Negative externalities in a large market

Lemma A10 (negative externalities with $w \in (0, 1)$)

In a trade equilibrium with $w \in (0, 1)$, G decreases in n for sufficiently large n.

Proof.

It suffices to show that $\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right) A(T; t, w) B(T; u)$ is decreasing for sufficiently large n. **Lemma A7** shows that φ is decreasing in n, while $n\varphi$ is increasing in n. Therefore, $\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}$ is increasing in n. Because $T(n, \varphi; w) \equiv \frac{wn\varphi}{1-(1-w)\varphi}$,

$$T' \equiv \frac{dT}{dn} = w\frac{(n\varphi)'\{1 - (1-w)\varphi\} + (1-w)\varphi' n\varphi}{\{1 - (1-w)\varphi\}^2}.$$

Using $\varphi' = \frac{(n\varphi)' - \varphi}{n}$, $T' = w\frac{(n\varphi)' - (1-w)\varphi^2}{\{1-(1-w)\varphi\}^2}$. Let φ be of order $n^{-\delta_\varphi}$ (from **Lemma A7** $\delta_\varphi = \frac{2}{3}$

if $u=1$ and $\delta_\varphi = 1$ if $u<1$.). Then $n\varphi$ is of order $n^{1-\delta_\varphi}$ and $(n\varphi)'$ is of order $n^{-\delta_\varphi}$. Hence, $T'>0$ for sufficiently large n. Because $A(T;t,w)B(T;u)$ is decreasing in T, $A(n\varphi;t)B(n\varphi;t)$ is decreasing in n for sufficiently large n.

We show that $\frac{d}{dn}\left\{\log\left(\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right)\right\}<0$ for sufficiently large n. First,

$$\frac{d}{dn}\left\{\log\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)\right\}=-\frac{1}{n}+\frac{1}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\left\{1-w\left(\frac{\varphi}{1-\varphi}+(n+1)\frac{\varphi'}{(1-\varphi)^2}\right)\right\}.$$

From $T'=w\frac{(n\varphi)'-(1-w)\varphi^2}{\{1-(1-w)\varphi\}^2}$, $(n\varphi)'=\frac{T'\{1-(1-w)\varphi\}^2}{w}+(1-w)\varphi^2$. Substitute $\varphi'=\frac{(n\varphi)'-\varphi}{n}$ to obtain

$$\begin{aligned}
&\frac{d}{dn}\left\{\log\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)\right\}\\
=\;&\frac{1+(n+1)\frac{w\varphi}{1-\varphi}}{n\left\{n-1-(n+1)\frac{w\varphi}{1-\varphi}\right\}}\\
&-\frac{1}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\frac{w}{(1-\varphi)^2}\left[\varphi(1-\varphi)+\frac{n+1}{n}\left\{\frac{T'\{1-(1-w)\varphi\}^2}{w}+(1-w)\varphi^2-\varphi\right\}\right]\\
=\;&\frac{1}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\left[\frac{1+(n+1)\frac{w\varphi}{1-\varphi}}{n}-\frac{w}{(1-\varphi)^2}\left\{\varphi(1-\varphi)+\frac{n+1}{n}\left((1-w)\varphi^2-\varphi\right)\right\}\right]\\
&-\frac{\frac{n+1}{n}}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\left(\frac{1-(1-w)\varphi}{1-\varphi}\right)^2 T'.
\end{aligned}$$

The expression in the square bracket can be written as

$$\begin{aligned}
&\frac{1}{n}+\frac{w\varphi}{1-\varphi}\left[\frac{n+1}{n}-\left\{1-\frac{n+1}{n}\frac{1-(1-w)\varphi}{1-\varphi}\right\}\right]\\
=\;&\frac{1}{n}+\frac{w\varphi}{1-\varphi}\left\{\frac{1}{n}+\frac{n+1}{n}\frac{1-(1-w)\varphi}{1-\varphi}\right\}\\
=\;&\frac{1}{n}\left[1+\frac{w\varphi}{1-\varphi}\left\{1+(n+1)\frac{1-(1-w)\varphi}{1-\varphi}\right\}\right].
\end{aligned}$$

Next,

$$\frac{d}{dn}\log A(T;t,w)=\frac{T'}{N_A D_A}\left\{(1-w+(w-t)c_s)D_A-(1-w+wc_s)N_A\right\},$$

where N_A is the numerator and D_A is the denominator of

$$A(T;t,w)\equiv\frac{1-tc_s+(1-w+(w-t)c_s)T(n,\varphi;w)}{1+(1-w+wc_s)T(n,\varphi;w)}.$$

Calculating this, $\frac{d}{dn}\log A(T;t,w) = -\frac{wtc_s(1-c_s)}{N_A D_A}T'$. Similarly, $\frac{d}{dn}\log B(T;u) = -\frac{1-u}{u\{1+T(1-u)\}}T'$. Therefore,

$$\begin{aligned}
&\frac{d}{dn}\left\{\log\left(\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right)\right\} \\
=\ & \frac{1}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\frac{1}{n}\left[1+\frac{w\varphi}{1-\varphi}\left\{1+(n+1)\frac{1-(1-w)\varphi}{1-\varphi}\right\}\right] \\
& -\left[\frac{\frac{n+1}{n}}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\left(\frac{1-(1-w)\varphi}{1-\varphi}\right)^2 + \frac{wtc_s(1-c_s)}{N_A D_A} + \frac{1-u}{u\{1+T(1-u)\}}\right]T'.
\end{aligned}$$

Recall that given φ of order $n^{-\delta_\varphi}$ with δ_φ being either 1 (for $u<1$) or $\frac{2}{3}$ (for $u=1$), T is of order $n^{1-\delta_\varphi}$ and T' is of order $n^{-\delta_\varphi}$. Hence,

$$N_A D_A = \{1 - tc_s + (1-w+(w-t)c_s)T\}\{1+(1-w+wc_s)T\}$$

is of order $n^{2(1-\delta_\varphi)}$. Also, $\frac{1-u}{u\{1+T(1-u)\}}$ is of order $n^{-(1-\delta_\varphi)}$ if $u<1$. Finally, $\frac{\frac{n+1}{n}}{n-1-(n+1)\frac{w\varphi}{1-\varphi}}\left(\frac{1-(1-w)\varphi}{1-\varphi}\right)^2$ is of order n^{-1}. Combining these, the second term of $\frac{d}{dn}\left\{\log\left(\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right)\right\}$ is of order $n^{\max\{-1,-2(1-\delta_\varphi)\}-\delta_\varphi}$ with $\delta_\varphi = \frac{2}{3}$ if $u=1$ (i.e. order $n^{-2+\frac{2}{3}} = n^{-\frac{4}{3}}$), and of order $n^{\max\{-1,-2(1-\delta_\varphi),-(1-\delta_\varphi)\}-\delta_\varphi}$ with $\delta_\varphi = 1$ if $u<1$ (i.e. order n^{-1}). The first term of $\frac{d}{dn}\left\{\log\left(\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right)\right\}$ is of order $n^{-2+1-\delta_\varphi} = n^{-(1+\delta_\varphi)}$. For $u<1$, $\delta_\varphi = 1$ and $-(1+\delta_\varphi) = -2 < -1$. For $u=1$, $\delta_\varphi = \frac{2}{3}$ and $-(1+\delta_\varphi) = -\frac{5}{3} < -\frac{4}{3}$. For both cases, the positive first term goes to zero faster than the negative second term as $n\to\infty$. Thus, for sufficiently large n,

$$\frac{d}{dn}\left\{\log\left(\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right)\right\} < 0.$$

■

Lemma A11 (negative externalities with $w=1$)

In a trade equilibrium with $w=1$,

(a) *G decreases in n for sufficiently large n.*

(b) *If $t<1$, or $[t=1$ and $\frac{4\tau_\varepsilon\tau_x}{\rho^2} < u < 1]$, then $\lim_{n\to\infty} G \in (0,\infty)$.*

(c) *If $t=u=1$, then given $1 < \frac{\rho^2}{\tau_\varepsilon\tau_x}$, $\lim_{n\to\infty} G = 0$ and $\lim_{n\to\infty} nG \in (0,\infty)$.*

(d) *If $t=1$ and $u < \min\left\{1, \frac{4\tau_\varepsilon\tau_x}{\rho^2}\right\}$, then $\lim_{n\to\infty} G = \lim_{n\to\infty} nG = 0$.*

Proof.
We prove **(c)** and **(d)** first. For **(c)**, **Lemma A8** shows that φ is constant. From

Lemma A5,

$$\begin{aligned} G &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{\varphi}{1-\varphi}\right)\frac{1-c_s}{1+c_sn\varphi}\right) \\ &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}\frac{1-c_s}{1+c_sn\varphi}\right). \end{aligned}$$

We have

$$\lim_{n\to\infty}\left(1+\frac{\rho^2}{\tau_v\tau_x}\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}\frac{1-c_s}{1+c_sn\varphi}\right)=1.$$

Also,

$$n\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}\frac{1-c_s}{1+c_sn\varphi}\right)=\log\left\{\left(1+\frac{\rho^2}{\tau_v\tau_x}\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}\frac{1-c_s}{1+c_sn\varphi}\right)^n\right\}$$

has a finite limit as $n\to\infty$. It is clear that G is strictly decreasing in n for sufficiently large n. For **(d)**, **Lemma A8** and **A9** show that φ increases so that $\frac{n-1}{n}-2\varphi$ hits zero with finite n and $\varphi<1$. Since $A(n\varphi;t)B(n\varphi;u)$ is decreasing in $n\varphi$, $\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}A(n\varphi;t)B(n\varphi;u)$ must be decreasing in n before it hits zero.

(a) For all other cases, **Lemma A8** shows that φ is decreasing in n, while $n\varphi$ is increasing in n. Therefore, by imposing $w=1$ in **Lemma A5**, we know that $\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}$ is increasing in n while $A(n\varphi;t)B(n\varphi;t)$ is decreasing in n.

We show that $\left(\frac{n-1}{n}-2\varphi\right)n^2\frac{d}{dn}\left\{\log\left(\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}A(n\varphi;t)B(n\varphi;u)\right)\right\}<0$ for sufficiently large n. First,

$$\begin{aligned} &\left(\frac{n-1}{n}-2\varphi\right)n^2\frac{d}{dn}\left\{\log\left(\frac{\frac{n-1}{n}-2\varphi}{1-\varphi}A(n\varphi;t)B(n\varphi;u)\right)\right\} \\ &= 1-n\left\{(n+1)(1-\varphi)\varphi'+c_sK(n-1-2n\varphi(n\varphi)')\right\}, \end{aligned}$$

where $K\equiv\frac{t}{\{1-tc_s+(1-t)c_sn\varphi\}(1+c_sn\varphi)}+\frac{1-u}{c_s\{1+(1-u)n\varphi\}}$ and $\varphi'\equiv\frac{\partial\varphi}{\partial n}$. Because $\varphi'=\frac{1}{n}((n\varphi)'-\varphi)$,

$$\begin{aligned} &(n+1)(1-\varphi)\varphi'+c_sK(n-1-2n\varphi(n\varphi)') \\ &= c_sK(n-1)-\frac{n+1}{n}(1-\varphi)\varphi+\left\{\frac{n+1}{n}(1-\varphi)-2c_sKn\varphi\right\}(n\varphi)'. \end{aligned} \tag{65}$$

It suffices to show that (65) is positive and not decreasing in n for sufficiently large n. If $u=1$ and $t<1$, then from **Lemma A8** $n\varphi$ increases at the rate $n^{\frac{1}{3}}$. Hence, $K=\frac{t}{\{1-tc_s+(1-t)c_sn\varphi\}(1+c_sn\varphi)}$ decreases at the rate $n^{-\frac{2}{3}}$ and $K(n-1)$ increases at the rate $n^{\frac{1}{3}}$. Because $Kn\varphi$ decreases at the rate $n^{-\frac{1}{3}}$, (65) is positive and increasing in n for sufficiently large n. For all other cases, from **Lemma A8** $n\varphi$ increases to a finite limit. Hence, K and $Kn\varphi$ are bounded and $(n\varphi)'$ approaches zero, while $K(n-1)$ is not bounded. Therefore, (65) is positive and increasing in n for sufficiently large n.

