Cuban Oranges and Rotten Cucumbers: Information Revelation and Bundling

Tomasz Sadzik* Chris Woolnough†

March 13th, 2023

Abstract

The paper examines information revelation and incentives to complete markets, in a model of strategic trade. It provides a tight condition on information structures ("*competitive information*") so that prices aggregate information about the assets distributed among the traders. It also tightly characterizes environments in which competition does not provide incentives for strategic traders to issue assets that complete markets ("*Cuban oranges and rotten cucumbers*"). Towards those goals, in our model of strategic trade with multiple assets, factors, and asymmetrically distributed information, the paper establishes uniqueness and provides a closed form characterization of the linear equilibrium.

Keywords: Information aggregation, Endogenously incomplete markets, Market microstructure.

*Department of Economics, UCLA. E-mail: tsadzik@econ.ucla.edu
†E-mail: cjw339@nyu.edu

1 Introduction

In facilitating the allocation of resources, asset markets rely on their ability to efficiently broker and communicate the information distributed in the economy. One mechanism driving information revelation in those markets is the information-based trade, with arbitrage pushing asset prices to reflect the underlying values. The second mechanism is financial innovation, guaranteeing that the market has sufficiently many assets, for the prices to reflect well the underlying factors. This paper investigates how the two mechanisms interact, in a strategic setting. In particular, we study under what conditions strategic traders (nearly) reveal information about the asset prices, as well as whether they have sufficient incentives to introduce rich set of assets to complete markets.

We consider a model of asset trade, in which some agents are asymmetrically informed about multiple aspects of the economy, or factors. They compete with each others and pure liquidity traders by trading simultaneously multiple assets. Our first main result provides a tight (sufficient and necessary) condition on the information structures so that, for a fixed set of assets, prices approximate asset values in the equilibrium of the trading game. The condition ("*competitive information*") formalizes the notion of competitive information and requires that no finite set of agents monopolizes any amount of information in the economy.

Our second main result provides a tight condition on the information structures so that spanning assets do not maximize the informed traders payoffs. Consequently, competition in financial innovation does not provide incentives for the traders to issue assets that complete the market. The condition of "*Cuban oranges and rotten cucumbers*" (CORC), explained in detail below, is satisfied when the liquidity demand and the amount of private information departs from an assortative order across factors—some factors in high demand, and others with informed traders enjoying high informational advantage.

The two main economic insights of the paper give differing answers about the extent to which competition among strategic traders leads to informational efficiency. The first result gives an overall positive answer, in line with the insights of the past literature and the high level intuition on the force of competition. The contribution is a tight and abstract, yet intuitive condition. In particular, surprisingly, the condition makes no reference to the higher order beliefs about the information of others, often implicitly in agreement in "replica economy" constructions. Intuitively, if a small group of agents monopolises some information—as measured by their contribution to the posterior variance of a joint estimate—then due to their strategic impact on price, each shades own trade. The shading limits how much of their information flows into the price. On the other hand, no information revelation would require each of many agents to limit their trade on the joint piece of information, implying rents for each, and unbounded profits overall.

The bulk of research on strategic information revelation is in the setting of a single asset and one factor, information about which is distributed among many agents. Our framework, with strategic trade over many assets and information about many factors, allows us to address the second main economic point, which is, broadly speaking, how informed payoffs respond to different structures of assets.

The second main economic result is to point out sharp limits on incentives to complete markets, in strategic trade. The argument is novel and goes as follows. The profits that the informed traders can make off a trade of a given factor are proportional to both the strength of the (liquidity) demand for the factor, as well as the distance between own informed expectation and the market price. High amount of information means high "profit margin" per trade. High demand means that large trade can be executed with little adverse effect on the price. When the two parameters are not assortative across the factors, informed traders might benefit from bundling a highly demanded asset with the one with high "profit margin". In an informationally complex economy, the chances of all factors ordered assortatively seem low.

On the light note, the rationale for bundling assets is reminiscent of the story of trading produce in the Eastern Block, eponymous for our condition. Around Christmas time, long lines would form for highly the coveted Cuban oranges (at state regulated, low profit margins). At the same time, the stores would want to unload the undemanded cucumbers rotting on the shelves. Ingenious salespeople understood how to benefit from bundling.

The argument behind the benefit of bundling shows not just the limits of, but a detriment of financial innovation. Some of the most successful examples of financial innovation, such as asset-backed securitization or tranching, have been motivated as ways to improve hedging and risk-sharing of the investors. While this is undoubtedly a dominant force, our argument points to the risk of diminishing informativeness and rent extraction. This seems especially pertinent for the Credit Default Swaps that bundle the highly demanded transparent, well rated bonds with the ones characterized by highly asymmetric information.

The technical innovations that underpin our economic results are as follows. With the analysis of strategic trade in divisible setting notoriously difficult, looking into market completeness requires going beyond single asset and factor setting. The model we propose is a version of the static Kyle (1985), but allowing for many payoff relevant factors and many assets, linear in factors. The signals are asymmetric and allow for a broad range of informational patterns, both about the fundamentals and signals of others, with the restriction that all the signals are Normal. Despite this multidimensional and asymmetric setting, we derive a closed formula characterizing the unique linear equilibrium. In particular, the matrix of price impacts ("*Kyle lambda*"), or market learning parameters admits a tractable characterization despite the multidimensional learning with endogenous asymmetric signals (strategies).

The literature on information revelation and strategic foundations of a fully revealing Rational Expectations Equilibrium goes back to Wilson (1977), in an indivisible single object setting, and Hellwig (1980) and Wilson (1985), in a divisible exchange setting like ours.[1]. The papers find sufficient conditions such that information gets aggregated, when the number of agents becomes large. Unlike in this paper, the results depend critically on the symmetry assumptions between the traders.

Aside from double auctions, Roberts and Postlewaite (1976), Jackson (1992), and Jackson and Manelli (1997) show that Walrasian mechanisms are difficult to manipulate in large replica economies. Gul and Postlewaite (1992) and McLean and Postlewaite (2002) identify crucial variability of beliefs and information-smallness requirements sufficient for asymptotic efficiency. The information smallness requirement is closely related to our competitive information, but is weaker: while we require no group of agents to monopolise information, information smallness requires no single agent to monopolise information. This is dictated by our paper using a fixed market microstructure trading mechanism of market orders, while there the choice of a mechanism is open, and the optimal mechanism relies on the belief elicitation in the spirit of Cremer and McLean (1985).

The literature on market completeness focuses largely on competitive exchange markets, as well as the appropriate "spanning" conditions on prices that guarantee completeness (see Duffie (2010) for a textbook treatment and references therein). Anderson and Raimondo (2008) provide sufficient conditions on the primitives of an environment, rather than prices, for market completeness, in a competitive setting (see also Hugonnier, Malamud, and Trubowitz (2012)). At the same time, while the literature has long recognized strategic financial innovation as driving the spanning motive (see Allen, Gale et al. (1994) and Duffie and Rahi (1995) for surveys), it has offered few specific examples (see discussion in Duffie and Rahi (1995)). In this paper we provide requirements on the primitives that are required for market completeness, in a strategic exchange setting.

The two most closely related papers to ours are Lambert, Ostrovsky, and Panov

[1]Other important contributions include Milgrom (1981), Klemperer and Meyer (1989) , Pesendorfer and Swinkels (1997), Kremer (2002), Reny and Perry (2006). We refer to Vives (2011), Rostek and Weretka (2012) and Andreyanov and Sadzik (2021) for the analysis when agents have heterogenous values

(2018) and Carvajal, Rostek, and Weretka (2012). Lambert, Ostrovsky, and Panov (2018) analyzes similar informationally complex setting to ours and provides conditions that guarantee information aggregation, in economies with many replicas of each trader type. We do not focus on replica economies and our conditions are sufficient and necessary. Moreover, we consider a setting with multiple assets, allowing for the analysis of strategic market completeness.

Along with Allen and Gale (1991), Carvajal, Rostek, and Weretka (2012) stands out as providing a strategic model with insufficient conditions for strategic market completeness. The underlying settings and mechanisms in those papers are different than here. In Allen and Gale, strategic traders can miscoordinate, in a mixed equilibrium, on a beneficial but costly asset issue. In Carvajal, Rostek, and Weretka (2012), entrepreneurs can benefit from restricting the span of assets when investors have utilities exhibiting precautionary savings motive.

On the modelling side, the papers most related to ours is the work building on the classic model of Kyle (1985). The defining feature is the trade based on market orders, cleared by competitive market makers.[2] While the original model has only one trader and one factor, Admati and Pfleiderer (1988), Holden and Subrahmanyam (1994), Foster and Viswanathan (1996), Back, Cao, and Willard (2000) extended it to multiple informed, yet symmetric traders. Foster and Viswanathan (1994), Dridi and Germain (2009), and Colla and Mele (2010) consider asymmetrically informed traders, and Sadzik and Woolnough (2021) allow for multiple factors. Finally, while Caballe and Krishnan (1994) and Pasquariello (2007) allow for multiple assets, they maintain the assumption of symmetry among the agents.

2 Baseline Model

There are N payoff relevant factors, or fundamentals, $(x_1, ..., x_N) \in \mathbb{R}^N$, whose values are not initially known to the market participants. There are I strategic informed traders, and each trader i observes a multidimensional signal $s_i \in \mathbb{R}^{k_i}$, which is informative of the factors (and other signals). In the market, there are also nonstrategic, uninformed liquidity traders, who trade to hedge the aggregate shocks $(u_1, ..., u_N) \in \mathbb{R}^N$ to the factors, and an uninformed market maker. All those market participants trade M assets, with values $(v_1, ..., v_M) \in \mathbb{R}^M$ determined by the factors.

Let x, s, u, v be the corresponding vectors of factors, signals, liquidity shocks, and

[2]Along with limit orders, and the corresponding model of competition in demand schedules, market orders account for highest volume of assets exchanged on centralized exchanges; see ?

values (of dimensionalities $N, \sum_{i=1}^{I} k_i, N, M$). The key assumption that drives much of the tractability of our results is that all those variables are Normally distributed, with means normalized to 0.[3] Specifically, let Θ be the $N \times M$ matrix of linear coefficients that describe how the asset values depend on the factors,

$$v = \Theta^T x. \tag{1}$$

We shall assume that the assets are linearly independent: Matrix Θ has rank M.[4]

Let Σ be the covariance matrix of the vector of all the signals and the factors, of dimension $\sum_{i=1}^{I} k_i + N$. The matrix describes the covariances between the factors, as well as how the information about them is distributed in the market, how much the traders know about the information of others, etc. We only assume, without loss of generality, that each agent's signals are linearly independent, and so there are no redundancies in signals, and that each factor matters for the values and is distinguishable by the agents,[5] and so there are no redundant factors. Let Σ_x be the submatrix of Σ of covariances between the factors, Σ_s the submatrix of covariances between the signals. Note that the covariance matrix Σ_v of the values is $\Theta^T \Sigma_x \Theta$. Finally, matrix Σ_u describes the correlations in the aggregate liquidity shocks across different factors, of dimension N.

We note that both the information distributed in the market, as well as the liquidity shocks are defined over the factors, which are primitive in the model, and not directly over the assets. This allows us to analyze and compare the properties of markets for different sets of assets.

In the rest of the paper we identify a set of assets with the matrix of linear coefficients Θ. We call assets *non-overlapping* if each row of Θ has exactly one non-zero entry. Assets are *spanning* if $M = N$, and so the rank of Θ equals the number of factors. A special case of spanning assets has one asset corresponding to one factor, with $\Theta = Id_N$, the N-dimensional identity matrix.