(b) From **Lemma A8**, $\lim_{n\to\infty}\varphi=0$. If $u<1$, then $n\varphi$ and $A(n\varphi;t)B(n\varphi;u)$ have positive finite limits. If $u = 1$, then $B(n\varphi; u) = 1$ and $\lim_{n\to\infty} n\varphi = \infty$ and hence $\lim_{n\to\infty} A(n\varphi; t) = 1 - t$. ■

7.4.7 Price-taking equilibrium

Lemma A12 (price-taking equilibrium)

(**a**) *A price-taking equilibrium exists for all* $n \geq 1$.

(**b**) *If* $w < 1$, *the optimal market size is smaller than that under the Nash assumption.*

(c) *If* $w = 1$, *the optimal market size is smaller than that under the Nash assumption if* (i) $t < 1$ *or* (ii) $t = 1$ *and* $\frac{4\tau_\varepsilon \tau_x}{\rho^2} \leq u < 1$ *or* (iii) $t = u = 1$ *and* $1 < \frac{\rho^2}{\tau_\varepsilon \tau_x}$.

(d) *The optimal market size is finite.*

Proof.
(a) (37) with $\lambda = 0$ is satisfied for all $n \geq 1$.
(b) By **Lemma A5**,

$$\exp\left(2\rho G^{pt}\right) = 1 + \{\exp(2\rho G) - 1\} \frac{1}{\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}} \frac{n}{1+n} \frac{1-\varphi}{1-(1-w)\varphi}. \tag{66}$$

From the proof of **Lemma A4**, $1 - \widetilde{\lambda} = \frac{1-(1-w)\varphi}{1-\varphi}\frac{n+1}{n}\left(\frac{n-1}{n} - \frac{n+1}{n} w \frac{\varphi}{1-\varphi}\right) < 1$. Hence, $\exp(2\rho G^{pt}) = \frac{\exp(2\rho G) - \widetilde{\lambda}}{1-\widetilde{\lambda}}$. The sign of $\frac{d\exp\left(2\rho G^{pt}\right)}{dn}$ equals the sign of

$$\left(\frac{d\exp(2\rho G)}{dn} - \widetilde{\lambda}'\right)\left(1 - \widetilde{\lambda}\right) + \widetilde{\lambda}'\left\{\exp(2\rho G) - \widetilde{\lambda}\right\} = \widetilde{\lambda}'\{\exp(2\rho G) - 1\} + \frac{d\exp(2\rho G)}{dn}\left(1 - \widetilde{\lambda}\right),$$

where $\widetilde{\lambda}' \equiv \frac{d\widetilde{\lambda}}{dn}$. From the proof of **Lemma A10** and **A11**, the second term is negative for sufficiently large n. Also, it is easy to verify $\widetilde{\lambda}' < 0$ for $\widetilde{\lambda} \equiv \left(\frac{\left(w - \frac{1-w}{n}\right)\varphi + \frac{1}{n}}{1-\varphi}\right)^2$, which implies that the first term is also negative. Thus, G^{pt} is decreasing in n for sufficiently large n.

To compare n^* and n^*_{pt}, use

$$\exp(2\rho G) = \left(1 - \widetilde{\lambda}\right)\exp\left(2\rho G^{pt}\right) + \widetilde{\lambda}.$$

Because $\exp(2\rho G)$ is a weighted average of $\exp(2\rho G^{pt})$ and 1, where $\exp(2\rho G^{pt}) > 1$ and the weight on 1 is $\widetilde{\lambda} \in (0, 1)$, $n^*_{pt} < n^*$ if $\widetilde{\lambda}$ is decreasing in n. By **Lemma A8**, φ is decreasing in n, and so is $\widetilde{\lambda}$. This proves (**b**).

(c) With $w = 1$, (66) becomes

$$\exp\left(2\rho G^{pt}\right) = 1 + \{\exp(2\rho G) - 1\} \frac{1-\varphi}{\frac{n-1}{n} - 2\varphi} \frac{n(1-\varphi)}{1+n}$$

and

$$\widetilde{\lambda} = \left(\frac{\varphi + \frac{1}{n}}{1-\varphi}\right)^2.$$

Also, we can write

$$\begin{aligned} 1 - \widetilde{\lambda} &= \left(1 + \frac{\varphi + \frac{1}{n}}{1-\varphi}\right)\left(1 - \frac{\varphi + \frac{1}{n}}{1-\varphi}\right) \\ &= \frac{1+n}{n} \frac{\frac{n-1}{n} - 2\varphi}{(1-\varphi)^2}. \end{aligned}$$

Hence, $\exp(2\rho G^{pt}) = \frac{\exp(2\rho G) - \widetilde{\lambda}}{1-\widetilde{\lambda}}$ whenever $\widetilde{\lambda} < 1$. The sign of $\frac{d\exp\left(2\rho G^{pt}\right)}{dn}$ equals the sign of

$$\left(\frac{d\exp(2\rho G)}{dn} - \widetilde{\lambda}'\right)\left(1 - \widetilde{\lambda}\right) + \widetilde{\lambda}'\left\{\exp(2\rho G) - \widetilde{\lambda}\right\} = \widetilde{\lambda}'\{\exp(2\rho G) - 1\} + \frac{d\exp(2\rho G)}{dn}\left(1 - \widetilde{\lambda}\right),$$

where $\widetilde{\lambda}' \equiv \frac{d\widetilde{\lambda}}{dn}$. For all cases where a trade equilibrium exists for large n, φ is not increasing in n. Hence, $\widetilde{\lambda}$ is decreasing in n, which implies that G^{pt} is decreasing in n for sufficiently large n. For the remaining case, φ is either constant or converges to one by **Lemma A8**. Therefore, $G^{pt} = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x}\frac{n(1-\varphi)}{1+n}B(n\varphi; u)\right)$ converges to zero. Therefore, G^{pt} is decreasing in n for sufficiently large n and $n^*_{pt} \equiv \arg\max_n G^{pt}$ exists.

To compare n^* and n^*_{pt}, use $\exp(2\rho G) = \left(1 - \widetilde{\lambda}\right)\exp(2\rho G^{pt}) + \widetilde{\lambda}$. Because $\exp(2\rho G)$ is a weighted average of $\exp(2\rho G^{pt})$ and 1, where $\exp(2\rho G^{pt}) > 1$ and $\widetilde{\lambda} \in (0,1)$, $n^* >(<) n^*_{pt}$ if $\widetilde{\lambda}$ is decreasing (increasing) in n. Note that $\widetilde{\lambda} = \left(\frac{\varphi + \frac{1}{n}}{1-\varphi}\right)^2$ is decreasing in n whenever φ is either constant or decreasing in n. From **Lemma A3(a)** and **Lemma A8**, the conditions (i)-(iii) in (**c**) imply φ is either constant or decreasing in n and that trade equilibrium exists for sufficiently large n.

(d) The remaining parameter values to be checked is the ones for which extreme illiquidity arises in trade equilibrium. From

$$G^{pt} = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x}\frac{n}{1+n}\frac{1-\varphi}{1-(1-w)\varphi}A(T; t, w)B(T; u)\right),$$

it is clear that G^{pt} approaches zero when φ approaches one. ■

Remark on price-taking equilibrium.

In any equilibrium of the form (6), the inference from the price is affected by the ratio of constants $\beta_s, \beta_e, \beta_p$, but not by their level. Because the price impact affects the three constants proportionally, it does not distort the information revealed by prices. This can be confirmed by observing that the three equations (39), (40), (41) show that ratios $\frac{\beta_p}{\beta_s}$ and $\frac{\beta_e}{\beta_s}$ are not affected by the value of λ. This implies that the characterization of k, φ, and p^* are the same in both trade equilibrium and price-taking equilibrium.[34] Thus, the optimal order in a price-taking equilibrium is obtained by replacing $\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi$ of the optimal order in a trade equilibrium with $1 - \varphi$. Therefore, a trade equilibrium converges to a price-taking equilibrium as n increases if and only if $\lim_{n\to\infty} \varphi = 0$. For $t = w = 1$, this does not occur if $u = 1$ or $u < \min\left\{\frac{4\tau_\varepsilon \tau_x}{\rho^2}, 1\right\}$. For all the other parameter values, the convergence occurs but achieving this convergence incurs the welfare loss because the optimal market size is finite.

In a price-taking equilibrium, traders do not internalize the impact of their demand on prices. As a result, a price-taking equilibrium always exists and the three constants $\beta_s, \beta_e, \beta_p$ are larger compared with those in a trade equilibrium. This implies that under the Nash assumption trading volume is reduced relative to that under the price-taking assumption, while the amount of information sharing is not affected as discussed above. Because beliefs are the same in both equilibria but there is less trading (hence less risk-sharing) in a trade equilibrium, the GFT are smaller in a trade equilibrium than in a price-taking equilibrium.

In the old version of the paper we show that the Nash implementation of the price-taking allocation is possible by using a subsidy which takes a form of volume-discount. Our result here indicates that when the market suffers from extreme negative externalities, this subsidy makes the optimal market size larger and therefore might have a large welfare impact. However, for the other cases, the subsidy makes the optimal market size smaller.

7.4.8 Role of t, w, u

Lemma A13 ($t, w, u \to 0^+$)

(a) $\lim_{t\to 0^+} k = \infty$, $\lim_{t\to 0^+} \varphi = 0$ and $\lim_{t\to 0^+} G = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v \tau_x}\frac{n-1}{n}u\right)$.

(b) $\lim_{w\to 0^+} k = \frac{1}{\sqrt{t}}\frac{1-tc_s}{1-c_s}$, $\lim_{w\to 0^+} \varphi = \left\{1 + \frac{\alpha u N_u}{t}\left(\frac{1-tc_s}{1-c_s}\right)^2\right\}^{-1}$ *and*

$$\lim_{w\to 0^+} G = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v \tau_x}\frac{n-1}{n}\left(1 - tc_s\right)u\right).$$

(c) *Optimal market size goes to infinity as* $t \to 0^+$ *or* $w \to 0^+$.

(d) $\lim_{u\to 0^+} k = \frac{1}{\sqrt{t}}\left(\frac{1-tc_s}{1-c_s} - w\frac{n}{1+n}\right)$, $\lim_{u\to 0^+} \varphi = 1$. *For any fixed* n *for which trade equilibrium exists,* $\underline{u} \in (0,1)$ *defined by* $\frac{n+1}{n-1} = \alpha \underline{u} N_{\underline{u}} k^2 + 1 - w$ *exists and* $\forall u \leq \underline{u}$ *the second order condition* (30) *is violated.*

[34] See page 623 in Madhavan (1992) for more discussion on this point.

If $w<1$, *then* β_s, β_e, β_p *have positive limits as* $u \searrow \underline{u}$.

If $w=1$, *then* $\lim_{u\searrow\underline{u}} \beta_s = \lim_{u\searrow\underline{u}} \beta_e = \lim_{u\searrow\underline{u}} \beta_p = 0$ *and* $\lim_{u\searrow\underline{u}} \frac{1}{n\beta_p} = \infty$.

Proof.
Recall $\varphi = (1+\alpha u N_u k^2)^{-1}$, where k is characterized by

$$F(k) \equiv \left(\alpha u N_u k^2 + w\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + \ wn\left\{(1-u)\sqrt{t}k - \left(\frac{1-tc_s}{1-c_s} - w\right)\right\} = 0.$$

Also,

$$\begin{aligned} T(n,\varphi;w) &\equiv \frac{wn\varphi}{1-(1-w)\varphi}, \\ A(T;t,w) &\equiv \frac{1-tc_s+(1-w+(w-t)c_s)T(n,\varphi;w)}{1+(1-w+wc_s)T(n,\varphi;w)}, \\ B(T;u) &\equiv \frac{u}{1+(1-u)T(n,\varphi;w)}, \\ G &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right). \end{aligned}$$

(a) First, if $u=1$, $F(k)$ becomes

$$F(k) = \left(\alpha k^2 + w\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) - wn\left(\frac{1-tc_s}{1-c_s} - w\right) = 0.$$

From **Lemma A6** the solution is unique. The solution must satisfy $\sqrt{t}k > \frac{1-tc_s}{1-c_s}$. This implies $\lim_{t\to0^+} k = \infty$. The limit solution must satisfy

$$\lim_{t\to0^+}\left(\alpha k^2 + w\right)\left(\sqrt{t}k - \frac{1}{1-c_s}\right) = wn\left(\frac{1}{1-c_s} - w\right).$$

Thus, $\sqrt{t}k$ must approach $\frac{1}{1-c_s}$ from above. Therefore, for sufficiently small t, $k \sim t^{-\frac{1}{2}}$ and $\varphi = \frac{1}{1+\alpha k^2} = \frac{k^{-2}}{k^{-2}+\alpha} \sim \frac{t}{t+\alpha}$.