In the paper we use the standard notation of matrix and vector multiplication, with $(\Theta^T x)_m = \sum_{n=1}^{N} \Theta^T_{mn} x_n$, vector and vector multiplication, with $x^T u = \sum_{n=1}^{N} x_n u_n$, and with "T" representing the transpose. For any positive semi-definite square matrices A and B of the same dimension we write $A \geq B$ if $A - B$ is positive semi-definite. For any two numbers x, y denote their maximum and minimum as $x \wedge y$ and $x \vee y$. Finally, for a matrix Θ, $\overline{\Theta}$ is the matrix with column lengths normalized to one, $\overline{\Theta}_{n,m} = \Theta_{n,m}/\|\Theta_{\cdot,m}\|$.

[3]The normalization of the means is made without loss of generality, in order to simplify notation.

[4]Without this assumption we would have to take stance on how the liquidity traders and the informed traders coordinate to trade two copies of the same asset or, more generally, on the sets of linearly dependent assets.

[5]Formally, the matrix Σ_{xs} of covariances of factors and all the informed traders' signals has rank N, and no column of Θ is identically zero.

2.1 Trade and Payoffs

After observing signal s_i, each informed trader i submits their vector of demands, $d_i(s_i) \in \mathbb{R}^M$, for all the assets. Liquidity shock u is realized, and liquidity traders submit the vector of their aggregate demands, $d_L(u) \in \mathbb{R}^M$. Market maker observes the vector of total demands for each asset, $d = \sum_{i=1}^{I} d_i(s_i) + d_L(v) \in \mathbb{R}^M$, and sets a vector of prices $P(d)$ for the assets, at which they clear the market. Informed trader i collects payoffs $\pi_i = d_i(s_i)^T(v - P(d))$, the long term value of their portfolio net the cost of acquiring it. Liquidity traders collect their aggregate payoffs, $\pi_L = -\left(x^T(u - \Theta d_L(u)\right)^2$, which are the negative square distance between values of their liquidity shock and their portfolio.

It follows that the liquidity demand is non-strategic, and depends only on the asset structure and the distribution of factors,

$$d_L(u) = -\left(\Theta^T \Sigma_x \Theta\right)^{-1} \Theta^T \Sigma_x u. \tag{2}$$

In particular, liquidity demand is linear in liquidity shocks. Hence, models with a single liquidity trader, or a fringe of liquidity traders with an aggregate shock u are equivalent.

For a fixed set of assets, let Σ_{*u} be the covariance matrix of liquidity demands for the assets. The matrix is a scalar variance, in the classic case of a single asset. When each factor corresponds to a separate factor, then $\Sigma_{*u} = \Sigma_u$.

Finally, note that as long as liquidity traders believe that prices are the unbiased estimates of the asset values, $\mathbb{E}[P^T(x-P)] = 0$, then a liquidity trader's demand would be unchanged, were their utilities replaced with $\pi_L^P = -\left(x^T u - P^T(d)d^L(u)\right)^2$, with portfolio evaluated at the market prices instead of the realized values—or, indeed, at a linear combination of the two.[6]

2.2 Linear Equilibrium

We use the essentially the same solution concept of a linear equilibrium as as Kyle '85.

Definition 1 *A profile of demand functions $\{d_i(\cdot)\}_{i=1}^{I}$ and a pricing rule $P(\cdot)$ are an* equilibrium *if the following two conditions are satisfied:*

(i) For any trader i and any signal s_i the demand $d_i(s_i)$ maximizes the expected payoff, given the strategies for the other traders $\{d_j(\cdot)\}_{j\neq i}$ and the pricing rule $P(\cdot)$.

[6]In Appendix we provide one brief microfoundation for this reduced form of the liquidity trader's utility. Roughly, each trader is small, and does not have superior information about the aggregate liquidity shock. They learn about their liquidity shock arriving at time two, which is correlated with the factors, and trades at time one to spread the impact of the liquidity shock between times one and two, given quadratic (or CARA) utility function.

(ii) For any vector of total demands d, the vector of prices $P(d)$ agrees with the vector of expected values of the assets conditional on the realized d, given the demand functions $\{d_i(\cdot)\}_{i=1}^I$.

Definition 2 *We say that an equilibrium is* linear*, if all the demand functions and the pricing rule are linear functions,*

$$d_i(s_i) = \beta_i^T s_i, \text{ for all } i$$
$$P(d) = \lambda^T d,$$

where β_i is a $k_i \times M$ matrix, with $(\beta_i)_{k,m}$ the derivative of trader i's demand for asset m on trader's k'th signal, and λ is a $M \times M$ matrix, with $\lambda_{m,\tilde{m}}$ the derivative of the price of $\tilde{m}$'th asset with respect to the demand for m'th asset.

The interpretation of the trading intensities β_i is similar as in the case of a single asset and signal. The interpretation of the price sensitivities λ is more delicate. Matrix λ is a generalization of "Kyle lambda", which with a single asset is a scalar, and represents the price impact of the trading volume—the inverse of the market depth. It's effect on the strategic trade, and on the inefficient demand reduction in the face of the price impact in particular is well understood. In our model, informed trader's demand for asset m has an impact on the prices of all other assets. This complicates the interpretation of λ as the inverse of the market depth. At the same time, the strategic problem of an informed trader goes beyond the one-dimensional trade-off between the volume of trade and the profit margin (the difference between the expected value and the price).

Let β be a matrix of all individual equilibrium trading intensities β_i, $i \leq I$, stacked on top of one another (of dimension $\sum_{i=1}^I k_i \times M$). Let $\Sigma_x^p = \mathbb{E}\left[(x - \mathbb{E}[x|d])(x - \mathbb{E}[x|d])^T\right]$ and $\Sigma_v^p := \mathbb{E}\left[(v - \mathbb{E}[v|d])(x - \mathbb{E}[v|d])^T\right]$ be the posterior covariance matrices about the factors and the asset values, conditional on the equilibrium vector of total order flows d.

3 Equilibrium and Information Revelation

The following is the first main result of the paper.

Theorem 1 *Fix an information structure Σ, liquidity trade Σ_u, and a set of assets Θ. There exists a unique linear equilibrium with a symmetric price impact matrix λ.*

In the equilibrium, the trading intensities and the price impact matrices are

$$\beta^T = \lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1}, \tag{3}$$
$$\lambda = D^{0.5}(D^{0.5}\Sigma_{*u}D^{0.5})^{-0.5}D^{0.5}, \tag{4}$$

where

$$D = \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{diag} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta,$$

and Σ_{xs} is the covariance between the factors and all the informed traders' signals, Σ_s is the covariance matrix between all the signals, and Σ_{diag} is the quasi-diagonal matrix, with the covariance matrices of individual traders' signals aligned along the diagonal (see Appendix).

The proof of the result is constructive, and establishes several properties on the way. First, it shows that given linear strategies of informed traders, pricing rule is linear; similarly, given a linear pricing rule and linear strategies of all but one trader, the best response is also linear. This result is fairly standard, given Normality of the distributions, total orders ("public signals") that are linear in strategies, and payoffs that are quadratic in own strategy and the price vector. Second, a novelty of the result is to show that, given price impacts, there exists a unique strategy profile with each informed trader best responding. Moreover, the trading intensities are the explicit function of price impacts and the primitives of the model. This holds despite the generality of the information structure, asset structure, or the liquidity trade, which may exhibit arbitrary correlations and interdependencies, as modelled by the primitives Σ, Θ, and Σ_u, respectively. Third, given the unique best response strategy profile of the informed traders, the proof shows that the matrix of learning parameters, or price impacts, satisfies a particular matrix quadratic equation. Importantly, the equation admits a unique symmetric solution, which depends only on the primitives of the model. This establishes existence and uniqueness of the market learning parameters in a multi-dimensional linear-Normal model, in which the correlations of the public signals (trading intensities) depend themselves endogenously on the learning.

Consider a market with given assets and factors, described by covariance matrix Σ_v and Θ (or, alternatively, Σ_x), and a given number of informed traders I. However, place no restrictions on what signals the agents get, and so how the information is distributed in the market (beside Σ_x being the marginal of Σ). We seek to characterize what information about the asset values can be revealed to the market when the assets are traded, across all the information structures. Specifically, the result below provides an explicit tight bound on how much information can be revealed, which depends only on the prior covariance Σ_v and the number of agents I.

Let a *Symmetric Model* be one in which each agent i knows the factors, $s_i = x$.

Proposition 1 *Fix a prior covariance of asset values Σ_v, and consider any information structure with I agents. In the linear equilibrium, the posterior covariance about the asset* *values satisfies*

$$\Sigma_v^p \geq \frac{1}{I+1}\Sigma_v. \tag{5}$$

Moreover, the bound is achieved in the symmetric model.

The maximal amount of information revealed to the market depends only on the prior uncertainty about the values, and the number of traders. In particular, there is limited scope for information design, and a way to distribute the information among the traders in the market. No shrewd design can lead to the revelation of values, with finitely many traders. Also, there is no way for the design to "trade" more revelation of one asset for less revelation of another asset. The bound on information revelation, across all assets, is achieved by a single information structure, in which all the traders are fully informed about the factors.

Even if several traders share and compete on trading away the same information, in a symmetric model, they do not reveal all the information about the values to the market. Hence, they make non-negligible profit: each trades a positive quantity (volume) at non-negative profit margin. Intuitively, the competition shares many features with the model of Cournot competition, albeit with an endogenous slope of the inverse demand function—or, in our case, the price impact.

At the same time, in the symmetric model, increasing the number of informed traders reduces the posterior covariance matrix—or the covariance between the prices and the values—down to zero, at a rate $(I+1)^{-1}$. This, again, is reminiscent of the price converging down to the (common, constant) marginal cost in the Cournot model. In the case when the assets are spanning the uncertainty, and so the matrix Θ is square and invertible, analogous results hold for the posterior covariance of the factors Σ_x^p

In the following result, we generalize the result on information revelation beyond a symmetric model. For that purpose, we consider a sequence of information structures $\{\Sigma_I\}$, parametrized by the number of agents I, with I converging to infinity. For simplicity, we assume that all structures are defined over the same space of factors, described by a fixed covariance matrix Σ_x and, consequently, all structures share a fixed matrix Σ_{*u} of liquidity trade (see (2)).[7] We also fix the set of assets, described by a matrix Θ.

We impose the following restriction on the sequence of information structures:[8]

Assumption 1 (Regularity) *For every $\varepsilon > 0$ there exists L such that for every infor-*

[7]Fixing the covariance matrix, or even the number of factors is not necessary for the results, but simplifies the statements and the interpretation.

[8]s' is a sub-vector of s if it is of lower dimension and can be obtained form s by eliminating some of s's components.

mation structure Σ_I *and any vector of signals* s*, there is a sub-vector* s' *of* s *of dimension* L *such that*

$$\mathbb{E}[(\mathbb{E}[x_n|s] - \mathbb{E}[x_n|s'])^2] < \varepsilon,$$

for every factor x_n

In other words, for any vector of signals and an approximation level, there is a sub-vector of a uniformly bounded size that carries approximately the same information.[9]

Example 1 *(Lambert et. al) Suppose there are* G *groups of traders. Each trader* i *belongs to one of the groups* $g \leq G$*, and receives a* k_g *dimensional signal* $s_i = \phi_g + \xi_{g,i}$ *where* ϕ_g *is common to the whole group, and* $\xi_{g,i}$ *is a noise that is independent of all other variables and is distributed identically within group* g*. The simplest example is one with only one group, and each trader observing a noisy signal of the factors:* $s_i = x + \xi_i$*.*

The model satisfies the regularity assumption: Only the finitely many variables ϕ_g *matter for the estimation of values and, from the Law of Large Numbers, for any group* $g \leq G$ *and* $\varepsilon > 0$*, there is a finite number of draws of signals that brings the posterior variance of* ϕ_g*'s estimate below* ε*.*

Consider the following condition on the information structures:

Condition 1 (Competitive Information) *For every* $\varepsilon > 0$*, every* $L > 0$*, every vector of signals* $s \in \mathbb{R}^{\sum_{i=1}^{I} k_i}$*, and every sub-vector* s' *of* s *with at most* L *signals eliminated,*

$$\mathbb{E}[(\mathbb{E}[x_n|s] - \mathbb{E}[x_n|s'])^2] < \varepsilon, \text{ for every factor } x_n,$$

for all except for finitely many information structures.

The condition requires that the additional information carried by any finite group of signals is negligible, in sufficiently large economies. This implies that no finite group of agents has a monopoly power on any significant amount of information, in a large economy. The condition thus captures the intuition that the information is distributed competitively in an economy.

The following is our second main result.

[9] Note that given the Normality of the distributions, the variance in the definition does not depend on the realizations of the signals, but only on their ex-ante distributions.

Theorem 2 *Consider a sequence of information structures $\{\Sigma_I\}$ that satisfies the regularity assumption, with a fixed covariance of factors Σ_x and fixed assets Θ. Then prices aggregate private information, i.e.,*

$$lim_{I \to \infty} \mathbb{E}[(P_m - \mathbb{E}[v_m|s])^2] = 0, \text{ for all assets } m \leq M, \tag{6}$$

if and only if the information structures satisfy the competitive information assumption.

The theorem shows that the condition of competitive information, placed on the exogenous information structures, is necessary and sufficient for prices to aggregate private information about the assets distributed in the economy, in the endogenous linear equilibrium of the trading game. We emphasize that the condition is not phrased in terms of the structure and details of the underlying model and, hence, is portable to other information environments.[10] In particular, it is not written in terms of explicit conditions on the structure of the signals—their correlations and informativeness—or on the structure of the liquidity trade. The condition carries a strong intuition of competitiveness of an informational environment and, we believe, should be readily verifiable in any given family of information structures.

Competitive information condition is phrased entirely in terms of the first-order information that the agents have about the fundamentals, and is silent about the higher-order information that the traders have about themselves. In particular, while it requires, roughly, that each trader's information about the fundamentals is replicated by many other traders, it does not require that the same is true of their information about other traders' information. This is in contrast with, say, models such as "replica" economies, where ever increasing groups of agents have signals that matter equally for the estimate of the fundamentals, and result in the same higher-order beliefs (see LOP).

This irrelevance of higher-order beliefs is an important consequence of the relatively simple structure of the trading game. The beliefs of all the other traders matter only to the extent of how they affect asset prices. Moreover, asset prices depend only on a simple statistic—the sum—of the submitted demands. At the same time, the simplicity of the trading game and the irrelevance of higher-order beliefs come at a price of the strength of the competitive information assumption. It is violated even in an environment in which finitely many traders share the same signals, and so each trader's information can be identified from the signals of others (compare Cremer and McLean, McLean and Postlewaite).

The intuition for the theorem is as follows. When the competitive information assumption is violated then, no matter the size of the economy, there is a finite group of L

[10]The same is true of the regularity condition.

traders who monopolise non-negligible amount of information about the fundamentals, and so assets. In any such economy, those L traders act as if they traded only among themselves, taking the trading strategies of everybody else as given. Put differently, in this auxiliary trading game, values of the assets are the original values net of the price impact of the equilibrium trade of everybody else. It follows from Proposition 1 that in this game only at most the fraction $L/(L+1)$ of the monopolised information is revealed, resulting in non-revelation of information, despite large economies.

Suppose the competitive information is satisfied and, by way of contradiction, some information is not revealed, in arbitrarily large economies. Then, from regularity, there is a group of agents of size L that has non-negligible amount of information about the assets. By trading on this information, traders in this group would be able to make non-negligible profits. This puts a lower bound on their equilibrium profits. Given competitive information, the complement of $I-L$ traders have essentially all the information about the assets. Hence, again, it contains the second group of L traders, who have non-negligible information about the assets and thus must make non-negligible profits in equilibrium. Proceeding in this way, the total equilibrium profits in a large economy would converge to infinity. This contradicts the bound on the total profits—say, by the bound on profits of a single, monopolistic trader with commitment power. In other words, large economies with competitive information must offer vanishing profit opportunities for traders, which is possible only when prices follow closely asset values.

Example 2 *Consider the information structures from Example 1. Suppose that, as the structures grow, the number of agents in each group $g \leq G$ converges to infinity. The condition implies competitive information and, hence, prices aggregate private information. The condition is also necessary for competitive information and information aggregation if the fundamental signals phi_g are uncorrelated.*

Example 3 *Consider Example 1 and suppose that, for some group g, each fundamental signal $\phi_{g,k}$ can be written as a linear combination of signals from other groups, plus noise that is uncorrelated with the factors. Information structures may satisfy competitive information even if the size of group g of traders remains bounded.*

4 Bundling Assets

We now turn attention to the assets and how they affect equilibrium trade. In particular, we consider how the assets impact information revelation about the factors.

Just as with competitive economies, assets must satisfy a spanning condition if prices are to reveal information about the hidden factors. Theorem 2 thus implies the following:

Corollary 1 *Consider a sequence of information structures* $\{\Sigma_I\}$ *that satisfies the regularity assumption, with a fixed covariance of factors* Σ_x *and fixed assets. Then prices aggregate private information about the factors, i.e.,*

$$lim_{I\to\infty}\mathbb{E}[(\mathbb{E}[x_n|P]-\mathbb{E}[x_n|s])^2]=0, \text{ for all factors } n\leq N, \tag{7}$$

if and only if the information structures satisfy the competitive information assumption, and there are at least N *linearly independent assets,* $M \geq N$*.*

The corollary provides the benchmark necessary and sufficient conditions on the information and asset structure under which asset prices aggregate private information about the factors. Information aggregation requires many traders with competitive information and sufficiently many, spanning assets. We emphasize that the conditions are on the primitives of the model: the informational environment and the instruments traded.

To gain further insight into the role of the assumptions, we investigate whether the appropriate comparative statics results hold also away from the limiting benchmark.

Proposition 2 *Consider an information structure* Σ *and a set of* M *assets* Θ*.*

i) Adding traders or signals results in less uncertainty about the asset values, in the linear equilibrium:

$$\Sigma_v^{p\Sigma'} \leq \Sigma_v^{p\Sigma}, \tag{8}$$

for the posterior covariance matrices in the two settings, for any matrix Σ' *that adds traders or signals to* Σ*,* $\Sigma'_{rc}=\Sigma_{rc}$*, for* (r,c) *lower than the dimension of* Σ*. It need not decrease uncertainty about the factors, when assets are not spanning,* $M<N$

ii) Adding assets results in less uncertainty about the factors, in the linear equilibrium:

$$\Sigma_x^{p\Theta'} \leq \Sigma_x^{p\Theta}, \tag{9}$$

for the posterior covariance matrices in the two settings, for any matrix Θ' *that adds assets to* Θ*,* $\Theta'_{nm}=\Theta_{nm}$*, for* $n\leq N, m\leq M$*.*

Adding traders, or having existing traders receive more information, results in more information about the assets passed on to the public. When assets are spanning, more information gets revealed about the factors too. Likewise, adding linearly independent assets enhances information about the factors revealed to the public. The results hold true for any fixed number of traders, and so away from the competitive limit, when traders act strategically, and for any information structure, with arbitrary details of the strategic interaction. In particular, the proposition provides a strong normative benchmark for the richness of the asset space in strategic settings, as measured by the yardstick of informational efficiency.

With Normally distributed variables, the reduction of the posterior covariance of the values is the price impact λ (learning parameters), times the covariance of the values and the vector of order flows (signal). Theorem 1 thus yields the following:

$$\Sigma_v^p = \Sigma_v - \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta. \tag{10}$$

The posterior covariance about the assets is as if the public directly observed the statistical estimates of the asset values given the signals, albeit each signal had twice larger variance, due to idiosyncratic noise. It follows that the informational effects of strategic incentives—how traders shade their asset demands, faced with limited competition for information and market power—is the same as the doubling of noise, in a non-strategic setting. Assets then serve to filter this information to the public, much as they would in a fully competitive setting. We also emphasize that the structure or the size of the liquidity trade plays no role in how much information is revealed to the public.

Given the equivalence of informational effects in the strategic and an appropriate nonstrategic setting, the idea behind the proposition is clear. Regarding the first part, adding signals always decreases posterior covariance about the values, in any nonstrategic learning problem. At the same time, when asset are not spanning, adding signals about other factors garbles the information about a factor of interest. Regarding the second part, with more assets, the information revealed by trade is the exogenous information in a nonstrategic setting, but expressed in a larger space of asset prices.

Consider the following simple example. Suppose there are two independent factors x_1, x_2, each with prior variance of one, and an asset with value $v_1 = x_1 + x_2$. If there is only one trader in the market that is fully informed about x_1, the market learns as if it observed the value estimate, based on the signal about x_1 with an additional noise with variance one. The posterior equilibrium variance about x_1 is a half. Suppose now that market is joined by the second trader that is fully informed about x_2. The market learns as if it observed the value estimate, based on the signals about x_1 and x_2, each with an additional noise with variance one. The posterior covariance matrix is

$$\Sigma_x^p = \begin{bmatrix} 0.75 & -0.25 \\ -0.25 & 0.75 \end{bmatrix},$$

and so the learning about the first factor is hampered. At the same time, the prior variance about asset value is two, the posterior variance with one trader is one and a half, and with two traders it is one. Finally, when a second linearly independent asset—say, $v_2 = x_2$—is added to the mix, the market learns as if it directly observed the signals about each factor (with increased variance), so that

$$\Sigma_x^{p,span} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

Many assets help with information revelation. However, would informed traders want to trade the large, spanning set of assets? To answer the question, we fix an information structure Σ as well as liquidity demand Σ_u, defined directly over the factors, and investigate how changing the sets of traded assets affects the sum of informed traders' payoffs, in equilibrium.

We restrict attention to the structures that satisfy the following three independence conditions. First, we say that Σ has *independent factors* if $(\Sigma)_{n,n'} = 0$, for $n \neq n'$. Second, Σ_u has *independent liquidity demands* if $(\Sigma_u)_{n,n'} = 0$, for $n \neq n'$. For such structures, we write $\sigma_{u,n}^2$ for $(\Sigma_u)_{n,n}$. Finally, Σ has *independent signals* if every signal has the form $s_{n,k} = x_n + \varepsilon_n + \varepsilon_{n,k}$, for some factor x_n, a factor-specific noise ε_n and an idiosyncratic noise $\varepsilon_{n,k}$, with all noises independent of other variables, and idiosyncratic noises identically distributed, for a given factor.

Private information in a structure that satisfies the above independence properties can be characterizes as follows. For each, n'th factor, the three relevant parameters are i) the number of agents that received a signal about this factor, $I_n \leq I$[11], ii) the informativeness of such signals, $Cor(s_{n,k}, x_n)$, and iii) the correlation between the signals, $Cor(s_{n,k}, s_{n,k'})$.