For $u<1$, for sufficiently small t, $\sqrt{t}k \in \left(\frac{1-tc_s}{1-c_s}, \frac{\frac{1-tc_s}{1-c_s}-w}{1-u}\right)$ if $u \in (w(1-c_s), 1)$, while $\sqrt{t}k \in \left(\frac{\frac{1-tc_s}{1-c_s}-w}{1-u}, \frac{1-tc_s}{1-c_s}\right)$ if $u \le w(1-c_s)$. Either way, $\lim_{t\to0^+}\sqrt{t}k \in (0,\infty)$. Hence for sufficiently small t, $k \sim t^{-\frac{1}{2}}$ and $\varphi = \frac{1}{1+\alpha u N_u k^2} = \frac{k^{-2}}{k^{-2}+\alpha u N_u} \sim \frac{t}{t+\alpha u N_u}$. Using

$$\lim_{t\to0^+} T(n,\varphi;w) = 0, \ \lim_{t\to0^+} A(T;t,w) = 1, \ \lim_{t\to0^+} B(T;u) = u$$

in the expression of G yields the result for $\lim_{t\to0^+} G$.

(b) From **Lemma A6** $F(k)$ has a unique solution for small w. Because only the cases

Lemma A7 (ii) or (iii) apply, $\sqrt{t}k \in \left(\frac{1-tc_s}{1-c_s}, \frac{\frac{1-tc_s}{1-c_s}-w}{1-u}\right)$. In fact,

$$\lim_{w\to 0^+} F(k) = \alpha u N_u k^2 \left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right).$$

Therefore $\lim_{w\to 0^+} k = \frac{1}{\sqrt{t}}\frac{1-tc_s}{1-c_s}$. Using this in $\varphi = (1+\alpha u N_u k^2)^{-1}$ yields $\lim_{w\to 0^+}\varphi$. Using

$$\lim_{w\to 0^+} T(n,\varphi;w) = 0,\ \lim_{w\to 0^+} A(T;t,w) = 1-tc_s,\ \lim_{w\to 0^+} B(T;u) = u$$

in the expression of G yields the result for $\lim_{w\to 0^+} G$.

(c) First, multiplicity does not occur in both limits. Recall that the negative impact of larger n comes from $A(T;t,w)B(T;u)$ decreasing in T and $T(n,\varphi;w)$ increasing in n. As $w \to 0^+$, both $T(n,\varphi;w) \equiv \frac{wn\varphi}{1-(1-w)\varphi}$ and $T' = w\frac{d}{dn}\left(\frac{n\varphi}{1-(1-w)\varphi}\right)$ approaches zero.

As $t \to 0^+$, $\varphi \to 0$. Since

$$\begin{aligned} T' &= w\frac{(n\varphi)'\{1-(1-w)\varphi\}+(1-w)\varphi' n\varphi}{\{1-(1-w)\varphi\}^2} \\ &= w\frac{(n\varphi'+\varphi)\{1-(1-w)\varphi\}+(1-w)\varphi' n\varphi}{\{1-(1-w)\varphi\}^2}, \end{aligned}$$

and

$$\varphi' = \alpha u(1-u)(\varphi k)^2,$$

T' approaches zero if φk does so. From the proof of (a), $\varphi k \sim t^{\frac{1}{2}}$. On the other hand, the positive impact of n remains for both limits because $\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi} \underset{t\to 0^+ \text{ or } w\to 0^+}{\longrightarrow} \frac{n-1}{n}$. Therefore, the optimal market size goes to infinity.

(d) Note that $\alpha^{\pm}_{n,m}$ defined in (57) both go to ∞ as $u \to 0^+$. Hence, there is a unique solution to $F(k)$ for sufficiently small u.[35] In fact,

$$\begin{aligned} \lim_{u\to 0^+} F(k) &= w\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + wn\left\{\sqrt{t}k - \left(\frac{1-tc_s}{1-c_s} - w\right)\right\} \\ &= w\left\{\sqrt{t}k(1+n) - \frac{1-tc_s}{1-c_s}(1+n) + wn\right\}. \end{aligned}$$

Therefore $\lim_{u\to 0^+} k = \frac{1}{\sqrt{t}}\left(\frac{1-tc_s}{1-c_s} - w\frac{n}{1+n}\right)$ and $\lim_{u\to 0^+}\varphi = \lim_{u\to 0^+}(1+\alpha u N_u k^2)^{-1} = 1$.

Suppose a trade equilibrium exists for a given n. Because the second order condition is satisfied,

$$\begin{aligned} \frac{n+1}{n-1} &< \frac{1-\varphi}{\varphi} + 1 - w \\ &= \alpha u N_u k^2 + 1 - w. \end{aligned}$$

[35] Note also that for Δ defined in (54), $\lim_{u\to 0^+}\Delta = 0$.

Because $\frac{n+1}{n-1} > 1 > 1 - w$, for sufficiently small $\underline{u} \in (0, 1)$, the condition above holds as the equality. With this $\underline{u}$,

$$\frac{n+1}{n-1} = \frac{1-\varphi}{\varphi} + 1 - w \Leftrightarrow \varphi = \left(\frac{n+1}{n-1} + w\right)^{-1} = \frac{n-1}{n+1+w\,(n-1)}.$$

From **Lemma A3**, the coefficients β_s, β_e, β_p on the optimal order are some constants times $\frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\varphi$. Hence,

$$\begin{aligned}
& \frac{n-1}{n} - \left(1 + w - \frac{1-w}{n}\right)\frac{n-1}{n+1+w\,(n-1)} \\
= \;& \frac{n-1}{n}\left\{1 - \frac{n\,(1+w) - (1-w)}{n+1+w\,(n-1)}\right\} \\
= \;& \frac{n-1}{n}\frac{n+1+w\,(n-1) - (n-1) - w\,(n+1)}{n+1+w\,(n-1)} \\
= \;& \frac{n-1}{n}\frac{2\,(1-w)}{n+1+w\,(n-1)}.
\end{aligned}$$

Therefore, if $w < 1$, the optimal order has positive coefficients in the limit $u \searrow \underline{u}$. If $w = 1$, the coefficients approach zero and the market becomes infinitely illiquid when equilibrium disappears. ■

Remark on the discontinuity at $w = 0$ and $w = 1$.

There is the discontinuity at $w = 0$ in terms of the second order condition. When $w = 0$, by Bayes rule $E_i[v_1] = \frac{\tau_\varepsilon}{\tau_1}s_i$ and $\tau_1 = \tau_v + \tau_\varepsilon$, where s_i is an identical signal for all traders.[36] Hence, (39) through (41) become $\beta_s = \frac{\tau_\varepsilon}{\lambda\tau+\rho}\frac{\tau}{\tau_1}\sqrt{t}$, $\beta_e = \frac{\rho}{\lambda\tau+\rho}$, $\beta_p = \frac{\tau}{\lambda\tau+\rho}$. Using $\lambda = \frac{1}{n\beta_p}$, obtain $\beta_p = \frac{n-1}{n}\frac{\tau}{\rho}$ and

$$q_i(p; H_i) = \frac{n-1}{n}\left\{\frac{\tau_\varepsilon}{\rho}\frac{\sqrt{t}}{1 + (1-t)\frac{\tau_\varepsilon}{\tau_v}}s_i - e_i - \frac{\tau}{\rho}p\right\}.$$

This is equivalent to (44) with $w \to 0^+$.[37] However, the second order condition when $w = 0$ is

$$0 < 1 + \frac{\tau}{\rho}\frac{2}{n\beta_p} = 1 + \frac{2}{n}\frac{n}{n-1} = \frac{n+1}{n-1},$$

which is satisfied for all $n > 1$. This is different from the limit of (30) as $w \to 0^+$, which is

[36] This can also be obtained by setting $w = 0$ in (32).

[37] Also, $\lim_{w\to 0}\frac{\tau_1}{\tau} = 1 + (1-t)\frac{\tau_\varepsilon}{\tau_v}$ from (43).

$\lim_{w \to 0^+} \varphi < \frac{n-1}{n+1}$. In the limit $w \to 0^+$, φ has a positive limit

$$\lim_{w \to 0^+} \varphi = \left(1 + \alpha u N_u \left(\lim_{w \to 0^+} k\right)^2\right)^{-1} \in (0,1),$$

where $\lim_{w \to 0^+} k = \frac{1+(1-t)\frac{\tau_\varepsilon}{\tau_v}}{\sqrt{t}}$. Hence, in the limit $w \to 0^+$, $\varphi < \frac{n-1}{n+1}$ still imposes an implicit restriction on the other parameters (t, u, n, α, c_s). For example, for sufficiently small α or u (i.e., less risk-sharing needs relative to speculation needs), the limit value of φ is close to one, which violates $\varphi < \frac{n-1}{n+1}$ given a finite n. With a finite number of traders, whether or not traders learn from prices makes a difference to the second order condition, no matter how little they learn from prices. Assuming $w = 0$ ignores this discontinuity.

There is another type of discontinuity at $w = 1$ as shown in **Lemma A13(c)**. Starting from parameter values for which a trade equilibrium exists, as a common shock to endowments become more important ($u \searrow \underline{u}$), the second order condition is eventually violated and any trade equilibrium disappears. When equilibrium disappears, the market becomes infinitely illiquid if and only if $w = 1$. In terms of trading volume, given iid errors in signals ($w = 1$), volume smoothly decreases in u to zero when equilibrium disappears. On the other hand, when there is correlation in signals ($w < 1$), there is non-zero trading when equilibrium disappears.

7.4.9 Parameterization of τ_x and τ_ε

Suppose that, instead of constant $(\tau_v, \tau_\varepsilon, \tau_x)$, we have $(\widetilde{\tau}_v, \widetilde{\tau}_\varepsilon, \widetilde{\tau}_x)$ that depends on n as follows:

$$\widetilde{\tau}_v = \tau_v n^{\delta_v},\ \widetilde{\tau}_\varepsilon = \tau_\varepsilon n^{\delta_\varepsilon},\ \widetilde{\tau}_x = \tau_x n^{\delta_x}.$$

Our baseline model has $(\delta_v, \delta_\varepsilon, \delta_x) = (0,0,0)$. We investigate how our main result changes as $(\delta_v, \delta_\varepsilon, \delta_x)$ changes. These parameters *exogenously* determine how the underlying information structure changes as n changes. For example, if $\delta_v > 0$ (< 0), then traders have more (less) precise prior knowledge about v_0 and v_1 as n increases. If $\delta_\varepsilon > 0$ (< 0), then traders have more (less) precise signals about v_1 as n increases. If $\delta_x < 0$ (> 0), then traders have more (less) volatile endowment shocks x_0 and x_1 as n increases. Noting that

$$\begin{aligned}
\widetilde{\alpha} &\equiv \frac{\rho^2}{\widetilde{\tau}_\varepsilon \widetilde{\tau}_x} = \frac{\rho^2}{\tau_\varepsilon \tau_x} n^{-(\delta_\varepsilon + \delta_x)} = \alpha n^{-(\delta_\varepsilon + \delta_x)}, \\
\widetilde{c}_s &\equiv \frac{\widetilde{\tau}_\varepsilon}{\widetilde{\tau}_v + \widetilde{\tau}_\varepsilon} = \frac{\tau_\varepsilon n^{\delta_\varepsilon}}{\tau_v n^{\delta_v} + \tau_\varepsilon n^{\delta_\varepsilon}} = \frac{\tau_\varepsilon}{\tau_v n^{(\delta_\varepsilon - \delta_v)} + \tau_\varepsilon},
\end{aligned}$$

the cubic equation that characterizes k changes to:

$$\begin{aligned}
F(k) &\equiv \left(\widetilde{\alpha} u N_u k^2 + w\right)\left(\sqrt{t}k - \frac{1 - t\widetilde{c}_s}{1 - \widetilde{c}_s}\right) + \ wn\left\{(1-u)\sqrt{t}k - \left(\frac{1 - t\widetilde{c}_s}{1 - \widetilde{c}_s} - w\right)\right\} \\
&= \left(\alpha u N_u k^2 n^{-(\delta_\varepsilon + \delta_x)} + w\right)\left(\sqrt{t}k - \frac{1 - t\widetilde{c}_s}{1 - \widetilde{c}_s}\right) + \ wn\left\{(1-u)\sqrt{t}k - \left(\frac{1 - t\widetilde{c}_s}{1 - \widetilde{c}_s} - w\right)\right\}.
\end{aligned}$$