For each factor n define *profitability* π_n,

$$\pi_n = \frac{Cor(s_{n,k}, x_n)\sqrt{I_n}}{2 + (I_n - 1)Cor(s_{n,k}, s_{n,k'})}. \tag{11}$$

Proposition 3 *Suppose an information structure satisfies the independence conditions.*

i) With M non-overlapping assets Θ, the total expected profits of the informed traders satisfy[12]

$$\Pi = \sum_{m=1}^{M} \sqrt{\sum_{n=1}^{N} \overline{\Theta}_{n,m}^2 \pi_n^2} \sqrt{\sum_{n=1}^{N} \overline{\Theta}_{n,m}^2 \sigma_{u,n}^2}. \tag{12}$$

ii) Total expected profits by the informed traders are maximized with spanning assets exactly when

$$\pi_n \sigma_{u,n} + \pi_{n'} \sigma_{u,n'} \geq \alpha_{n,n'}[(\pi_n \wedge \pi_{n'})(\sigma_{u,n} \wedge \sigma_{u,n'}) + (\pi_n \vee \pi_{n'})(\sigma_{u,n} \vee \sigma_{u,n'})]. \tag{CORC}$$

[11] We assume that an agent can get at most one signal about a given factor.

[12] Recall that $\overline{\Theta}$ is the matrix Θ with column lengths normalized to one.

for every pair of factors and parameters $\alpha_{n,n'} \in [1/2, 1]$ *that depend only on the details of the information structure,* $n, n' \leq N$.

The first part of the proposition provides an explicit formula for the total expected profits by informed traders, when each factor affects at most one asset. In particular, when assets correspond to factors, $\Theta = Id_N$, the profits are the sum of profitabilities multiplied by liquidity demands (standard deviations), for each factor,

$$\Pi = \sum_{n=1}^{N} \pi_n \sigma_{u,n}.$$

With bundled assets, profits satisfy an analogous formula, with both "profitability" and liquidity demand for each asset equal to a weighted square mean of the parameters for the underlying factors.

For each asset, the sum of informed profits equals the losses of the liquidity traders. Those uninformed liquidity losses stem from trading at the adversely affected price, and equal price impact times the volume of trade. The volume of liquidity trade thus has an analogous effect on informed profits as the demand volume, in a basic model of oligopolistic competition. Price impact is analogous to the profit margin and, as we show, is proportional to the profitability parameter, which depends on the structure of the private information and the degree of competition.

Correlation between the signals and a factor measures the informational advantage of the informed traders over the market. It has positive effect on the profitability of the factor, as it is proportional to the absolute difference between the factor price, absent any trade, and the conditional expected value of the factor. Correlation between the signals measures the degree of homogeneity of the informed traders' signals. High correlation translates into high competition between the traders and, hence, has negative effect on the profits. Finally, the effect of the number of traders informed about a factor has on its profitability is ambiguous. With poorly correlated signals and few traders, increasing the number of traders has a positive effect on profitability, as the positive effect of increased overall private information dominates. With highly correlated signals or many traders, more traders only increases competition, and dampens profitability.

An asset that depends on several factors averages, or bundles together the profitabilities and the liquidity demands of the underlying factors. The orders for the bundled asset come from traders informed about each of the underlying factors. They are proportional to the factor's effect on the asset and to the degree of informational competition for the factor, expressed in their profitability parameter. Total liquidity demand for the asset likewise comes from traders demanding liquidity for each of the underlying factors.

Our model of liquidity demand for a bundled asset is conservative, as it fully factors in the negative effect of increased risk of bundled assets. The typical (unsigned) liquidity order for a bundled asset, $\sigma_{*u,m}$, is an average and always strictly smaller than the sum of individual liquidity orders for the factors; it is equal to one of them when the individual orders are equal. Liquidity traders that want to insure against the shocks to a factor slash their demand if they can only buy an asset that exposes them to a shock to a different factor. Bundling may be profitable despite the overall negative effect on the liquidity demand and, hence, the inefficiency from reduced risk hedging by liquidity traders.

The second part of the proposition provides exact conditions for when bundling is profitable.[13] Spanning assets provide maximal total profits when the pairs of liquidity demands and profitabilities, for each factor, are close to assortative: There is a factor with lowest liquidity demand and profitability, one with second lowest liquidity demand and profitability, and so on. Bundling is optimal when near-assortativity fails for at least one pair of factors. The "fudge factor" is specific to each pair, but the ratio of the sum of products of liquidity demands and profitabilies across the factors and along the factors must be less than two.

The intuition for bundling goes back to the story of Cuban oranges and rotten cucumbers. Suppose there is large liquidity demand for factor one (Cuban oranges), but it's profitability is negligible: say, there is very little private information about it distributed in the market, above and beyond what is known publicly, by market makers. On the other hand, there is little liquidity demand for factor two (rotten cucumbers), but it has high profitability, or informed traders have large informational advantage about this factor. Trader separately, total informed profits are hamstrung by low profit margins for factor one, and low demand for factor two. When bundled, both the demand and the profit margin will be sizeable.

4.1 Strategic Bundling of the Assets

Proposition 3 established comparative statics result on the informed payoffs for different structures of assets. Here we turn to the incentives of strategic informed traders to issue assets. Whereas for a fixed sets of assets informed traders compete away information about them, under competitive information, the question here is whether competition between the traders would also result in a complete, spanning set of assets. In the rest of this section we consider *symmetric* information structures, which satisfy the three independence properties (see Section 4) and informed traders are symmetrically informed

[13] The proof provides the characterization of the parameters $\alpha_{n,n'}$, for $n, n' \leq N$

about the factors ($I_n = I$ for every factor n).

We consider the following asset issue game between the traders. Each trader $i \leq I$ chooses to issue $M_i \leq N$ linearly independent assets, which may be described by the matrix of linear coefficients Θ_i of dimension $N \times M_i$. The choices are made simultaneously, and the total issue cost for trader i is κM_i, and so linear in the number of assets, with $\kappa > 0$. Given the issued assets, the traders subsequently trade as in Section 3, for a set of $M \leq N$ linearly independent assets Θ that spans the same space as the set of all the issued assets.[14] For a given profile of pure asset issue strategies and a resulting matrix Θ of traded assets, payoffs to trader i are the expected payoffs from trading assets, in the unique linear equilibrium (Theorem 1).

Few comments are due. We show in Appendix that informed traders' expected profits as well as aggregate utilities of the liquidity traders in the trading game are the same for any two sets of assets with the same span. Moreover, aggregate liquidity traders utilities are strictly increasing in the span of assets traded. One implication is that the payoffs of the informed traders in the asset issue game are well defined, as the exact choice of Θ is irrelevant. More broadly, the result helps justify the choice of maximal span assets traded based on efficiency considerations, and abstracting away from the details of how traders coordinate on the traded assets.

We highlight that the game favors the liberal issue of the assets, as long as the issue cost parameter κ is small. No player can prevent or undo assets issued by others. At the same time, a single player can expand the set of assets: the span of assets issued by any individual trader is a lower bound on the span of assets eventually traded.

Theorem 3 *Suppose an information structure is independent and consider the asset issue game. i) If condition (CORC) is satisfied, then in any Nash equilibrium, the resulting assets are complete, with probability one.*

ii) If condition (CORC) is violated, then in any Nash equilibrium in weakly undominated strategies the resulting assets are complete, with probability one.

When the (CORC) condition is satisfied, the payoffs of each symmetric informed trader are maximized when assets are complete. Each trader can bring those profits about by issuing spanning assets.

We note that if issuing assets was costly, it would be possible for the players seeking to minimize costs to miscoordinate on the asset issue, resulting in non-spanning assets, with positive probability (see Allen Gale).[15] However, with vanishing costs, the probability

[14] Recall that a span of a matrix of assets Θ of dimension $N \times M$ is $span(\Theta) = \left\{\sum_{m=1}^{M} \lambda_m \Theta_{\cdot,m} | (\lambda_1, ..., \lambda_M) \in \mathbb{R}^M\right\}$.

[15] For example, consider a cost κM_i for issuing M_i assets, with a parameter $\kappa > 0$.

of mis-coordination vanishes as well, in any Nash equilibrium, or else each single player would rather "play it safe" and issue spanning assets herself.

When the (CORC) condition violated, the payoffs are maximized with non-spanning assets. Without any restrictions on strategies used, however, there would always be a trivial equilibrium where each trader issues spanning assets. The players would mis-coordinate on wrong strategies. Note that issuing spanning assets is a weakly dominated strategy.

The results shows that when players use weakly undominated strategies, the mis-coordination on issuing too many assets will not happen. It would never benefit a trader to issue assets that span larger space than the ones that maximize payoffs. The result provides a strong strategic justification for incomplete markets.

5 Appendix: Proofs

5.1 Proof of Theorem 1

We first argue that in an equilibrium matrix λ must be positive definite. If the matrix had an eigenvalue that is strictly negative, then any trader could make infinite profits, by submitting demands aligned with the corresponding eigenvector. With an eigenvalue of zero, demands aligned with the corresponding eigenvector have no price impact. For a trader that receives a signal about the value of the corresponding eigenvector, this would, again, result with infinite profits.[16]

Laet us now derive the optimality conditions for a trader i. Given the strategies of other traders nd the price impact, submitting demand d_i given signal s_i results in the expected utility

$$\mathbb{E}_i\left[d_i^T\left[v-P\left(d_i+\sum_{j\neq i}^{I}\beta_j^T s_j\right)\right]|s_i\right]=\mathbb{E}_i\left[d_i^T\left[\Theta^T x-\lambda^T\left(d_i+\sum_{j\neq i}^{I}\beta_j^T s_j\right)\right]|s_i\right],$$

and so the necessary first-order condition (FOC) is

$$\Theta^T\mathbb{E}_i\left[x|s_i\right]-\lambda^T\left(\sum_{j\neq i}^{I}\beta_j^T\mathbb{E}_i\left[s_j|s_i\right]\right)-(\lambda+\lambda^T)d_i=0,$$

and the second-order condition is that the matrix $\lambda+\lambda^T$ is positive definite, which is satisfied.

Using the projection theorem for the Normal variables, the definition of d_i, and the symmetry of λ, the FOC can be rewritten as

$$\Theta^T\Sigma_{xi}\Sigma_{ii}^{-1}s_i-\lambda\left(\sum_{j\neq i}^{I}\beta_j^T\Sigma_{ji}\Sigma_{ii}^{-1}s_i\right)-2\lambda\beta_i^T s_i=0, \tag{13}$$

where Σ_{xi} is the covariance between the factors and i's signals, and Σ_{ji} is the covariance between j's and i's signals. As the condition must hold for all signal vectors s_i, it is

[16] Recall the assumption that there are no redundant factors, and so there is some information among the traders about each of them.

equivalent to

$$\Theta^T \Sigma_{xi} \Sigma_{ii}^{-1} - \lambda \left(\sum_{j \neq i}^{I} \beta_j^T \Sigma_{ji} \Sigma_{ii}^{-1} \right) - 2\lambda \beta_i^T = 0, \tag{14}$$

$$\Theta^T \Sigma_{xi} \Sigma_{ii}^{-1} - \lambda \left(\sum_{j=1}^{I} \beta_j^T \Sigma_{ji} \Sigma_{ii}^{-1} \right) - \lambda \beta_i^T \Sigma_{ii} \Sigma_{ii}^{-1} = 0,$$

$$\lambda \left(\sum_{j=1}^{I} \beta_j^T \Sigma_{ji} \right) + \lambda \beta_i^T \Sigma_{ii} = \Theta^T \Sigma_{xi},$$

$$\left(\sum_{j=1}^{I} \beta_j^T \Sigma_{ji} \right) + \beta_i^T \Sigma_{ii} = \lambda^{-1} \Theta^T \Sigma_{xi}.$$

Hence, by stacking those equations for all the agents from left to right, with the matrix Σ_{diag} that has Σ_{ii} boxes on the diagonal, and Σ_s that is the covariance matrix between the signals, and Σ_{xs} the covariance matrix between the factors and signals, and β^T that has all the individual β_j^T stacked from left to right, we have

$$\beta^T \Sigma_{diag} + \beta^T \Sigma_s = \lambda^{-1} \Theta^T \Sigma_{xs}, \tag{15}$$

$$\beta^T = \lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s \right)^{-1}. \tag{16}$$

This is exactly the formula for the individual trading intensities in the statement of the theorem.

We now turn to the problem of the market maker. Projection Theorem for Normal variables implies that the matrix λ equals the matrix of covariances between the "public signals", i.e., total order flows, and the hidden asset values. Given the strategies of the informed traders and the liquidity trade with covariance Σ_{*u}, as well as the linearity of values in factors (see (1)), the formula is

$$\lambda = \Theta^T \Sigma_{xs} \beta \left[\beta^T \Sigma_s \beta + \Sigma_{*u} \right]^{-1}, \tag{17}$$

where β is the matrix of all the individual trading intensities β_i stacked from top to bottom. Consequently,

$$\beta^T \Sigma_s \beta + \Sigma_{*u} = \lambda^{-1} \Theta^T \Sigma_{xs} \beta,$$

And so, substituting for β from equation (15),

$$\lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_s \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_{xs}^T \Theta \lambda^{-1} + \Sigma_{*u} = \lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_{xs}^T \Theta \lambda^{-1},$$

$$\lambda^{-1} \Theta^T \Sigma_{xs} \left[\left(\Sigma_{diag} + \Sigma_s \right)^{-1} - \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_s \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \right] \Sigma_{xs}^T \Theta \lambda^{-1} = \Sigma_{*u},$$

$$\lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \left[\Sigma_{diag} + \Sigma_s - \Sigma_s \right] \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_{xs}^T \Theta \lambda^{-1} = \Sigma_{*u},$$

$$\lambda^{-1} \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_{diag} \left(\Sigma_{diag} + \Sigma_s \right)^{-1} \Sigma_{xs}^T \Theta \lambda^{-1} = \Sigma_{*u}. \tag{18}$$

Since there are no redundancies in signals, the matrix $(\Sigma_{diag} + \Sigma_s)^{-1} \Sigma_{diag} (\Sigma_{diag} + \Sigma_s)^{-1}$ has full rank. Moreover, since there are no redundancies in factors, the matrix Σ_{xs} has rank N, and Θ has rank M, from Assumption 1, the matrix $\Theta^T \Sigma_{xs}$ has rank M. It thus follows that the $M \times M$ matrix "between" the two λ^{-1}, in the last display equation, as well as Σ_{*u} are positive definite. Consequently, there is a unique positive definite matrix λ that solves the last display equation. It can be verified that the solution is as in the statement of the theorem.

5.2 Proof of Proposition 1

Fix the matrices Σ_v, Σ_x, and Θ.

Tightness of the bound. It follows immediately from Theorem 1 that in a symmetric model the traders use identical strategies in the equilibrium, with intensities β_1. The first-order condition (13) reduces to

$$\begin{aligned} \Theta^T s_i - (I-1)\lambda\beta_1^T s_i - 2\lambda\beta_1^T s_i &= 0, \\ (I+1)\lambda\beta_1^T &= \Theta^T. \end{aligned} \tag{19}$$

In a symmetric model, the formula for the posterior covariance matrix of the values takes the form

$$\Sigma_v^p = \Sigma_v - \lambda I \beta^T \Sigma_x \Theta = \Theta^T \Sigma_x \Theta - \frac{I}{I+1} \Theta^T \Sigma_x \Theta = \frac{1}{I+1} \Sigma_v.$$

The first equality above restates that the covariance reduction is the product of the gain parameter λ and the covariance of the signal and the values, for the special case of a symmetric model. The second equality follows from (19).

Sufficiency of the bound. First, in the Appendix we show that in any given information structure, linearly transforming ("renaming") signals of each agent i by an invertible matrix Y_i, as well as accordingly adjusting the cross-correlations across the agents as well as between the signals and factors, does not affect the distributions of the observable outcomes in equilibrium. Specifically, equilibrium trading intensities are linear transformations of those in the original model, but the distribution of order flows or asset prices are unchanged. Consequently, in the rest of the proof we assume that each agent's signals are linearly independent, and signal variances are normalized to one.

Second, we may assume that the values of the factors, and hence assets, are determined by the collective information of the agents, $x = \mathbb{E}[x|s]$. Otherwise, both the prior and the posterior covariance matrices of the values include the covariance $\Sigma_{noise} > 0$ of the noise that is unobservable by anyone in the economy. Replacing factors with their expectations, given the full vector of signals, does not affect the equilibrium, and results in the prior and posterior covariances of values decreased by $\Sigma_{noise} > 0$. With rescaling, this would result in more information revelation.

The Bayesian estimates of the values, given a vector of signals s, are given by

$$\mathbb{E}[v|s] = \Theta^T \Sigma_{xs} \Sigma_s^{-1} s.$$

Hence, the prior variance of the value is given by

$$\Sigma_v = \Theta^T \Sigma_{xs} \Sigma_s^{-1} \Sigma_s \Sigma_s^{-1} \Sigma_{xs}^T \Theta = \Theta^T \Sigma_{xs} \Sigma_s^{-1} \Sigma_{xs}^T \Theta,$$

whereas it follows from Corollary **??** that the amount of information revealed in equilibrium is given by

$$\Sigma_v - \Sigma_v^p = \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta.$$

Note that, given the normalization of private signals, the matrix Σ_{diag} is an identity matrix.

To establish the proof it suffices to show that

$$\left(\Sigma_{diag} + \Sigma_s\right)^{-1} \leq \frac{I}{I+1} \Sigma_s^{-1}. \tag{20}$$

Indeed, the covariance matrix Σ_s has ones on the diagonal, and the off-diagonal entries bounded by 1, in absolute value. It follows from the Gershgorin Circle Theorem that all the eigenvalues of Σ_s are bounded above by I. Consequently,

$$-\frac{1}{I}\Sigma_s + \Sigma_{diag} \geq 0,$$

$$\Sigma_s + \Sigma_{diag} \geq \Sigma_s + \frac{1}{I}\Sigma_s = \frac{I+1}{I}\Sigma_s,$$

which implies (20).

5.3 Proof of Theorem 2

Necessity. Consider a sequence $\{\Sigma_I\}$ of information structures that violates the competitive information assumption. There exist $\varepsilon > 0$, $L < \infty$, factor x_n, and an infinite sub-sequence of structures, such that each Σ_I has a vector of signals s'' of dimension L that carries at least ε information about x_n, in the sense that $\mathbb{E}[(\mathbb{E}[x_n|s] - \mathbb{E}[x_n|s'])^2] \geq \varepsilon$, for $s \in \mathbb{R}^{\sum_{i=1}^I k_i}$ and $s' = s/s'' \in \mathbb{R}^{\sum_{i=1}^I k_i - L}$ the complement of s'' in s.

Pick one such information structure Σ_I and fix the linear equilibrium, with trading intensities β and price impact λ. Let

$$v^* = v - \lambda(\beta_{s'}^T s' + d_L(u))$$

be the vector of values minus the vector of prices, when traders trade only on signals s' and not on s''. Vector v^* can be rewritten as $\Theta^T x^*$, for appropriately redefined factors

x^* (see (14) in the proof of Theorem 1). Let $\beta_{s''}$ be the associated matrix of equilibrium trade intensities on signals $s'' \in \mathbb{R}^L$. The first-order condition on $\beta_{s''}$ is identical to the first-order condition in an information structure Σ^* that has only signals s'', and the vector of factors is x^*, for the fixed price impact matrix λ (see (14)).

By definition, the posterior covariance Σ^p_{v*} about the values v^*, given $\beta_{s''}$, is the same as the posterior covariance Σ^p_v about the values v, given β, in the original structure Σ_I. Hence,

$$\begin{aligned}
\Sigma^p_v = \Sigma^p_{v*} \geq \frac{1}{L+1}\Sigma_{v*} &\geq \frac{1}{L+1}\mathbb{E}[(v-\lambda\beta^T_{s'}s')(v-\lambda\beta^T_{s'}s')^T] \\
&\geq \frac{1}{L+1}\mathbb{E}[(v-\mathbb{E}[v|s'])(v-\mathbb{E}[v|s'])^T] \\
&\geq \frac{1}{L+1}\mathbb{E}[(\mathbb{E}[v|s]-\mathbb{E}[v|s'])(\mathbb{E}[v|s]-\mathbb{E}[v|s'])^T].
\end{aligned}$$

The first inequality follows from Proposition 1.[17] [18] The second inequality follows since the right-hand side matrix is the covariance matrix of a vector $v^* - \lambda d_L(u)$, which leaves out the additional, uncorrelated liquidity noise from the vector v^*. The third inequality follows from the fact that $\mathbb{E}[v|s']$ has the form As' for a matrix A that, by definition, minimizes the covariance matrix of $v - \mathbb{E}[v|s']$. The last inequality follows, as the second one, from leaving out uncorrelated noise vector in v.

The proof of necessity thus follows because, for any $m \leq M$,

$$\mathbb{E}[(\mathbb{E}[v|s]-\mathbb{E}[v|s'])(\mathbb{E}[v|s]-\mathbb{E}[v|s'])^T]_{mm} \geq (\Theta_{nm})^2\mathbb{E}[(\mathbb{E}[x_n|s]-\mathbb{E}[x_n|s'])^2] \geq (\Theta_{nm})^2\varepsilon.$$

Sufficiency. To simplify notation, we assume that, in any information structure in the sequence, $v = \mathbb{E}[v|s]$. Also, given Proposition XX in Appendix X, we assume that each agent's signals are linearly independent, and signal variances are normalized to one. We start with two preliminary results.

Claim 1. There exists a positive definite $\overline{\lambda}$ such that $\lambda \leq \overline{\lambda}$, for any information structure in the sequence.

Recall that Σ_v is the covariance of the asset values, for assets described by Θ and a fixed covariance of factors Σ_x in the sequence. It is enough to show that matrices D, in the definition of λs are bounded above (in the sense that the difference is positive semi-definite). Indeed,

[17] The decrease in the variance is $\Theta^T\Sigma^*_{xs}\left(\Sigma^*_{diag}+\Sigma^*_s\right)^{-1}\Sigma^{*T}_{xs}\Theta$.

[18] Formally, we need that the bound holds for any matrix λ, not just the equilibrium one. This is true, since $\lambda\beta^T$ has λ canceling out.

$$D = \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{diag} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta$$
$$\leq \Theta^T \Sigma_{xs} \left(\Sigma_{diag} + \Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta \leq \Theta^T \Sigma_{xs} \left(\Sigma_s\right)^{-1} \Sigma_{xs}^T \Theta = \Sigma_v.$$

Claim 2. There exists $\overline{\pi}$ such that the sum of informed traders' equilibrium expected payoffs is bounded by $\overline{\pi}$, for any information structure in the sequence.