Also, ex ante gains from trade become

$$\begin{aligned} G &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\widetilde{\tau}_v\widetilde{\tau}_x}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right) \\ &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}n^{-(\delta_v+\delta_x)}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right), \end{aligned}$$

where

$$\begin{aligned} T\left(n,\varphi;w\right) &\equiv \frac{wn\varphi}{1-(1-w)\varphi}, \\ A(T;t,w) &\equiv \frac{1-t\widetilde{c}_s+(1-w+(w-t)\widetilde{c}_s)\,T\left(n,\varphi;w\right)}{1+(1-w+w\widetilde{c}_s)\,T\left(n,\varphi;w\right)}, \\ B(T;u) &\equiv \frac{u}{1+(1-u)\,T\left(n,\varphi;w\right)}. \end{aligned}$$

To restrict our attention to $(\delta_\varepsilon,\delta_x)$, we impose the following restriction:

$$\delta_v=\delta_\varepsilon.$$

This means that precision of prior knowledge about v and that of private signals move in the same direction at the same rate as n increases. This can occur when both public signal and private signals are generated by the same source. This assumption leaves the ratio

$$\widetilde{c}_s\equiv\frac{\widetilde{\tau}_\varepsilon}{\widetilde{\tau}_v+\widetilde{\tau}_\varepsilon}=\frac{\tau_\varepsilon}{\tau_v n^{(\delta_\varepsilon-\delta_v)}+\tau_\varepsilon}=\frac{\tau_\varepsilon}{\tau_v+\tau_\varepsilon}=c_s$$

independent of n and simplifies the analysis. With this assumption, we have

$$F\left(k\right)=\left(\alpha uN_u k^2 n^{-(\delta_\varepsilon+\delta_x)}+w\right)\left(\sqrt{t}k-\frac{1-tc_s}{1-c_s}\right)+wn\left\{(1-u)\sqrt{t}k-\left(\frac{1-tc_s}{1-c_s}-w\right)\right\}=0, \tag{67}$$

and

$$G=\frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}n^{-(\delta_\varepsilon+\delta_x)}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right), \tag{68}$$

$$\varphi=\left\{1+\alpha uN_u k^2 n^{-(\delta_\varepsilon+\delta_x)}\right\}^{-1}. \tag{69}$$

where the expression for $T\left(n,\varphi;w\right)$, $A(T;t,w)$ and $B(T;u)$ are same as in the baseline case. Because only the sum $\delta_\varepsilon+\delta_x$ matters, we let

$$\delta\equiv\delta_\varepsilon+\delta_x,$$

and focus on how the main results are affected by different values of δ. In the proof of **Lemma A14**, we characterize $k\sim n^{\delta_k}$, $\varphi\sim n^{\delta_\varphi}$, $G\sim n^{\delta_G}$ as $n\to\infty$.

Lemma A14 (Parameterization of τ_x and τ_ε)

(a) $\lim_{n\to\infty}\varphi = 0$, $\lim_{n\to\infty} n\varphi = \infty$, $\lim_{n\to\infty} G = 0$ *if and only if either one of the following conditions hold:*

(i) $0 < tw < u = 1$ *and* $\delta \in (0,2)$.

(ii) $0 < tw < 1$, $u < 1$ *and* $\delta \in (0,1)$.

(iii) $twu = 1$ *and* $\delta \in \left(-\frac{1}{2}, 0\right)$.

(iv) $u < tw = 1$ *and* $\delta = \delta_k \in (-1,0)$.

For each case above:

(b) *If* δ *is below the lower bound, then* $\lim_{n\to\infty} G = \infty$.

(c) *If* δ *is above the upper bound,* $\lim_{n\to\infty}\varphi = 1$ *and extreme negative externalities arise.*

Proof.
Conjecture $k \sim n^{\delta_k}$, $\varphi \sim n^{\delta_\varphi}$, $G \sim n^{\delta_G}$ as $n \to \infty$. We characterize $(\delta_k, \delta_\varphi, \delta_G)$.

(a)(i)
$[w < u = 1]$
First, (67), (69) and (68) become

$$\begin{aligned} F(k) &= \left(\alpha k^2 n^{-\delta} + w\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) - wn\left(\frac{1-tc_s}{1-c_s} - w\right) = 0, \\ \varphi &= \left(1 + \alpha k^2 n^{-\delta}\right)^{-1}, \end{aligned}$$

$$G = \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v \tau_x} n^{-\delta}\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)\right).$$

From the conjecture $k \sim n^{\delta_k}$ and $F(k)$, $\delta_k < 0$ implies that $F(k) < 0$ for sufficiently large n. This cannot be the limit and hence $\delta_k \geq 0$. There are two cases to consider. If $2\delta_k - \delta \geq 0$, then

$$2\delta_k - \delta + \delta_k = 1 \Leftrightarrow \delta_k = \frac{1}{3}(1+\delta).$$

This occurs when

$$2\delta_k \geq \delta \Leftrightarrow \frac{2}{3}(1+\delta) \geq \delta \Leftrightarrow \delta \leq 2.$$

If $2\delta_k - \delta < 0$, then $\delta_k = 1$. This occurs when $2\delta_k = 2 < \delta$. Therefore, $\delta_k(\delta)$ is given by

$$\delta_k(\delta) = \begin{cases} \frac{1+\delta}{3} & \text{if } \delta \leq 2 \\ 1 & \text{if } \delta > 2 \end{cases}.$$

For φ, there are two cases to consider. If $2\delta_k - \delta > 0$, then $\delta_\varphi = \delta - 2\delta_k$. If $2\delta_k - \delta \leq 0$, then $\lim_{n\to\infty}\varphi \in (0,1]$ and $\delta_\varphi = 0$. Note that $2\delta_k(\delta) = \delta$ has only one solution $\delta = 2$ and

$2\delta_k - \delta \gtreqless 0 \Leftrightarrow \delta \lesseqgtr 2$. Therefore, $\delta_\varphi(\delta)$ is given by

$$\delta_\varphi(\delta) = \begin{cases} \frac{\delta-2}{3} & \text{if } \delta \leq 2 \\ 0 & \text{if } \delta > 2 \end{cases}.$$

Thus, $\lim_{n\to\infty} \varphi = 0$ if and only if $\delta < 2$. Because $n\varphi$ is of order $\delta_\varphi + 1$, $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta_\varphi(\delta) + 1 > 0 \Leftrightarrow \frac{\delta+1}{3} > 0 \Leftrightarrow \delta > -1$. Therefore, $\lim_{n\to\infty} \varphi = 0$ and $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta \in (-1, 2)$.

Finally, for $\delta \in (-1, 2)$, consider G. Because $\lim_{n\to\infty} \left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right) = 1$ and

$$\begin{aligned}\lim_{n\to\infty} A(T;t,w) &= \frac{1 - tc_s + (1-w+(w-t)c_s)\lim_{n\to\infty} T(n,\varphi;w)}{1 + (1-w+wc_s)\lim_{n\to\infty} T(n,\varphi;w)} \\ &= \frac{1-w+(w-t)c_s}{(1-w+wc_s)},\end{aligned}$$

$\delta_G = -\delta$. Therefore, $\lim_{n\to\infty} G = 0$ if $\delta \in (0, 2)$.

[$w = u = 1$, $t < 1$]

First, (67), (69) and (68) become

$$\begin{aligned} F(k) &= \left(\alpha k^2 n^{-\delta} + 1\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) - n\left(\frac{1-tc_s}{1-c_s} - 1\right) = 0, \\ \varphi &= \left(1 + \alpha k^2 n^{-\delta}\right)^{-1}, \\ G &= \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x}n^{-\delta}\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{\varphi}{1-\varphi}\right)A(T;t,1)\right). \end{aligned}$$

where $T(n,\varphi;1) = n\varphi$. The analysis is analogous to the case studied above [$w < u = 1$].

(a)(ii)

[$w < 1$, $u < 1$]

From, (67),

$$F(k) = \left(\alpha u N_u k^2 n^{-\delta} + w\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + \ wn\left\{(1-u)\sqrt{t}k - \left(\frac{1-tc_s}{1-c_s} - w\right)\right\} = 0.$$

This implies

$$\lim_{n\to\infty} k \in \left(\min\left\{\frac{1}{\sqrt{t}}\frac{1-tc_s}{1-c_s}, \frac{1}{\sqrt{t}(1-u)}\frac{1-tc_s}{1-c_s}\right\}, \max\left\{\frac{1}{\sqrt{t}}\frac{1-tc_s}{1-c_s}, \frac{1}{\sqrt{t}(1-u)}\frac{1-tc_s}{1-c_s}\right\}\right),$$

and hence $\lim_{n\to\infty} k$ is bounded away from zero and infinity. Therefore, $\delta_k(\delta) = 0$ for any δ.

From (69)

$$\varphi = \left\{1 + \alpha u\left(1 + (1-u)n\right)k^2 n^{-\delta}\right\}^{-1},$$

there are two cases to consider. If $2\delta_k - \delta + 1 > 0$, then $\delta_\varphi = \delta - 2\delta_k - 1$. If $2\delta_k - \delta + 1 \leq 0$, then $\lim_{n\to\infty} \varphi \in (0,1]$ and $\delta_\varphi = 0$. Because $\delta_k = 0$, $\delta_\varphi(\delta)$ is given by

$$\delta_\varphi(\delta) = \begin{cases} \delta - 1 & \text{if } \delta \leq 1 \\ 0 & \text{if } \delta > 1 \end{cases}.$$

Thus, $\lim_{n\to\infty} \varphi = 0$ if and only if $\delta < 1$. Because $n\varphi$ is of order $\delta_\varphi + 1$, $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta_\varphi(\delta) + 1 > 0 \Leftrightarrow \delta > 0$. Therefore, $\lim_{n\to\infty} \varphi = 0$ and $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta \in (0,1)$.

Finally, for $\delta \in (0,1)$, consider G. From (68),

$$G = \frac{1}{2\rho} \log\left(1 + \frac{\rho^2}{\tau_v \tau_x} n^{-\delta} \left(\frac{n-1}{n} - \frac{n+1}{n} \frac{w\varphi}{1-\varphi}\right) A(T;t,w) B(T;u)\right).$$

Because $\lim_{n\to\infty} \left(\frac{n-1}{n} - \frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right) = 1$, $\lim_{n\to\infty} A(T;t,w) = \frac{1-w+(w-t)c_s}{(1-w+wc_s)}$, and $T(n,\varphi;w) \equiv \frac{wn\varphi}{1-(1-w)\varphi}$ is of order $\delta_\varphi + 1$, $B(T;u) = \frac{u}{1+(1-u)T(n,\varphi;w)}$ is of order $-(\delta_\varphi + 1) = -\delta$. This implies $\delta_G = -2\delta$ and hence $\lim_{n\to\infty} G = 0$ if $\delta \in (0,1)$.

[$w = 1, u < 1, t < 1$]

First, (67), (69) and (68) become

$$\begin{aligned} F(k) &= \left(\alpha u N_u k^2 n^{-\delta} + 1\right)\left(\sqrt{t}k - \frac{1-tc_s}{1-c_s}\right) + n\left\{(1-u)\sqrt{t}k - \left(\frac{1-tc_s}{1-c_s} - 1\right)\right\} = 0, \\ \varphi &= \left\{1 + \alpha u\left(1 + (1-u)n\right)k^2 n^{-\delta}\right\}^{-1}, \\ G &= \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x} n^{-\delta}\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{\varphi}{1-\varphi}\right) A(T;t,1)\frac{u}{1+(1-u)n\varphi}\right). \end{aligned}$$

The analysis is analogous to the case studied above [$w < 1$, $u < 1$].

(a)(iii) $w = t = u = 1$.