It is sufficient to show that the sum of expected payoffs from trade on any of the M assets remains bounded. Hence, we focus on he case of a single asset, $M = 1$. For any information structure, the sum of payoffs is bounded by above by the payoffs in the case when the informed traders pooled their information s, and then chose cooperatively the trading strategy for everybody, given that the price impact λ is consistent with the strategy. Representing the sum of informed traders' demands $d_I(s)$ as a linear function of the conditional expectation $\tilde{v} = \mathbb{E}[v|s]$ of the value plus independent noise, $b\tilde{v} + \eta$, the total expected payoffs are bounded by

$$max_{b,\eta} \left\{\mathbb{E}\left[\left(b\tilde{v} + \eta\right)\left(\tilde{v} - \lambda(b\tilde{v} + \eta)\right)\right]\right\} = max_{b,\eta} \left\{\mathbb{E}\left[\left(b\tilde{v} + \eta\right)\left(\tilde{v} - \frac{b\sigma_{\tilde{v}}^2(b\tilde{v} + \eta)}{b^2\sigma_{\tilde{v}}^2 + \sigma_\eta^2 + \sigma_{*u}^2}\right)\right]\right\}$$
$$= max_{b,\eta} \left\{b\sigma_{\tilde{v}}^2 \left[1 - \frac{b^2\sigma_{\tilde{v}}^2 + \sigma_\eta^2}{b^2\sigma_{\tilde{v}}^2 + \sigma_\eta^2 + \sigma_{*u}^2}\right]\right\} = max_b \left\{b\sigma_{\tilde{v}}^2 \frac{\sigma_{*u}^2}{b^2\sigma_{\tilde{v}}^2 + \sigma_{*u}^2}\right\} = \frac{\sigma_{\tilde{v}}\sigma_{*u}}{2}.$$

This establishes the proof of the claim, since $\sigma_{\tilde{v}} \leq \sigma_v$.

Given the two claims, the proof by contradiction is as follows. Suppose that a sequence of information structures satisfies the competitive information assumption but, for some asset m, $lim_{I\to\infty}\mathbb{E}[P_m - \mathbb{E}[v_m|s])^2] \geq \varepsilon_1 > 0$; by focusing on a sub-sequence, we may assume that the inequality is satisfied for all the structures in the sequence. Since P_m is a conditional expectation of v, $\mathbb{E}[P_m^2] \leq \sigma_v^2 - \varepsilon_1$, in any information structure. Therefore, for any vector of signals s'

$$\mathbb{E}[(v_m - \mathbb{E}[P_m|s'])^2] \geq \mathbb{E}[v_m^2] + \mathbb{E}[\mathbb{E}[P_m|s']^2] - 2\sqrt{\mathbb{E}[v_m^2]\mathbb{E}[\mathbb{E}[P_m|s']^2]}$$
$$= (\sqrt{\mathbb{E}[v_m^2]} - \sqrt{\mathbb{E}[\mathbb{E}[P_m|s']^2]})^2 \geq (\sqrt{\mathbb{E}[v_m^2]} - \sqrt{\mathbb{E}[P_m^2]})^2 \geq (\sigma_v - \sqrt{\sigma_v^2 - \varepsilon_1})^2 := \varepsilon_2.$$

From regularity, let L be such that for any vector of signals s', there is a subvector s'' of s of dimension L such that $\mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])^2] < \varepsilon_2/8$. From competitive information, for any $\overline{L}$ there is $I(\overline{L})$ such that for any structure $I \geq I(\overline{L})$, vector of signals s and every subvector s' of s with at most $\overline{L}$ signals eliminated, $\mathbb{E}[(\mathbb{E}[v_m|s'] - v_m)^2] < \varepsilon_2/8$. Consequently, for any $\overline{L}$ and structure $I \geq I(\overline{L})$, as well as any subvector s' with at most

$\overline{L}$ signals eliminated, there is a subvector s'' of s' of dimension L such that

$$\begin{aligned}
&\mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)^2] \\
&= \mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])^2] + \mathbb{E}[(\mathbb{E}[v_m|s'] - v_m)^2] + \mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])(\mathbb{E}[v_m|s'] - v_m)] \\
&\leq \mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])^2] + \mathbb{E}[(\mathbb{E}[v_m|s'] - v_m)^2] + 2\sqrt{\mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])^2]\mathbb{E}[(\mathbb{E}[v_m|s'] - v_m)^2]} \\
&= (\sqrt{\mathbb{E}[(\mathbb{E}[v_m|s''] - \mathbb{E}[v_m|s'])^2]} + \sqrt{\mathbb{E}[(\mathbb{E}[v_m|s'] - v_m)^2]})^2 < \varepsilon_2/2.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\mathbb{E}[\mathbb{E}[v_m - P_m|s'']^2] \qquad (21) \\
&= \mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)^2] + \mathbb{E}[(v_m - \mathbb{E}[P_m|s''])^2] + 2\mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)(v_m - \mathbb{E}[P_m|s''])] \\
&\geq \mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)^2] + \mathbb{E}[(v_m - \mathbb{E}[P_m|s''])^2] - 2\sqrt{\mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)^2]\mathbb{E}[(v_m - \mathbb{E}[P_m|s''])]^2} \\
&= (\sqrt{\mathbb{E}[(v_m - \mathbb{E}[P_m|s''])]^2} - \sqrt{\mathbb{E}[(\mathbb{E}[v_m|s''] - v_m)^2]})^2 > (\sqrt{\varepsilon_2} - \sqrt{\varepsilon_2/2})^2 := \varepsilon_3 > 0.
\end{aligned}$$

The rest of the proof is as follows. For any vector $s'' = (s''_1, ..., s''_L)$ of dimension L for which (21) holds, $\mathbb{E}[v_m - P_m|s''] = \sum_{l=1}^{L} \gamma_l s''_l$ implies that there is $s''_{*l} \in s''$ such that

$$\mathbb{E}[(\mathbb{E}[v_m - P_m|s''_{*l}])^2] \geq \mathbb{E}[(\gamma_l s''_{*l})^2] = max_{l \leq L}\mathbb{E}[(\gamma_l s''_l)^2] \geq \frac{1}{L^2}\mathbb{E}\left[\left(\sum_{l=1}^{L} \gamma_l s''_l\right)^2\right] \geq \frac{\varepsilon_3}{L^2}. \qquad (22)$$

The trader that corresponds to a signal s''_{*l} for which (22) holds has an available strategy that consists in trading only asset m based on this signal. Given an upper bound $\overline{\lambda}_{mm}$ of the impact of m's order flow on it's price (Claim 1), the strategy would yield expected profits $\frac{\varepsilon_3}{L^2 \lambda_{mm}} := \varepsilon_4 > 0$. Hence, ε_4 is also a lower bound on this trader's expected equilibrium profits, based solely on the information s''_{*l}.

Finally, pick $\overline{L} > \frac{\overline{\pi}}{\varepsilon_4} L$ and consider any information structure $I \geq I(\overline{L})$. Let s'''^0 be a full vector of signals in the structure. From regularity, pick an L-dimensional sub-vector s''^1 for which (21) holds, and, hence, a player in this group receiving signal s''^1_{*l} with equilibrium expected payoffs off the trade on this signal bounded below by ε_4. Let s'''^1 be the sub-vector of s'''^0 that has signals in s''^1 omitted. Again, from regularity, pick an L-dimensional sub-vector s''^2 of s'''^1, and a player receiving a signal in this group with equilibrium payoffs off the trade of this signal bounded below by ε_4. (If this trader was also receiving signal s''^1_{*l} then the equilibrium payoffs are bounded below by $2\varepsilon_4$, given independence of the signals.)

Proceeding inductively, there is a vector $(s''^1_{*l}, s''^2_{*l}, ..., s''^{\overline{L}/L}_{*l})$ of $\overline{L}/L$ signals, with equilibrium payoffs off the trade on them bounded below by $\varepsilon_4 \overline{L}/L > \overline{\pi}$. Contradiction establishes the proof.

5.4 Proof of Proposition 2

The proofs of both parts of the proposition rely on the following:

Claim 1 *The following matrix inequality holds*

$$\begin{pmatrix} Z & Y \end{pmatrix} \begin{pmatrix} A & B^T \\ B & C \end{pmatrix}^{-1} \begin{pmatrix} Z & Y \end{pmatrix}^T \geq ZA^{-1}Z^T,$$

when A and C are symmetric positive definite matrices and B is such that the square block matrix is positive definite.

Note that for the matrix inequality in the claim to be well defined, the dimensions of the matrices are: Z is $n_1 \times n_2$, Y is $n_1 \times n_3$, A is $n_2 \times n_2$, C is $n_3 \times n_3$, and B is $n_3 \times n_2$.

Proof. (Claim) Let $\bar{C} = (C - BA^{-1}B^T)^{-1}$ be the inverse of Schur's complement of A, which is symmetric positive definite. Using matrix block inversion we get

$$\begin{pmatrix} A & B^T \\ B & C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B^T\bar{C}BA^{-1} & -A^{-1}B^T\bar{C} \\ -\bar{C}BA^{-1} & \bar{C} \end{pmatrix}^{-1}.$$

Using this,

$$\begin{aligned} \begin{pmatrix} Z & Y \end{pmatrix} \begin{pmatrix} A & B^T \\ B & C \end{pmatrix}^{-1} \begin{pmatrix} Z & Y \end{pmatrix}^T &= ZA^{-1}Z^T + ZA^{-1}B^T\bar{C}BA^{-1}Z^T \\ &- Y\bar{C}BA^{-1}Z^T - ZA^{-1}B^T\bar{C}Y^T + Y\bar{C}Y^T \\ &= ZA^{-1}Z^T + (ZA^{-1}B^T - Y)\bar{C}(ZA^{-1}B^T - Y)^T, \end{aligned}$$

where $(ZA^{-1}B^T - Y)\bar{C}(ZA^{-1}B^T - Y)^T$ is symmetric positive definite. This establishes the claim.

Part i) Consider adding an extra informed trader and denote their vector of signals s', so that $\Sigma_{s's'}$, $\Sigma_{ss'}$, and $\Sigma_{xs'}$ represent the covariance matrix of the new signals, the covariance of the original and the new signals, and the covariance of the factors and the new signals. From 10, reduction in the posterior covariance of the values in the linear equilibrium with the new trader equals

$$\Theta^T \begin{pmatrix} \Sigma_{xs} & \Sigma_{xs'} \end{pmatrix} \left(\begin{pmatrix} \Sigma_{diag} & 0 \\ 0 & \Sigma_{s's'} \end{pmatrix} + \begin{pmatrix} \Sigma_s & \Sigma_{ss'}^T \\ \Sigma_{ss'} & \Sigma_{s's'} \end{pmatrix} \right)^{-1} \begin{pmatrix} \Sigma_{xs} & \Sigma_{xs'} \end{pmatrix}^T \Theta.$$

Part i) of the proposition therefore follows from Claim 1.

Part ii) The posterior covariance matrix $\Sigma_x^{p\Theta}$ satisfies

$$\Sigma_x^{p\Theta} = \Sigma_x - \Sigma_{xs}\beta^{\Theta} \left[(\beta^{\Theta})^T \Sigma_s \beta^{\Theta} + \Sigma_{*u}^{\Theta} \right]^{-1} (\beta^{\Theta})^T \Sigma_{xs}^T,$$

where Σ^{Θ}_{*u} and β^{Θ} are the covariance of the liquidity trade and equilibrium trading intensities, for the model with assets Θ. $\Sigma^{p\Theta'}_{x}$ satisfies the analogous equation.

From the definition of λ (see (18) in the proof of Theorem 1),

$$\begin{aligned}\left[(\beta^{\Theta})^T\Sigma_s\beta^{\Theta}+\Sigma^{\Theta}_{*u}\right]^{-1} &= \left[(\lambda^{\Theta})^{-1}\Theta^T\Sigma_{xs}(\Sigma_s+\Sigma_{diag})^{-1}\Sigma_s(\Sigma_s+\Sigma_{diag})^{-1}\Sigma_{xs}^T\Theta(\lambda^{\Theta})^{-1}+\Sigma^{\Theta}_{*u}\right]^{-1}\\ &= \left[(\lambda^{\Theta})^{-1}\Theta^T\Sigma_{xs}(\Sigma_s+\Sigma_{diag})^{-1}\Sigma_{xs}^T\Theta(\lambda^{\Theta})^{-1}\right]^{-1}\\ &= \lambda^{\Theta}\left[\Theta^T\Sigma_{xs}(\Sigma_s+\Sigma_{diag})^{-1}\Sigma_{xs}^T\Theta\right]^{-1}\lambda^{\Theta}.\end{aligned}$$

Consequently, the reduction in the posterior covariance about the factors satisfies

$$\begin{aligned}&\Sigma_{xs}\beta^{\Theta}\left[(\beta^{\Theta})^T\Sigma_s\beta^{\Theta}+\Sigma^{\Theta}_{*u}\right]^{-1}(\beta^{\Theta})^T\Sigma_{xs}^T\\ &\quad=\Sigma_{xs}\left(\Sigma_{diag}+\Sigma_s\right)^{-1}\Sigma_{xs}^T\Theta\left[\Theta^T\Sigma_{xs}(\Sigma_s+\Sigma_{diag})^{-1}\Sigma_{xs}^T\Theta\right]^{-1}\Theta^T\Sigma_{xs}\left(\Sigma_{diag}+\Sigma_s\right)^{-1}\Sigma_{xs}^T\\ &\quad=Y\tilde{\Theta}\left(\tilde{\Theta}^T\tilde{\Theta}\right)^{-1}\tilde{\Theta}^TY,\end{aligned}$$

where

$$\begin{aligned}Y&=\left[\Sigma_{xs}\left(\Sigma_{diag}+\Sigma_s\right)^{-1}\Sigma_{xs}^T\right]^{0.5},\\ \tilde{\Theta}&=Y\Theta.\end{aligned}$$

An analogous equation for the reduction in posterior covariance, under assets Θ'.

Since Y is a positive definite matrix, it suffices to establish that

$$\Theta'\left(\Theta'^T\Theta'\right)^{-1}\Theta'^T\geq\Theta\left(\Theta^T\Theta\right)^{-1}\Theta^T, \tag{23}$$

in case when Θ' adds assets to Θ, $\Theta'_{nm}=\Theta_{nm}$, for $n\leq N, m\leq M$. Part ii) of the proposition therefore follows from Claim 1.

5.5 Proof of Proposition 3

As a preliminary step, we establish the following:

Claim 2 *Fix an information structure and consider two sets of assets, Θ and $\Theta\Gamma$, for an $M\times M$ matrix Γ of rank M. If β and λ are the linear equilibrium with assets Θ, then $\beta\Gamma^{T-1}$ and $\Gamma^T\lambda\Gamma$ are the linear equilibrium with assets $\Theta\Gamma$, and the expected payoffs of each informed trader remain the same.*

Proof. (Claim) Let Σ^{Θ}_{*u} and Σ^{Γ}_{*u} be the two covariance matrices under the two sets of assets. By substituting assets $\Theta\Gamma$ into formula (2), it follows that

$$\Sigma^{\Gamma}_{*u}=\Gamma^{-1}\Sigma^{\Theta}_{*u}\Gamma^{T-1}.$$

Given λ for assets Θ, the price impact for assets $\Theta\Gamma$ follows directly from substituting in the statement of Theorem 1, or, alternatively, in formula (18) in the proof of the Theorem. Direct substitution establishes also the formula for the trade intensities.

Fix trader i with signal vector s_i, who submits order $d_i = \beta_i^T s_i$, for assets Θ. Let w be the vector of expected equilibrium values minus the price impact of other traders, $w = \mathbb{E}[v - \lambda\beta^T s_{-i}]$. Expected profits for player i thus equal $d_i^T(w - \lambda\beta_i^T s_i)$. Substituting the corresponding parameters $\Theta\Gamma$, λ^Γ, and β^Γ, for the new set of assets, establishes that the expected payoffs for trader i remain unchanged. This establishes the claim. ■

Step 1. Towards the proof of the proposition, observe that market makers make zero expected profits in equilibrium. This follows from the condition on prices in the definition of an equilibrium (Definition 1). Consequently, in any information structure and any equilibrium, the sum of informed traders' expected payoffs equals the expected losses of liquidity traders. In a linear equilibrium, expected losses of liquidity traders conditional on a vector $d_L(u)$ of liquidity demands equal $d_L(u^T)\lambda d_L(u)$. Hence the unconditional total expected profits of informed traders Π are

$$\Pi = \mathbb{E}[d_L(u^T)\lambda d_L(u)] = \sum_{m=1}^{M}\left(\sum_{m'=1}^{M}\lambda_{m,m'}(\Sigma_{*u})_{m',m}\right). \tag{24}$$

In particular, suppose that information structure satisfies independent factors and liquidity demands, and the assets are non-overlapping, so that $\Theta^T\Theta$ is a diagonal matrix. In this case, the covariance matrix Σ_{*u} of liquidity demands for the assets is diagonal (see (2)).[19] Thus, the total expected profits simplify to

$$\Pi = \sum_{m=1}^{M}\lambda_{m,m}\sigma_{*u,m}^2. \tag{25}$$

Step 2. Let us establish part i) of the proposition. Given Claim 2, the profits from trading assets Θ are the same as from trading assets $\overline{\Theta}$, with column lengths normalized to one. Since the factors and the liquidity trade are independent, it follows that i) the first-order-optimality conditions have no interdependence across the factors, and that ii) market learns about asset m only from it's own total order, i.e., λ is a diagonal matrix (see Theorem 1). Let β_n be the trade intensity on signal $s_{n,k}$, by any player.

Let[20]

$$X_n = \frac{Cov(x_n, s_{n,k})}{Var(s_{n,k})} = Cor(s_{n,k}, x_n)^2,\ C_n = \frac{Cov(s_{n,k}, s_{n,k'})}{Var(s_{n,k})} = Cor(s_{n,k}, s_{n,k'}).$$

The first-order-condition for trade on signal about a factor n, which affects the value of

[19] As in the case of factor demand, we denote $\sigma_{*u,m} = (\Sigma_{*u})_{m,m}$.

[20] Recall that the variances of the factors are normalized to one.

the asset m is[21]

$$0 = \overline{\Theta}_{n,m} X_n - (I_n - 1)\lambda_{m,m}\beta_n C_n - 2\lambda_{m,m}\beta_n,$$

$$\lambda_{m,m}\beta_n = \frac{\overline{\Theta}_{n,m} X_n}{2 + (I_n - 1)C_n}. \tag{26}$$

From the Projection Theorem, the learning parameter is defined as

$$\lambda_{m,m} = \frac{\sum_{n=1}^{N} I_n \beta_n \overline{\Theta}_{n,m}}{\sigma^2_{*u,m} + \sum_{n=1}^{N}(I_n^2 \beta_n^2 Cov(s_{n,k}, s_{n,k'}) + I_n \beta_n^2 Var(\varepsilon_{n,k}))}$$

$$\lambda^2_{m,m}\sigma^2_{*u,m} = \sum_{n=1}^{N}(I_n \lambda_{m,m}\beta_n \overline{\Theta}_{n,m} - I_n^2 \lambda^2_{m,m}\beta_n^2 Cov(s_{n,k}, s_{n,k'}) - I_n \lambda^2_{m,m}\beta_n^2 Var(\varepsilon_{n,k}))$$

$$= \sum_{n=1}^{N} \frac{\overline{\Theta}^2_{n,m}\big(I_n X_n(2 + (I_n - 1)C_n) - I_n^2 X_n^2 Cov(s_{n,k}, s_{n,k'}) - I_n X_n^2 Var(\varepsilon_{n,k})\big)}{(2 + (I_n - 1)C_n)^2}$$

$$= \sum_{n=1}^{N} \frac{\overline{\Theta}^2_{n,m} I_n X_n\big(2 + (I_n - 1)C_n - I_n C_n - (1 - C_n)\big)}{(2 + (I_n - 1)C_n)^2} = \sum_{n=1}^{N} \frac{\overline{\Theta}^2_{n,m} I_n X_n}{(2 + (I_n - 1)C_n)^2}.$$

The last derivation used the first-order formulas (26), $C_n = X_n Cov(s_{n,k}, s_{n,k'})$, and $1 - C_n = X_n Var(\varepsilon_{n,k})$.

Given the definition of profitabilities π_n, the last displayed line implies that

$$\lambda_{m,m} = \frac{1}{\sigma_{*u,m}} \sqrt{\sum_{n=1}^{N} \overline{\Theta}^2_{n,m}\pi_n^2}.$$

On the other hand, volume of liquidity trade for asset m, $\sigma^2_{*u,m}$, follows from formula (2),

$$\sigma^2_{*u,m} = \sum_{n=1}^{N} \overline{\Theta}^2_{n,m}\sigma^2_{u,n}. \tag{27}$$

Given the formula (25) for the total expected informed profits, the last two expressions establish the step.

Step 3. We show that for any $2 < K < N$

$$max_{\{p_n \geq 0, p_1 + ... + p_K = 1\}} \sqrt{p_1\pi_1^2 + ... + p_K\pi_K^2}\sqrt{p_1\sigma^2_{u,1} + ... + p_K\sigma^2_{u,K}} \leq \sum_{n=1}^{K} \pi_n \sigma_{u,n}, \tag{28}$$

as long as the inequality is satisfied for $K = 2$.

If for any two factors $n, n' \leq K$, $\pi_n \geq \pi_{n'}$ and $\sigma_{u,n} \geq \sigma_{u,n'}$, then the maximum in (28) requires $p_{n'} = 0$. Consequently, we may assume that πs and σ_us have opposite orders: $\pi_1 \leq ... \leq \pi_K$ and $\sigma_{u,1} \geq ... \geq \sigma_{u,K}$, without loss of generality.

[21] Assets are non-overlapping, and so there is a single such asset m

To simplify notation, we also assume below that $K = 3$.[22] Fix an optimal vector (p_1, p_2, p_3) and parameters $x, y \in [0,1], x + y = 1$. We have

$$\left.\frac{d((p_1 - x\varepsilon)\pi_1^2 + (p_2 + \varepsilon)\pi_2^2 + (p_3 - y\varepsilon)\pi_3^2)((p_1 - x\varepsilon)\sigma_{u,1}^2 + (p_2 + \varepsilon) + (p_3 - y\varepsilon)\sigma_{u,3}^2)}{d\varepsilon}\right|_{\varepsilon=0}$$
$$= (\pi_2^2 - x\pi_1^2 - y\pi_3^2)(p_1\sigma_{u,1}^2 + ... + p_3\sigma_{u,3}^2) + (\sigma_{u,2}^2 - x\sigma_{u,1}^2 - y\sigma_{u,3}^2)(p_1\pi_1^2 + ... + p_3\pi_3^2),$$

as well as

$$\left.\frac{d^2((p_1 - x\varepsilon)\pi_1^2 + (p_2 + \varepsilon)\pi_2^2 + (p_3 - y\varepsilon)\pi_3^2)((p_1 - x\varepsilon)\sigma_{u,1}^2 + (p_2 + \varepsilon) + (p_3 - y\varepsilon)\sigma_{u,3}^2)}{(d\varepsilon)^2}\right|_{\varepsilon=0}$$
$$= 2(\pi_2^2 - x\pi_1^2 - y\pi_3^2)(\sigma_{u,2}^2 - x\sigma_{u,1}^2 - y\sigma_{u,3}^2).$$

Let x_A be such that $\pi_2^2 - x_A\pi_1^2 - (1 - x_A)\pi_3^2 = 0$ and, hence, the expression is positive when $x > x_A$, and let x_B be such that $\sigma_{u,2}^2 - x_B\sigma_{u,1}^2 - (1 - x_B)\sigma_{u,3}^2 = 0$ and, hence, the expression is positive when $x < x_B$. It follows that for $x \in [(x_A \vee x_B), (x_A \wedge x_B)]$ the second derivative above is weakly positive.