First, (67), (69) and (68) become

$$\begin{aligned} F(k) &= \left(\alpha k^2 n^{-\delta} + 1\right)(k-1) = 0, \\ \varphi &= \left(1 + \alpha k^2 n^{-\delta}\right)^{-1}, \\ G &= \frac{1}{2\rho}\log\left(1 + \frac{\rho^2}{\tau_v\tau_x} n^{-\delta}\left(\frac{n-1}{n} - \frac{n+1}{n}\frac{\varphi}{1-\varphi}\right)\frac{1-c_s}{1+c_s n\varphi}\right). \end{aligned}$$

Because $k = 1$, $\delta_k(\delta) = 0$ and

$$\delta_\varphi(\delta) = \begin{cases} \delta & \text{if } \delta < 0 \\ 0 & \text{if } \delta \geq 0 \end{cases}.$$

Thus, $\lim_{n\to\infty} \varphi = 0$ and $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta \in (-1,0)$. For this case, $\delta_G = -(\delta + \delta_{n\varphi}) = -(\delta + 1 + \delta) = -(1+2\delta)$. Therefore, to have $\delta_G < 0$ and $\lim_{n\to\infty} G = 0$, we need $\delta > -\frac{1}{2}$.

(a)(iv) $w = t = 1$, $u < 1$.

First, (67), (69) and (68) become

$$\begin{aligned} F(k) &= \left(\alpha u N_u k^2 n^{-\delta} + 1\right)(k-1) + n(1-u)k = 0, \\ \varphi &= \left\{1 + \alpha u \left(1 + (1-u)n\right) k^2 n^{-\delta}\right\}^{-1}, \\ G &= \frac{1}{2\rho} \log\left(1 + \frac{\rho^2}{\tau_v \tau_x} n^{-\delta} \left(\frac{n-1}{n} - \frac{n+1}{n}\frac{\varphi}{1-\varphi}\right) \frac{1-c_s}{1+c_s n\varphi} \frac{u}{1+(1-u)n\varphi}\right). \end{aligned}$$

From $F(k)$, $k \in (0,1)$. So $\delta_k \leq 0$. Consider two cases: $\delta_k < 0$ and $\delta_k = 0$.

First, suppose $\delta_k < 0$. Then $\lim_{n\to\infty} k = 0$ and from $F(k)$, either

$$0 < 2\delta_k - \delta + 1 = 1 + \delta_k \Leftrightarrow \delta_k = \delta > -1,$$

or

$$0 \geq 2\delta_k - \delta + 1 \text{ and } 1 + \delta_k = 0 \Leftrightarrow \delta_k = -1 \text{ and } \delta \geq -1.$$

Note that the former case is possible only for $\delta \in (-1, 0)$ while the latter is possible for all $\delta \geq -1$. Hence, for $\delta \in (-1,0)$, $\delta_k(\delta)$ can take two values δ and -1.

Next, suppose $\delta_k = 0$. If $0 \geq -\delta + 1$, then $\lim_{n\to\infty} F(k) = \infty$. Therefore, $\delta < 1$ and either

$$0 < -\delta + 1 = 1 \Leftrightarrow \delta = 0,$$

or

$$\lim_{n\to\infty} (1-k) = 0 \text{ and } 0 < -\delta + 1 + \delta_{1-k} = 1 \Leftrightarrow \delta_{1-k} = \delta < 0.$$

Therefore, $\delta_k(\delta)$ is given by

$$\delta_k(\delta) = \begin{cases} 0 & \text{if } \delta < -1 \\ \{-1, 0\} & \text{if } \delta = -1 \\ \{-1, \delta, 0\} & \text{if } \delta \in (-1, 0) \\ \{-1, 0\} & \text{if } \delta = 0 \\ -1 & \text{if } \delta > 0 \end{cases}.$$

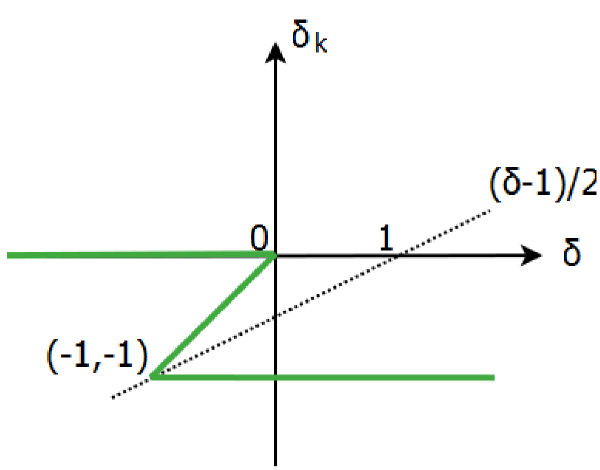


Figure A3. The order of k as a correspondence of δ.

For φ, there are two cases to consider. If $2\delta_k - \delta + 1 > 0$, then $\delta_\varphi = \delta - 2\delta_k - 1$. If $2\delta_k - \delta + 1 \leq 0$, then $\lim_{n\to\infty} \varphi \in (0,1]$ and $\delta_\varphi = 0$. Note that $\delta_k(\delta) = \frac{\delta-1}{2}$ has only one solution $\delta = -1$ and

$$\begin{aligned} 2\delta_k - \delta + 1 &> 0 \Leftrightarrow [\delta \leq -1 \text{ and } \delta_k = 0] \text{ or } [\delta \in (-1,0] \text{ and } \delta_k(\delta) \in \{\delta, 0\}], \\ 2\delta_k - \delta + 1 &= 0 \Leftrightarrow [\delta = -1 \text{ and } \delta_k = -1], \\ 2\delta_k - \delta + 1 &< 0 \Leftrightarrow \delta > 0 \text{ or } [\delta \in (-1,0] \text{ and } \delta_k(\delta) = -1]. \end{aligned}$$

Therefore, $\delta_\varphi < 0$ and $\lim_{n\to\infty} \varphi = 0$ if and only if $[\delta \leq -1$ and $\delta_k = 0]$ or $[\delta \in (-1,0]$ and $\delta_k(\delta) \in \{\delta, 0\}]$. For these cases,

$$\delta_\varphi(\delta) = \begin{cases} \delta - 1 & \text{if } \delta \leq -1 \text{ and } \delta_k = 0 \\ \{\delta - 1, -\delta - 1\} & \text{if } \delta \in (-1, 0] \end{cases}.$$

Because $n\varphi$ is of order $\delta_\varphi + 1$, $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta_\varphi(\delta) + 1 > 0 \Leftrightarrow \delta_\varphi(\delta) = -\delta - 1$ for $\delta \in (-1, 0)$. Therefore, $\lim_{n\to\infty} \varphi = 0$ and $\lim_{n\to\infty} n\varphi = \infty$ if and only if $\delta = \delta_k \in (-1, 0)$.

Finally, for $\delta \in (-1, 0)$, consider G. Because $\lim_{n\to\infty} \left(\frac{n-1}{n} - \frac{n+1}{n} \frac{w\varphi}{1-\varphi} \right) = 1$ and

$$\frac{1-c_s}{1+c_s n\varphi} \frac{u}{1+(1-u)\, n\varphi} \sim n^{-2\delta_{n\varphi}},$$

$\delta_G = -(\delta + 2\delta_{n\varphi}) = -(\delta + 2(-\delta)) = \delta < 0$. Therefore, $\lim_{n\to\infty} G = 0$ if $\delta \in (-1, 0)$.

(b) From the analysis of **(a)**, it is easy to verify that for all the cases, any δ below the lower bound leads to $\delta_G > 0$ and hence $\lim_{n\to\infty} G = \infty$.

(c) For all the cases, any δ above the upper bound makes $(1 + (1-u)n)\, k^2 n^{-\delta}$ approach zero as $n \to \infty$. Therefore, $\varphi = \left\{1 + \alpha u \left(1 + (1-u)n\right) k^2 n^{-\delta}\right\}^{-1} \to 1$. The second order condition

$$1 < n \text{ and } \frac{n+1}{n-1} < \frac{1-\varphi}{\varphi} + 1 - w$$

will be violated for some finite n because $\lim_{n\to\infty} \frac{n+1}{n-1} = 1 > \lim_{n\to\infty} \left(\frac{1-\varphi}{\varphi} + 1 - w\right) = 1 - w$. ■

Remark on the assumption about δ.

Because $\delta \equiv \delta_\varepsilon + \delta_x < 0$ (> 0) corresponds to the exogenous assumption that larger n makes the environment more (less) noisy, it is natural to expect with sufficiently small (large) δ, both φ and $n\varphi$ decrease (increase) in n and gains from trade increase (decrease) in n. The analysis in the proof of **Lemma A14** confirms this. What is more interesting in **Lemma A14** is that for each parameter value (t, u, w), there is a non-empty range of δ, for which gains from trade smoothly decrease to zero as n increases to infinity. If this is the case, with the arbitrary small fixed participation cost, the net gains from trade goes to zero for a finite n. In the baseline case with $\delta = 0$, this occurred only for $w = t = u = 1$, but **Lemma A14** shows that the implication obtained from this case is more generally applicable if we relax the value of δ. Also, in the baseline case with $\delta = 0$, extreme negative externalities occurred only for $u < w = t = 1$. **Lemma A14** shows that with sufficiently large δ, φ increases in n and extreme negative externalities arise for the other cases too, as long as $0 < tw$.

Lemma A14 shows that $\delta < 0$ is a necessary condition for $\lim_{n\to\infty} G = \infty$, i.e. for the optimal market size to be infinite.[38] Is $\delta < 0$ a reasonable assumption? We argue that it may not be for the following reason. Our analysis shows that ex ante gains from trade are larger when traders face higher risk. Hence, low values of $(\tau_\varepsilon, \tau_x)$ creates the higher incentive to participate in the market. Suppose traders are heterogenous in $(\tau_\varepsilon, \tau_x)$. Then the traders with low values of $(\tau_\varepsilon, \tau_x)$ have more incentive to participate in the market than those with high values of $(\tau_\varepsilon, \tau_x)$. If everyone faces the same level of participation cost, then the average participating traders should have lower values of $(\tau_\varepsilon, \tau_x)$ than those who do not participate. If the participation cost goes down, the additional traders would *raise* the average value of $(\tau_\varepsilon, \tau_x)$ in the market. Thus, once we consider the ex ante incentive of traders to join the market, it is likely that the average precision $(\tau_\varepsilon, \tau_x)$ in the market increases in n. Therefore, we view $\delta \geq 0$ as a sensible assumption in our symmetric environment. This supports our focus on the case with negative externalities. If δ is strictly increasing in n and $\delta(n)$ exceeds upper bounds identified in **Lemma A14** for sufficiently large n, the market suffers from extreme negative externalities and trading volume and gains from trade drop to zero for sufficiently large but finite n.

7.4.10 Equilibrium with transaction fees

Suppose that traders face the quadratic transaction cost:

$$-\frac{c}{2}q_i^2,\ c > 0$$

on top of the payment (or receipt) of pq_i. With this transaction fee, the first order condition (36) and the second order condition (37) become

$$E_i[v] - \frac{\rho}{\tau}(q_i + e_i) = p_i + 2\lambda_c q_i + cq_i = p + (\lambda_c + c)\, q_i, \tag{70}$$

$$2\lambda_c + c + \frac{\rho}{\tau} > 0, \tag{71}$$

where

$$\lambda_c \equiv \frac{1}{n\beta_p^c}.$$

Thus, we have

$$q_i^c(p; H_i) = \frac{E_i[v] - p - \frac{\rho}{\tau}e_i}{\lambda_c + c + \frac{\rho}{\tau}}. \tag{72}$$

Obviously, all the analysis goes through by replacing λ with $\lambda_c + c$ in (39), (40), (41). Importantly, because the form of transaction cost maintains the linearity of equilibrium, all the informational property of equilibrium is not affected by the value of c.

[38] For (iii) $\delta < -\frac{1}{2}$ is sufficient. For (iv) $\delta < -1$ is sufficient.

Lemma A15 (Transaction fees $\frac{c}{2}q_i^2$)

(a) *The optimal order is $\frac{\rho}{\rho+c\tau}$ times the order in a trade equilibrium without fees.*

The second order condition is same as (30)*.*

(b) *The interim GFT are $G_i^c = \left(1-\widetilde{\lambda}_c\right) G_i^{pt}$, where $\widetilde{\lambda}_c$ is increasing in c and*

$$\begin{aligned}
\widetilde{\lambda}_c &= \left(\frac{c\tau}{\rho+c\tau} + \left(1-\frac{c\tau}{\rho+c\tau}\right)\sqrt{\widetilde{\lambda}}\right)^2 + \frac{c\tau}{\rho+c\tau}\left(1-\frac{c\tau}{\rho+c\tau}\right)\left(1-\sqrt{\widetilde{\lambda}}\right)^2 \in \left[\widetilde{\lambda},1\right], \\
\widetilde{\lambda} &\equiv \left(\frac{\left(w-\frac{1-w}{n}\right)\varphi+\frac{1}{n}}{1-\varphi}\right)^2 \in [0,1].
\end{aligned}$$

(c) *The ex ante GFT are*

$$G^c = \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)\left(1-\frac{c\tau}{\rho+c\tau}\right)A(T;t,w)B(T;u)\right).$$

(d) *Fee revenue is maximized by setting $c(n)=\frac{\rho}{\tau}$. With this fee level,*

$$G^{c(n)} = \frac{1}{2\rho}\log\left(1+\frac{1}{2}\frac{\rho^2}{\tau_v\tau_x}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)A(T;t,w)B(T;u)\right). \tag{73}$$

Proof.
(a) (39), (40), (41) become

$$\beta_s^c = \frac{\tau_\varepsilon}{(\lambda_c+c)\tau+\rho}\frac{1-\varphi}{1+(1-w)(nw-1)\varphi}\frac{\tau}{\tau_1}\sqrt{t},$$

$$\beta_e^c = \frac{\rho}{(\lambda_c+c)\tau+\rho}\left(1-\frac{w\varphi}{1+(1-w)(nw-1)\varphi}\frac{\tau_\varepsilon}{\rho}N_u\frac{\beta_e^c}{\beta_s^c}\frac{\tau}{\tau_1}\sqrt{t}\right),$$

$$\beta_p^c = \frac{\tau}{(\lambda_c+c)\tau+\rho}\left(1-\frac{w\varphi}{1+(1-w)(nw-1)\varphi}\frac{(n+1)\tau_\varepsilon}{\tau_1}\frac{\beta_p^c}{\beta_s^c}\sqrt{t}\right).$$

Solving for $\frac{\beta_p^c}{\beta_s^c}$, we obtain the same expression for $\frac{\beta_p}{\beta_s}$ as before:

$$\frac{\beta_p^c}{\beta_s^c} = \frac{\tau_1}{\sqrt{t}\tau_\varepsilon}\frac{1+(1-w)(nw-1)\varphi}{1+\{wn-(1-w)\}\varphi}.$$

Thus,

$$\begin{aligned}\beta_p^c &= \frac{\tau}{(\lambda_c + c)\tau + \rho}\left(1 - \frac{w\varphi(n+1)}{1 + \{wn - (1-w)\}\varphi}\right) \\ &= \frac{\tau}{\left(\frac{1}{n\beta_p^c} + c\right)\tau + \rho}\left(1 - \frac{w\varphi(n+1)}{1 + \{wn - (1-w)\}\varphi}\right).\end{aligned}$$

Hence,

$$\begin{aligned}\beta_p^c\left(\left(\frac{1}{n\beta_p^c} + c\right)\tau + \rho\right) &= \tau\frac{1 + \{wn - (1-w)\}\varphi - w\varphi(n+1)}{1 + \{wn - (1-w)\}\varphi}. \\ \frac{\tau}{n} + \beta_p^c(c\tau + \rho) &= \tau\frac{1-\varphi}{1 + \{wn - (1-w)\}\varphi}. \\ \beta_p^c &= \frac{\tau}{c\tau + \rho}\left(\frac{1-\varphi}{1 + \{wn - (1-w)\}\varphi} - \frac{1}{n}\right) \\ &= \frac{\rho}{c\tau + \rho}\frac{\tau}{\rho}\frac{\frac{n-1}{n} - \varphi\left(1 + w - \frac{1-w}{n}\right)}{1 + \{wn - (1-w)\}\varphi} \\ &= \frac{\rho}{c\tau + \rho}\beta_p.\end{aligned}$$

Because $\frac{\beta_p^c}{\beta_s^c} = \frac{\beta_p}{\beta_s}$ and $\frac{\beta_p^c}{\beta_e^c} = \frac{\beta_p}{\beta_e}$,

$$\begin{aligned}\beta_s^c &= \frac{\beta_p^c}{\beta_p}\beta_s = \frac{\rho}{c\tau + \rho}\beta_s, \\ \beta_e^c &= \frac{\beta_p^c}{\beta_p}\beta_e = \frac{\rho}{c\tau + \rho}\beta_e.\end{aligned}$$

Since $\frac{\beta_e^c}{\beta_s^c} = \frac{\beta_e}{\beta_s}$, all the informational property of equilibrium remains the same.

From the second order condition (71),

$$0 < 2\lambda_c + c + \frac{\rho}{\tau} \Leftrightarrow 0 < 1 + \frac{\tau}{\rho}(2\lambda_c + c).$$

$$\begin{aligned}
1+\frac{\tau}{\rho}\left(2\lambda_c+c\right) &= 1+\frac{\tau}{\rho}c+2\frac{\tau}{\rho}\frac{1}{n\beta_p^c} \\
&= 1+\frac{\tau}{\rho}c+2\frac{\tau}{\rho}\frac{1}{n\frac{\rho}{c\tau+\rho}\beta_p} \\
&= 1+\frac{\tau}{\rho}c+2\frac{c\tau+\rho}{\rho}\frac{\tau}{\rho}\frac{1}{n}\frac{\rho}{\tau}\frac{1+\{wn-(1-w)\}\varphi}{\frac{n-1}{n}-\varphi\left(1+w-\frac{1-w}{n}\right)} \\
&= 1+\frac{\tau}{\rho}c+2\left(1+\frac{\tau}{\rho}c\right)\frac{1+\{wn-(1-w)\}\varphi}{n-1-n\varphi\left(1+w-\frac{1-w}{n}\right)} \\
&= \left(1+\frac{\tau}{\rho}c\right)\left\{1+2\frac{1+\{wn-(1-w)\}\varphi}{n-1-n\varphi\left(1+w-\frac{1-w}{n}\right)}\right\} \\
&= \left(1+\frac{\tau}{\rho}c\right)\frac{(n+1)\{1-\varphi(1-w)\}}{n-1-n\varphi\left(1+w-\frac{1-w}{n}\right)}.
\end{aligned}$$

This is positive if and only if $n-1-n\varphi\left(1+w-\frac{1-w}{n}\right)>0$, which yields the same condition as before.

(b) Interim profit is

$$E_i[v]\left(q_i+e_i\right)-\frac{\rho}{2}Var_i\left[v\right]\left(q_i+e_i\right)^2-pq_i-\frac{c}{2}q_i^2.$$

By substituting (72) into this, obtain

$$\begin{aligned}
\Pi_i^c &= \left(1-\left(\frac{\left(\lambda_c+c\right)\tau}{\rho+\left(\lambda_c+c\right)\tau}\right)^2\right)\left(\frac{\tau}{2\rho}\left(E_i[v]-p\right)^2+pe_i\right) \qquad (74) \\
&\quad+\left(\frac{\left(\lambda_c+c\right)\tau}{\rho+\left(\lambda_c+c\right)\tau}\right)^2\left(E_i[v]e_i-\frac{\rho}{2\tau}e_i^2\right)-\frac{c}{2}\left\{\frac{\tau}{\rho+\left(\lambda_c+c\right)\tau}\left(E_i[v]-p-\frac{\rho}{\tau}e_i\right)\right\}^2.
\end{aligned}$$

Note that $\Pi_i^{c=0}=\Pi_i$ because the first two terms of (74) are the same with (47) when $\lambda_c+c=\lambda$. Computing $\left(E_i[v]-p-\frac{\rho}{\tau}e_i\right)^2$ in the third term yields

$$\begin{aligned}
\left(E_i[v]-p-\frac{\rho}{\tau}e_i\right)^2 &= \left(E_i[v]-p\right)^2-\frac{2\rho}{\tau}\left(E_i[v]-p\right)e_i+\left(\frac{\rho}{\tau}\right)^2e_i^2 \\
&= \frac{2\rho}{\tau}\left\{\frac{\tau}{2\rho}\left(E_i[v]-p\right)^2+pe_i-\left(E_i[v]e_i-\frac{\rho}{2\tau}e_i^2\right)\right\}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\Pi_i^c &= \left(1-\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2\right)\left(\frac{\tau}{2\rho}\left(E_i[v]-p\right)^2+pe_i\right) \\
&\quad+\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2\left(E_i[v]e_i-\frac{\rho}{2\tau}e_i^2\right) \\
&\quad-\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2}\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2\left\{\left(\frac{\tau}{2\rho}\left(E_i[v]-p\right)^2+pe_i\right)-\left(E_i[v]e_i-\frac{\rho}{2\tau}e_i^2\right)\right\} \\
&= \left\{1-\left(1+\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2}\right)\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2\right\}\Pi_i^{pt} \\
&\quad+\left(1+\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2}\right)\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2\Pi_i^{nt}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
G_i^c &\equiv \Pi_i^c-\Pi_i^{nt} \\
&= \left(1-\widetilde{\lambda}_c\right)G_i^{pt},
\end{aligned}$$

where

$$\widetilde{\lambda}_c \equiv \left(1+\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2}\right)\left(\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau}\right)^2.$$

Next, we evaluate $\widetilde{\lambda}_c$. First, use $\lambda_c=\frac{1}{n\beta_p^c}=\frac{\rho+c\tau}{\rho}\frac{1}{n\beta_p}=\left(1+\frac{c\tau}{\rho}\right)\lambda$ to obtain

$$\begin{aligned}
\frac{(\lambda_c+c)\,\tau}{\rho+(\lambda_c+c)\,\tau} &= \frac{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda+c\right\}\tau}{\rho+\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda+c\right\}\tau} \\
&= \frac{\tau\lambda+c\tau\left(1+\frac{\tau}{\rho}\lambda\right)}{\rho+\tau\lambda+c\tau\left(1+\frac{\tau}{\rho}\lambda\right)} \\
&= \frac{\frac{\tau}{\rho}\lambda+c\frac{\tau}{\rho}\left(1+\frac{\tau}{\rho}\lambda\right)}{\left(1+c\frac{\tau}{\rho}\right)\left(1+\frac{\tau}{\rho}\lambda\right)} \\
&= \frac{c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}}+\frac{1}{1+c\frac{\tau}{\rho}}\frac{\lambda\tau}{\rho+\lambda\tau} \\
&= \frac{\frac{\lambda\tau}{\rho+\lambda\tau}+c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}} \geq \frac{\lambda\tau}{\rho+\lambda\tau},
\end{aligned}$$

with equality if $c=0$. This proves $\widetilde{\lambda}_c=\left(1+\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2}\right)\left(\frac{(\lambda_c+c)\tau}{\rho+(\lambda_c+c)\tau}\right)^2 \geq \left(\frac{\lambda\tau}{\rho+\lambda\tau}\right)^2=\widetilde{\lambda}$.