It follows that if an interior vector (p_1, p_2, p_3) is optimal then $x_A = x_B$, so that the second derivative above is non-positive, for any x, y. But then, with derivatives in ε of all order equal to zero, when $x = x_A, y = 1 - x_A$, there is another optimal vector (p_1', p_2', p_3'), with at least one coefficient on the boundary of the unit interval. Hence, (28) holds as long as the inequality is satisfied for $\underline{N} = 2$.

Step 4. We show that for any weights $p_1, p_2 \geq 0$, $p_1 + p_2 = 1$,

$$\frac{1}{2}[(\pi_1 \wedge \pi_2)(\sigma_{u,1} \wedge \sigma_{u,2}) + (\pi_1 \vee \pi_2)(\sigma_{u,1} \vee \sigma_{u,2})] \leq \sqrt{p_1\pi_1^2 + p_2\pi_2^2}\sqrt{p_1\sigma_{u,1}^2 + p_2\sigma_{u,2}^2}$$
$$\leq (\pi_1 \wedge \pi_2)(\sigma_{u,1} \wedge \sigma_{u,2}) + (\pi_1 \vee \pi_2)(\sigma_{u,1} \vee \sigma_{u,2}), \tag{29}$$

For the left inequality, for $p_1 = p_2 = 1/2$ we have

$$\left(\frac{1}{2}\pi_1^2 + \frac{1}{2}\pi_2^2\right)\left(\frac{1}{2}\sigma_{u,1}^2 + \frac{1}{2}\sigma_{u,2}^2\right) = \frac{1}{4}(\pi_1^2 + \pi_2^2)(\sigma_{u,1}^2 + \sigma_{u,2}^2)$$
$$= \frac{1}{4}[(\pi_1 \wedge \pi_2)(\sigma_{u,1} \wedge \sigma_{u,2}) + (\pi_1 \vee \pi_2)(\sigma_{u,1} \vee \sigma_{u,2})]^2$$
$$+ \frac{1}{4}[(\pi_1 \wedge \pi_2)(\sigma_{u,1} \vee \sigma_{u,2}) - (\pi_1 \vee \pi_2)(\sigma_{u,1} \wedge \sigma_{u,2})]^2.$$

For the right inequality, suppose first that profitabilities and liquidity demands are assortative, i.e., the right-hand side of (29) is $\pi_1\sigma_{u,1} + \pi_2\sigma_{u,2}$. We have

$$(p_1\pi_1^2 + p_2\pi_2^2)(p_1\sigma_{u,1}^2 + p_2\sigma_{u,2}^2) \leq (\pi_1 \wedge \pi_2)^2(\sigma_{u,1} \wedge \sigma_{u,2})^2 \leq (\pi_1\sigma_{u,1} + \pi_2\sigma_{u,2})^2.$$

[22] Otherwise, in what follows, consider analogous perturbations along any three coefficients $n_1, n_2, n_3 \leq K$.

Suppose now that profitabilities and liquidity demands are non-assortative, i.e., the right-hand side of (29) is $\pi_1\sigma_{u,2}+\pi_2\sigma_{u,1}$. We have

$$\begin{aligned}
&(p_1\pi_1^2+p_2\pi_2^2)(p_1\sigma_{u,1}^2+p_2\sigma_{u,2}^2)-(\pi_1\sigma_{u,2}+\pi_2\sigma_{u,1})^2\\
&\quad=-2\pi_1\sigma_{u,1}\pi_2\sigma_{u,2}-(1-p_1p_2)(\pi_1^2\sigma_{u,2}^2+\pi_2^2\sigma_{u,1}^2)+p_1^2\pi_1^2\sigma_{u,1}^2+p_2^2\pi_2^2\sigma_{u,2}^2\\
&\quad\leq-(1-p_1p_2)(\pi_1^2\sigma_{u,1}^2+\pi_2^2\sigma_{u,2}^2)+p_1^2\pi_1^2\sigma_{u,1}^2+p_2^2\pi_2^2\sigma_{u,2}^2\leq 0,
\end{aligned}$$

where the first inequality follows from omitting the first negative term, and non-assortative $(\pi_n,\sigma_{u,n})$ being equivalent to $\pi_1^2\sigma_{u,1}^2+\pi_2^2\sigma_{u,2}^2\leq\pi_1^2\sigma_{u,2}^2+\pi_2^2\sigma_{u,1}^2$, whereas the second inequality follows from $p_1^2+2_2^2\leq 1-p_1p_2$. This establishes (29).

Step 5. We now establish part ii) of the proposition. For any two factors n and n' define $\alpha_{n,n'}$ so that

$$\begin{aligned}
&max_{\{p_n,p_{n'}\geq 0,p_n+p_{n'}=1\}}\sqrt{p_n\pi_n^2+p_{n'}\pi_{n'}^2}\sqrt{p_n\sigma_{u,n}^2+p_{n'}\sigma_{u,n'}^2}\\
&\qquad\qquad=\alpha_{n,n'}[(\pi_n\wedge\pi_{n'})(\sigma_{u,n}\wedge\sigma_{u,n'})+(\pi_n\vee\pi_{n'})(\sigma_{u,n}\vee\sigma_{u,n'})].
\end{aligned}$$

It follows from step 5 that $\alpha_{n,n'}\in[1/2,1]$.

When condition (CORC) is violated for some factors n and n', then it follows from step 2 that there is a way to linearly bundle the two factors into one asset and increase total expected profits of informed traders.

Conversely, suppose the condition (CORC) is satisfied for every two factors, and consider any bundled assets. Given Claim 2, we may assume that the bundled assets are non-overlapping. Step 2 provides the formula for the total expected informed profits, and steps 3 and 4 establish that it is enough to consider assets that bundle only two factors. However, given the definition of parameters $\alpha_{n,n'}$, bundling any two factors may not increase profits, when condition (CORC) holds.

References

Admati, Anat R and Paul Pfleiderer. 1988. “A theory of intraday patterns: Volume and price variability.” *The Review of Financial Studies* 1 (1):3–40.

Allen, Franklin and Douglas Gale. 1991. “Arbitrage, short sales, and financial innovation.” *Econometrica: Journal of the Econometric Society* :1041–1068.

Allen, Franklin, Douglas Gale et al. 1994. *Financial innovation and risk sharing.* MIT press.

Anderson, Robert M and Roberto C Raimondo. 2008. “Equilibrium in continuous-time financial markets: endogenously dynamically complete markets." *Econometrica* 76 (4):841–907.

Andreyanov, Pasha and Tomasz Sadzik. 2021. "Robust mechanism design of exchange." *The Review of Economic Studies* 88 (2):521–573.

Back, Kerry, C. Henry Cao, and Gregory A. Willard. 2000. "Imperfect competition among informed traders." *The Journal of Finance* 55 (5):2117–2155.

Caballe, Jordi and Murugappa Krishnan. 1994. "Imperfect competition in a multi-security market with risk neutrality." *Econometrica: Journal of the Econometric Society* :695–704.

Carvajal, Andrés, Marzena Rostek, and Marek Weretka. 2012. "Competition in financial innovation." *Econometrica* 80 (5):1895–1936.

Colla, Paolo and Antonio Mele. 2010. "Information linkages and correlated trading." *The Review of Financial Studies* 23 (1):203–246.

Cremer, Jacques and Richard P McLean. 1985. "Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent." *Econometrica* 53 (2):345–61.

Dridi, Ramdan and Laurent Germain. 2009. "Noise and competition in strategic oligopoly." *Journal of Financial Intermediation* 18 (2):311–327.

Duffie, Darrell. 2010. *Dynamic asset pricing theory.* Princeton University Press.

Duffie, Darrell and Rohit Rahi. 1995. "Financial market innovation and security design: An introduction." *Journal of Economic Theory* 65 (1):1–42.

Foster, F. Douglas and S. Viswanathan. 1994. "Strategic trading with asymmetrically informed traders and long-lived information." *Journal of Financial and Quantitative Analysis* 29 (4):499–518.

Foster, F Douglas and S Viswanathan. 1996. "Strategic trading when agents forecast the forecasts of others." *The Journal of Finance* 51 (4):1437–1478.

Gul, Faruk and Andrew Postlewaite. 1992. "Asymptotic efficiency in large exchange economies with asymmetric information." *Econometrica* 60 (6):1273–1292.

Hellwig, Martin F. 1980. "On the aggregation of information in competitive markets." *Journal of Economic Theory* 22 (3):477–498.

Holden, Craig W and Avanidhar Subrahmanyam. 1994. "Risk aversion, imperfect competition, and long-lived information." *Economics Letters* 44 (1-2):181–190.

Hugonnier, Julien, Semyon Malamud, and Eugene Trubowitz. 2012. "Endogenous completeness of diffusion driven equilibrium markets." *Econometrica* 80 (3):1249–1270.

Jackson, Matthew O. 1992. "Incentive compatibility and competitive allocations." *Economics Letters* 40 (3):299–302.

Jackson, Matthew O and Alejandro M Manelli. 1997. "Approximately competitive equilibria in large finite economies." *Journal of Economic Theory* 77 (2):354–376.

Klemperer, Paul D and Margaret A Meyer. 1989. "Supply function equilibria in oligopoly under uncertainty." *Econometrica* 57 (6):1243–1277.

Kremer, Ilan. 2002. "Information aggregation in common value auctions." *Econometrica* 70 (4):1675–1682.

Kyle, Albert S. 1985. "Continuous auctions and insider trading." *Econometrica* 53 (6):1315–35.

Lambert, Nicolas S., Michael Ostrovsky, and Mikhail Panov. 2018. "Strategic Trading in Informationally Complex Environments." *Econometrica* 86 (4):1119–1157.

McLean, Richard and Andrew Postlewaite. 2002. "Informational size and incentive compatibility." *Econometrica* 70 (6):2421–2453.

Milgrom, Paul R. 1981. "Rational expectations, information acquisition, and competitive bidding." *Econometrica: Journal of the Econometric Society* :921–943.

Pasquariello, Paolo. 2007. "Imperfect competition, information heterogeneity, and financial contagion." *The Review of Financial Studies* 20 (2):391–426.

Pesendorfer, Wolfgang and Jeroen M Swinkels. 1997. "The loser's curse and information aggregation in common value auctions." *Econometrica* :1247–1281.

Reny, Philip J and Motty Perry. 2006. "Toward a strategic foundation for rational expectations equilibrium." *Econometrica* 74 (5):1231–1269.

Roberts, Donald John and Andrew Postlewaite. 1976. "The incentives for price-taking behavior in large exchange economies." *Econometrica* 44 (1):115–27.

Rostek, Marzena and Marek Weretka. 2012. “Price inference in small markets.” *Econometrica* 80 (2):687–711.

Sadzik, Tomasz and Chris Woolnough. 2021. “Rational destabilization in a frictionless market.” *Journal of Economic Theory* 192:105169.

Vives, Xavier. 2011. “Strategic supply function competition with private information.” *Econometrica* 79 (6):1919–1966.

Wilson, Robert. 1977. “A bidding model of perfect competition.” *The Review of Economic Studies* 44 (3):511–518.

———. 1985. “Incentive efficiency of double auctions.” *Econometrica* 53 (5):1101–1115.