Next,

$$
\begin{aligned}
1+\frac{\rho}{\tau}\frac{c}{(\lambda_c+c)^2} &= 1+\frac{\rho}{\tau}\frac{c}{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda+c\right\}^2} \\
&= \frac{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda+c\right\}^2+c\frac{\rho}{\tau}}{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda+c\right\}^2} \\
&= \frac{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda\frac{\tau}{\rho}+c\frac{\tau}{\rho}\right\}^2+c\frac{\tau}{\rho}}{\left\{\left(1+\frac{c\tau}{\rho}\right)\lambda\frac{\tau}{\rho}+c\frac{\tau}{\rho}\right\}^2} \\
&= \frac{\left\{\lambda\frac{\tau}{\rho}+c\frac{\tau}{\rho}\left(\lambda\frac{\tau}{\rho}+1\right)\right\}^2+c\frac{\tau}{\rho}}{\left\{\lambda\frac{\tau}{\rho}+c\frac{\tau}{\rho}\left(\lambda\frac{\tau}{\rho}+1\right)\right\}^2} \\
&= \frac{\left\{\lambda\tau+c\frac{\tau}{\rho}(\lambda\tau+\rho)\right\}^2+c\tau\rho}{\left\{\lambda\tau+c\frac{\tau}{\rho}(\lambda\tau+\rho)\right\}^2} \\
&= \frac{\left(\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}\right)^2+c\frac{\tau}{\rho}\left(\frac{\rho}{\lambda\tau+\rho}\right)^2}{\left(\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}\right)^2}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\widetilde{\lambda}_c &= \frac{\left(\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}\right)^2+c\frac{\tau}{\rho}\left(\frac{\rho}{\lambda\tau+\rho}\right)^2}{\left(\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}\right)^2}\left(\frac{\frac{\lambda\tau}{\rho+\lambda\tau}+c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}}\right)^2 \\
&= \frac{\left(\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}\right)^2+c\frac{\tau}{\rho}\left(1-\frac{\lambda\tau}{\lambda\tau+\rho}\right)^2}{\left(1+c\frac{\tau}{\rho}\right)^2} \\
&= \left(\frac{\frac{\lambda\tau}{\lambda\tau+\rho}+c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}}\right)^2+\frac{c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}}\frac{1}{1+c\frac{\tau}{\rho}}\left(1-\frac{\lambda\tau}{\lambda\tau+\rho}\right)^2.
\end{aligned}
$$

From **Lemma A4**,

$$
\sqrt{\widetilde{\lambda}}=\frac{\lambda\tau}{\rho+\lambda\tau}=\frac{\left(w-\frac{1-w}{n}\right)\varphi+\frac{1}{n}}{1-\varphi}.
$$

This proves (b).

(c) It suffices to show that $1-\widetilde{\lambda}_c=\left(1-\frac{c\tau}{\rho+c\tau}\right)\left(1-\widetilde{\lambda}\right)$. Let $\widetilde{c}\equiv\frac{c\frac{\tau}{\rho}}{1+c\frac{\tau}{\rho}}=\frac{c\tau}{\rho+c\tau}$. Then

$$
\widetilde{\lambda}_c=\left(\widetilde{c}+(1-\widetilde{c})\sqrt{\widetilde{\lambda}}\right)^2+\widetilde{c}(1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)^2
$$

and

$$
\begin{aligned}
1-\widetilde{\lambda}_c &= \left\{1-\left(\widetilde{c}+(1-\widetilde{c})\sqrt{\widetilde{\lambda}}\right)\right\}\left(1+\widetilde{c}+(1-\widetilde{c})\sqrt{\widetilde{\lambda}}\right)-\widetilde{c}(1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)^2 \\
&= (1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)\left(1+\widetilde{c}+(1-\widetilde{c})\sqrt{\widetilde{\lambda}}\right)-\widetilde{c}(1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)^2 \\
&= (1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)\left\{1+\widetilde{c}+(1-\widetilde{c})\sqrt{\widetilde{\lambda}}-\widetilde{c}\left(1-\sqrt{\widetilde{\lambda}}\right)\right\} \\
&= (1-\widetilde{c})\left(1-\sqrt{\widetilde{\lambda}}\right)\left(1+\sqrt{\widetilde{\lambda}}\right) \\
&= (1-\widetilde{c})\left(1-\widetilde{\lambda}\right).
\end{aligned}
$$

(d) The quantity traded is given by

$$
q_i^c = \frac{\frac{\rho}{\rho+c\tau}\sqrt{t}\left\{\frac{n-1}{n}-\left(1+w-\frac{1-w}{n}\right)\varphi\right\}}{1+\varphi(1-w)(nw-1)+(1-t)\frac{\tau_\varepsilon}{\tau_v}\left\{1+\varphi(nw-(1-w))\right\}}\left\{\frac{\tau_\varepsilon}{\rho}(s_i-\overline{s})-k(e_i-\overline{e})\right\}.
$$

Therefore, taking n as given, the fee revenue $\frac{c}{2}\sum_{i=1}^{n+1}\left(q_i^c\right)^2$ is maximized by setting c equals to

$$
\arg\max_c \frac{c}{(\rho+c\tau)^2}.
$$

The first order condition is

$$
\frac{\rho+c\tau-2c\tau}{(\rho+c\tau)^3}=0 \Leftrightarrow c=\frac{\rho}{\tau}.
$$

The second order condition is obviously satisfied. With this fee level, $1-\frac{c\tau}{\rho+c\tau}=\frac{1}{2}$. ■

Remark on the effect of transaction fees on gains from trade.

The impact of transaction fees in the form of quadratic volume tax $(\frac{c}{2}q_i^2)$ appears in the expression of gains from trade in a multiplicative way:

$$
G^c = \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\left(\frac{n-1}{n}-\frac{n+1}{n}\frac{w\varphi}{1-\varphi}\right)\left(1-\frac{c\tau}{\rho+c\tau}\right)A(T;t,w)B(T;u)\right).
$$

Note that $1-\frac{c\tau}{\rho+c\tau}=\frac{\rho}{\rho+c\tau}$ depends on n. If τ is increasing in n, the discount of gains from trade due to transaction fees is bigger for the larger market. This means that the optimal market size becomes smaller relative to the baseline case without transaction fees. However, **Lemma A15 (d)** shows that at the level of c that maximizes the fee revenue, $\frac{\rho}{\rho+c\tau}=\frac{1}{2}$. Therefore, if intermediaries choose c to maximize profits after observing n, the relevant ex ante GFT is (73). Thus, GFT is reduced to half for each n, and all the analysis for the optimal market size remains unchanged.

7.4.11 Sequentially opening markets

Suppose there are two markets and they open sequentially. Traders in the second market observe a market-clearing price in the first market before they trade. This implies that there is additional public information in the second market and the first market imposes informational externality on the second market. We characterizes equilibrium and gains from trade in the second market assuming $t = u = w = 1$. So **Lemma A2** applies to the market 1. Suppose that the first market has n_1+1 traders and the second market has n_2+1 traders. Each trader in the second market has the information set $H_{i2} = \{s_i, e_i, p_1, p\}$, where p_1 is the publicly observed market-clearing price in the first market and p is the equilibrium price in the second market to be determined. Ex ante gains from trade for the second market $G_2(n_2; n_1)$ depends on n_1. We define the optimal second market size for given n_1 by

$$n_2^*(n_1) \equiv \arg\max_{n_2} G_2(n_2; n_1).$$

Recall that the second order condition

$$\frac{n_1+1}{n_1-1} < \frac{\rho^2}{\tau_\varepsilon \tau_x} = \frac{1-\varphi}{\varphi}$$

imposes the lower bound on the market size:

$$\underline{n} \equiv \frac{1}{1-2\varphi} < n_1.$$

Using $\underline{n}$, the optimal size for the first market (see **Lemma A2**) is

$$n_1^* = \underline{n}\left(1+\sqrt{1+\frac{1}{\underline{n}c_s\varphi}}\right).$$

Lemma A16 (Sequentially opening markets)

Assume $t = u = w = 1$.

(a) *Two markets can coexist only if*

$$\frac{n_1+n_2+2}{n_1+n_2-2} < \frac{\rho^2}{\tau_\varepsilon \tau_x}.$$

In the second market,

$$\begin{aligned} q_{i2}(p;H_{i2}) &= \left(\frac{n_2-1}{n_2}-2\varphi\right)\left\{\frac{\tau_\varepsilon}{\rho}s_i - e_i - \frac{\tau_2}{\rho}\frac{1}{1+n_2\varphi}p + \frac{\tau_1}{\rho}\frac{(1+n_1)\varphi}{(1+n_1\varphi)(1+n_2\varphi)}p_1\right\}, \\ \tau_1 &= \tau_v + \tau_\varepsilon(1+n_1\varphi), \\ \tau_2 &= \tau_v + \tau_\varepsilon\{1+(1+n_1+n_2)\varphi\}, \\ q_{i2}(p^*;H_{i2}) &= \left(\frac{n_2-1}{n_2}-2\varphi\right)\left\{\frac{\tau_\varepsilon}{\rho}(s_i-\overline{s})-(e_i-\overline{e})\right\}. \end{aligned}$$

(b) *Ex ante GFT and the optimal market size are*

$$\begin{aligned} G_2(n_2;n_1) &= \frac{1}{2\rho}\log\left(1+\frac{\rho^2}{\tau_v\tau_x}\frac{\frac{n_2-1}{n_2}-2\varphi}{1-\varphi}\frac{1-c_s}{1+c_s(1+n_1+n_2)\varphi}\right), \\ n_2^*(n_1) &= \underline{n}\left(1+\sqrt{1+\frac{1}{\underline{n}c_s\varphi}+\frac{1+n_1}{\underline{n}}}\right). \end{aligned}$$

(c) $G_2(n_1;n_1) < G_1(n_1)$ *and* $n_2^*(n_1) > n_1^*$ *for all* $n_1 > \underline{n}$.

(d) *There exists unique* $\widehat{n}_1 \in (n_1^*, \infty)$ *defined by* $n_2^*(\widehat{n}_1) = \widehat{n}_1$ *such that*

$$n_1 \lesseqgtr n_2^*(n_1) \Leftrightarrow n_1 \lesseqgtr \widehat{n}_1.$$

(e) *There exists unique* $\widetilde{n} \in (\underline{n}, n_1^*)$ *defined by* $G_2(n_2^*(\widetilde{n});\widetilde{n}) = G_1(\widetilde{n})$ *such that*

(i) *if* $n_1 > \widetilde{n}$, *then* $G_2(n_2;n_1) < G_1(n_1)$ *for all* $n_2 > \underline{n}$,

(ii) *if* $n_1 \in (\underline{n}, \widetilde{n})$, *then* $G_2(n_2;n_1) = G_1(n_1)$ *has two solutions* $n_2^{\pm}$ *and*

$$G_2(n_2;n_1) > G_1(n_1) \Leftrightarrow n_2 \in \left(n_2^-, n_2^+\right).$$

Proof.
(a) From **Lemma A2**,

$$p_1 = \frac{\tau_\varepsilon(1+n_1\varphi)}{\tau_v+\tau_\varepsilon(1+n_1\varphi)}\left(\overline{s}_1 - \frac{\rho}{\tau_\varepsilon}\overline{e}_1\right),\ \varphi = \left(1+\frac{\rho^2}{\tau_\varepsilon\tau_x}\right)^{-1},$$

where $\overline{s}_1$ and $\overline{e}_1$ are the average of signals and endowment among $1+n_1$ traders in the first market. Using $c_s = \frac{\tau_\varepsilon}{\tau_v+\tau_\varepsilon}$, this can be written as

$$p_1 = \frac{c_s\left(1+n_1\varphi\right)}{1+c_s n_1\varphi}\left(v+\overline{\varepsilon}_1-\frac{\rho}{\tau_\varepsilon}\overline{e}_1\right).$$

This is informationally equivalent to a public signal

$$z_1 \equiv \frac{1+c_s n_1\varphi}{c_s\left(1+n_1\varphi\right)}p_1 = v+\overline{\varepsilon}_1-\frac{\rho}{\tau_\varepsilon}\overline{e}_1.$$

The informational value of this signal is measured by

$$\begin{aligned}\left(Var\left[z_1|v\right]\right)^{-1} &= \left(1+n_1\right)\left(\frac{1}{\tau_\varepsilon}+\left(\frac{\rho}{\tau_\varepsilon}\right)^2\frac{1}{\tau_x}\right)^{-1} \\ &= \left(1+n_1\right)\tau_\varepsilon\varphi.\end{aligned}$$

Given a conjecture

$$q_{i2}\left(p;H_{i2}\right) = \beta_{s2}s_i-\beta_{e2}e_i+\beta_1 p_1-\beta_{p2}p,$$

the market-clearing condition satisfies

$$q_{i2}+\sum_{j\neq i}\left(\beta_{s2}s_j-\beta_{e2}e_j+\beta_1 p_1\right) = n_2\beta_{p2}p.$$

Hence, trader i constructs a signal

$$\begin{aligned}h_{i2} &\equiv \frac{1}{\beta_{s2}}\left(\beta_{p2}p-\beta_1 p_1-\frac{q_{i2}}{n_2}\right) \\ &= v+\overline{\varepsilon}_i-\frac{\beta_{e2}}{\beta_{s2}}\overline{e}_i,\end{aligned}$$

where $\overline{\varepsilon}_i$ and $\overline{e}_i$ are the average among n_2 traders other than i. The informational value of this signal is measured by

$$\begin{aligned}\left(Var\left[h_{i2}|v\right]\right)^{-1} &= n_2\left(\frac{1}{\tau_\varepsilon}+\left(\frac{\beta_{e2}}{\beta_{s2}}\right)^2\frac{1}{\tau_x}\right)^{-1} \\ &= n_2\tau_\varepsilon\varphi_2,\end{aligned}$$

where $\varphi_2 \equiv \left(\frac{1}{\tau_\varepsilon}+\left(\frac{\beta_{e2}}{\beta_{s2}}\right)^2\frac{1}{\tau_x}\right)^{-1}$. Therefore, in the second market, each trader i has three signals about v:

$$\begin{aligned} s_i &= v + \varepsilon_i, \\ z_1 &= v + \overline{\varepsilon}_1 - \frac{\rho}{\tau_\varepsilon}\overline{e}_1, \\ h_{i2} &= v + \overline{\varepsilon}_i - \frac{\beta_{e2}}{\beta_{s2}}\overline{e}_i. \end{aligned}$$

Note that $[v, s_i, e_i, z_1, h_{i2}]^\top$ is jointly normal with mean zero and a variance-covariance matrix

$$\begin{bmatrix} \frac{1}{\tau_v} & \frac{1}{\tau_v} & 0 & \frac{1}{\tau_v} & Cov[v, h_{i2}] \\ & \frac{1}{\tau_v} + \frac{1}{\tau_\varepsilon} & 0 & \frac{1}{\tau_v} & Cov[v, h_{i2}] \\ & & \frac{1}{\tau_x} & 0 & Cov[v, h_{i2}] \\ & & & \frac{1}{\tau_v} + \frac{1}{(1+n_1)\tau_\varepsilon \varphi} & Cov[v, h_{i2}] \\ & & & & Var[h_{i2}] \end{bmatrix}.$$

The rest of the analysis follows the proof of **Lemma A3**. In particular, because noise in signals are mutually independent,

$$\begin{aligned} \tau_2 &\equiv (Var[v|H_{i2}]) \\ &= \tau_v + \tau_\varepsilon \{1 + (1 + n_1)\varphi + n_2\varphi_2\}. \end{aligned}$$

It is easy to verify that

$$\frac{\beta_{e2}}{\beta_{s2}} = \frac{\beta_e}{\beta_s} = \frac{\rho}{\tau_\varepsilon} \text{ and } \varphi_2 = \varphi.$$

Using $\lambda_2 = \frac{1}{n_2\beta_{p2}}$, the second order condition is satisfied if and only if $2\lambda_2 + \frac{\rho}{\tau_2} > 0 \Leftrightarrow 0 < 1 + \frac{\tau_2}{\rho}\frac{2}{n_2\beta_p}$. Evaluating this shows

$$1 < n_2 \text{ and } \frac{n_2 + 1}{n_2 - 1} < \frac{1 - \varphi}{\varphi}.$$

This can be written as

$$\frac{1}{1 - 2\varphi} < n_2.$$

Because $\frac{1}{1-2\varphi} < n_1$ is necessary for the first market, two markets can coexist only if

$$\begin{aligned}\frac{2}{1-2\varphi} &< n_1+n_2 \\ &\Leftrightarrow 2 < n_1+n_2-2\left(n_1+n_2\right)\varphi \\ &\Leftrightarrow 2\left(n_1+n_2\right)\varphi < n_1+n_2-2 \\ &\Leftrightarrow \varphi < \frac{n_1+n_2-2}{2\left(n_1+n_2\right)} \\ &\Leftrightarrow \frac{2\left(n_1+n_2\right)}{n_1+n_2-2} < \frac{1}{\varphi} \\ &\Leftrightarrow \frac{2\left(n_1+n_2\right)}{n_1+n_2-2}-1 < \frac{1-\varphi}{\varphi} \\ &\Leftrightarrow \frac{n_1+n_2+2}{n_1+n_2-2} < \frac{\rho^2}{\tau_\varepsilon \tau_x}.\end{aligned}$$

(b) The proof follows that of **Lemma A4** and **A5** where $\widetilde{\lambda}$ is replaced with

$$\widetilde{\lambda}_2 \equiv \left(\frac{\lambda_2\tau_2}{\rho+\lambda_2\tau_2}\right)^2 = \left(\frac{1+n_2\varphi}{n_2\left(1-\varphi\right)}\right)^2.$$

Using

$$1-\widetilde{\lambda}_2 = \frac{1}{\left(1-\varphi\right)^2}\frac{1+n_2}{n_2}\left(\frac{n_2-1}{n_2}-2\varphi\right)$$

and

$$\begin{aligned}\frac{\tau_v}{\tau_2} &= \frac{\tau_v}{\tau_v+\tau_\varepsilon\left\{1+\left(1+n_1+n_2\right)\varphi\right\}} \\ &= \frac{1-c_s}{1+c_s\left(1+n_1+n_2\right)\varphi}\end{aligned}$$

yields the expression of $G_2\left(n_2;n_1\right)$. Clearly,

$$\begin{aligned}n_2^*\left(n_1\right) &\equiv \arg\max_{n_2} G_2\left(n_2;n_1\right) \\ &= \arg\max_{n_2}\frac{\frac{n_2-1}{n_2}-2\varphi}{1+c_s\left(1+n_1+n_2\right)\varphi}.\end{aligned}$$

The first order condition for this is

$$\frac{1}{n_2^2}\left\{1+c_s\left(1+n_1+n_2\right)\varphi\right\}-\left(\frac{n_2-1}{n_2}-2\varphi\right)c_s\varphi=0.$$

This defines

$$\left(1-2\varphi\right)n_2^2-2n_2-\left(\frac{1}{c_s\varphi}+1+n_1\right)=0.$$

This has two solutions and only the larger one satisfies the second order condition. Hence,

$$\begin{aligned}
n_2^*(n_1) &= \frac{1+\sqrt{1+(1-2\varphi)\left(\frac{1}{c_s\varphi}+1+n_1\right)}}{1-2\varphi} \\
&= \underline{n}\left(1+\sqrt{1+\frac{1}{\underline{n}c_s\varphi}+\frac{1+n_1}{\underline{n}}}\right).
\end{aligned}$$

(c) Obvious from the expression of $G_2(n_2;n_1)$ and $n_2^*(n_1)$.
(d) First,

$$\begin{aligned}
\frac{n_2^*(n_1)}{n_1} &= \frac{\underline{n}}{n_1}\left(1+\sqrt{1+\frac{1}{\underline{n}c_s\varphi}+\frac{1+n_1}{\underline{n}}}\right) \\
&= \frac{\underline{n}}{n_1}\left(1+\sqrt{\frac{1+c_s\varphi(1+\underline{n})}{\underline{n}c_s\varphi}+\frac{n_1}{\underline{n}}}\right).
\end{aligned}$$

Let $x \equiv \frac{n_1}{\underline{n}} > 1$. Then

$$\begin{aligned}
\frac{n_2^*(n_1)}{n_1} &\gtreqless 1 \Leftrightarrow x \lesseqgtr 1+\sqrt{\frac{1+c_s\varphi(1+\underline{n})}{\underline{n}c_s\varphi}+x} \\
&\Leftrightarrow (x-1)^2 \lesseqgtr \frac{1+c_s\varphi(1+\underline{n})}{\underline{n}c_s\varphi}+x \\
&\Leftrightarrow x^2-3x-\frac{1+c_s\varphi}{\underline{n}c_s\varphi} \lesseqgtr 0.
\end{aligned}$$

The equation defined by $\frac{n_2^*(n_1)}{n_1}=1$ has only one positive solution

$$x^+ = \frac{3}{2}+\sqrt{\left(\frac{3}{2}\right)^2+\frac{1+c_s\varphi}{\underline{n}c_s\varphi}} > 3.$$

Letting $\widehat{n} \equiv \underline{n}x^+$,

$$n_2^*(n_1) \gtreqless n_1 \Leftrightarrow n_1 \lesseqgtr \widehat{n}.$$

From the expression of n_1^* and $\widehat{n}$, it is clear that $n_1^* < \widehat{n}$.
(e) From the expression of $G_1(n_1)$ and $G_2(n_2;n_1)$,

$$G_1(n_1) \gtreqless G_2(n_2;n_1) \Leftrightarrow \frac{\frac{n_1-1}{n_1}-2\varphi}{1+c_s n_1\varphi} \gtreqless \frac{\frac{n_2-1}{n_2}-2\varphi}{1+c_s(1+n_1+n_2)\varphi}.$$

The latter is equivalent to

$$\frac{n_2}{n_1}\left\{(1-2\varphi)\,n_1-1\right\} \quad \gtreqless \quad \left\{(1-2\varphi)\,n_2-1\right\}\frac{1+c_s n_1\varphi}{1+c_s\left(1+n_1+n_2\right)\varphi}$$
$$\Leftrightarrow \quad \frac{n_2}{n_1} \gtreqless \frac{n_2-\underline{n}}{n_1-\underline{n}}\frac{c_s\varphi+\frac{1}{n_1}}{c_s\varphi+\frac{1}{n_1}+c_s\varphi\frac{1+n_2}{n_1}}.$$

Note that given both n_1 and n_2 are greater than $\underline{n}$, $\frac{n_2}{n_1}\leq 1$ implies $\frac{n_2-\underline{n}}{n_1-\underline{n}}\leq 1$. Combined with $\frac{c_s\varphi+\frac{1}{n_1}}{c_s\varphi+\frac{1}{n_1}+c_s\varphi\frac{1+n_2}{n_1}}<1$, this implies that $n_2\leq n_1 \Rightarrow G_1\left(n_1\right)>G_2\left(n_2;n_1\right)$. From (d), if $n_1\geq\widehat{n}$, then $n_2^*\left(n_1\right)\leq n_1$. Therefore, to have $G_1\left(n_1\right)=G_2\left(n_2^*\left(n_1\right);n_1\right)$, it must be that $n_1<\widehat{n}$.

If $n_1=n_1^*$, $G_1\left(n_1^*\right)>G_1\left(n_2^*\left(n_1^*\right)\right)>G_2\left(n_2^*\left(n_1^*\right);n_1^*\right)$. For $n_1\in\left[n_1^*,\widehat{n}\right)$, $n_1<n_2^*\left(n_1\right)$. Hence, $G_1\left(n_1\right)>G_1\left(n_2^*\left(n_1\right)\right)>G_2\left(n_2^*\left(n_1\right);n_1\right)$ for all $n_1\in\left[n_1^*,\widehat{n}\right)$. In particular, at $n_1=n_1^*$ it must be that

$$\frac{n_1^*-\underline{n}}{n_1^*}>\frac{n_2^*\left(n_1^*\right)-\underline{n}}{n_2^*\left(n_1^*\right)}\frac{c_s\varphi+\frac{1}{n_1^*}}{c_s\varphi+\frac{1}{n_1^*}+c_s\varphi\frac{1+n_2^*\left(n_1^*\right)}{n_1^*}}.$$

For $n_1\in\left(\underline{n},n_1^*\right)$, consider

$$\frac{n_1-\underline{n}}{n_1}=\frac{n_2^*\left(n_1\right)-\underline{n}}{n_2^*\left(n_1\right)}\frac{c_s\varphi+\frac{1}{n_1}}{c_s\varphi+\frac{1}{n_1}+c_s\varphi\frac{1+n_2^*\left(n_1\right)}{n_1}}.$$

The left hand side is increasing in n_1 from zero to $\frac{n_1^*-\underline{n}}{n_1^*}$. The right hand side is also increasing but from strictly positive value to the value strictly smaller than $\frac{n_1^*-\underline{n}}{n_1^*}$. Therefore, there is $\widetilde{n}\in\left(\underline{n},n_1^*\right)$ that satisfies $G_2\left(n_2^*\left(\widetilde{n}\right);\widetilde{n}\right)=G_1\left(\widetilde{n}\right)$. For the uniqueness of $\widetilde{n}$, notice that $G_1\left(n_1\right)$ is strictly increasing for $n_1\in\left(\underline{n},n_1^*\right)$ while $G_2\left(n_2^*\left(n_1\right);n_1\right)$ is strictly decreasing. ■