The Magic of Markets: Information Acquisition and Aggregation with Strategic Traders*

Paymon Khorrami† João Thereze‡

March 8, 2025

Abstract

We study financial markets in which strategic traders can acquire information at a cost. In the context of large markets (as the number of traders $n \to \infty$), we develop a sufficient statistic for aggregate information collection that depends on the information technology only through the marginal cost of signal precision at the prior. In contrast to Grossman and Stiglitz (1980), the large-market limit can be informationally efficient, which happens if and only if the marginal cost of signal precision at the prior is zero. Unlike the previous literature, our framework can deliver any level of price informativeness. We obtain closed-form expressions for liquidity, volume, price volatility, and price informativeness, and we show how these measures comove in ways that are at odds with Noisy Rational Expectations Equilibria. Finally, we develop a method to identify time-variation in price informativeness by decomposing changes in observables into changes in fundamental and non-fundamental uncertainty.

*We thank Christian Opp, Cecilia Parlatore, Shri Santosh, and Vish Viswanathan for helpful discussions.

†The Fuqua School of Business, Duke University, Durham, NC 27701 — paymon.khorrami@duke.edu.

‡The Fuqua School of Business, Duke University, Durham, NC 27701 — joao.thereze@duke.edu.

1 Introduction

What determines whether or not financial markets are informationally-efficient? This question poses a challenge to the standard theoretical framework. On the one hand, efficiency is precluded in noisy rational expectations models (Grossman and Stiglitz, 1980): traders would not incur costs to acquire information in a fully efficient market, because they could simply obtain the information from market prices for free; but if no trader acquires information, then markets cannot be informative. In much of the rest of finance, on the other hand, information efficiency is assumed, with the microfoundations and microstructure ignored. Neither of these two approaches is able to address the information efficiency question, making them ill-suited to explain when markets reveal information and what information they reveal.

We provide a simple framework to understand information efficiency. For that, we study information-acquisition in the large-market limit of a strategic trading model (Kyle, 1985; Lambert et al., 2018) rather than a perfectly competitive Noisy Rational Expectations Equilibrium (NREE). This seemingly minor difference leads to a vastly different result. Depending on the information technology, any amount of aggregate information-acquisition and price informativeness can be justified, from completely uninformative to fully-informative prices. By not arbitrarily capping price informativeness, we are able to answer the question of what makes markets efficient and informative. Our framework also bears implications that are substantively different from the previous the literature: price informativeness, market liquidity, and volatility can sometimes exhibit patterns opposite to those in an NREE. Finally, our framework allows us to map these objects to their critical determinants, where we find a special role for the marginal cost of an initial piece of information.

Model and Results. There are n ex-ante identical and risk-neutral strategic traders, as well as some noise traders and an uninformed, competitive, market maker trading a risky asset. Strategic traders have access to an information technology allowing them to learn about the asset value at a cost. After acquiring information, strategic traders submit market orders, anticipating their price impact, and competitive market makers set a break-even price upon observing the order flow. We are interested in the large-market limit of this model ($n \to \infty$), in which individual price impacts vanish, resembling perfectly competitive markets.

In our model, the information technology is represented by the cost of obtaining a signal with precision z above the prior, given by $c(z)$. In small markets, the details of the information technology matter to determine equilibrium. However, we show that when the number of traders approaches infinity, the critical object that characterizes price informativeness is $\chi := c'(0)$—the marginal cost of precision at the prior. More generally, we show that the large-market quantity of aggregate information is an increasing function of the following sufficient statistic:

$$\Gamma := \frac{\text{fundamental volatility} \times \text{noise trader volatility}}{\chi}$$

Our main result is that prices are fully informative if and only if $\chi = 0$. In a large market, each trader ends up acquiring a small amount of information z. Intuitively, this is a consequence of the strategic substitutability of information: when there are more traders who learn the same, prices are more revealing, reducing individual incentives for acquisition. When prices are arbitrarily revealing, information acquisition incentives are virtually absent. Therefore, it is immediate that a necessary condition for efficiency would be that a small piece of information can be obtained at a vanishingly small marginal cost. The key contribution is to show this is also sufficient: regardless of the details of the cost function, as long as the marginal cost of information at the prior is zero, prices will become arbitrarily informative as the market grows. Although each individual trader acquires virtually no information, the aggregate level of information is high. In this case, liquidity and volume are also infinite. Through this result, our paper provides a renewed sense of the "magic of markets" when information is costly.

We proceed to show that our model can produce any level of price informativeness. To that end, we characterize the large-market equilibrium for all values of Γ. As Γ falls, the aggregate level of information declines monotonically and can take any value between 0 and ∞. The fact that the level of information is increasing in both fundamental volatility and noise is intuitive: both uncertainties allow strategic traders to "conceal" their information when trading, allowing greater information rents, thereby augmenting incentives for information acquisition.

Finally, we highlight model predictions which are at odds with NREE. We start by obtaining closed-form expressions for the trading volume, liquidity, price informativeness, and price volatility. The co-movement of these quantities differs markedly with and without strategic incentives. For example, while an increase in noise trading volatility reduces informativeness in NREE models, it always increases informativeness in our model. The reason is that strategic incentives boost information-acquisition incentives when noise is increased, leading to larger trades, which in turn reveal more information. Moreover, liquidity is non-monotonic in information costs in our model: it is decreasing when information costs are low, as in NREE, but it positively co-moves with information cost when those are high. This is because when information costs are high, there is little private information in the market, and therefore very low adverse selection. As adverse selection vanishes, all the order flow is uninformed and the market maker is willing to trade against any quantity of demand at the uninformed price—that is, liquidity is again infinite. These results call attention to a broader implication of our model: the microfoundation of price-taking behavior matters for equilibrium variables.

Applications. We conclude with two applications that illustrate how one could leverage our framework to learn about time-variation in informativeness. First, we show that price volatility can be used to proxy for informativeness under some circumstances, justifying a common empirical practice (e.g., Roll, 1988; Campbell et al., 2023). Specifically, as long as fundamental uncertainty remains fixed, price informativeness is strictly increasing in price volatility. That this monotonicity property fails in NREE models (Dávila and Parlatore, 2023) emphasizes the importance of microfounding competition in addressing questions of market efficiency.

Our second application identifies the dynamics of informativeness when both fundamental and non-fundamental uncertainty vary. Suppose the researcher has access to a high-frequency dataset of price volatility and trading volume, which are two readily available quantities. From this hypothetical data, we show how to obtain time series for (i) the underlying fundamental uncertainty and noise trader volatility; (ii) the market liquidity; and (iii) the level of price informativeness. Thus, our framework allows the researcher to infer the degree of market efficiency. Under the arguably empirically-relevant assumption of information-rich markets, a setting our framework accommodates easily, our identification argument provides a clean set of equations for the evolution of the variables of interest, which can be straightfowardly computed from observables.

Related literature. A vast literature studies whether prices aggregate "dispersed information", starting from the formulation of NREE in Grossman and Stiglitz (1980). These papers investigate what determines price informativeness when information is exogenous (Grossman, 1976; Hellwig, 1980) or endogenous (Verrecchia, 1982). We depart from this literature by defining competitive equilibrium as the large-market limit of a strategic trading model a la Kyle (1985) and, more recently, Lambert et al. (2018).[1] The

[1]Recent contributions to this strategic trading literature include the study of belief disagreement (Han and Kyle, 2018; Kyle et al., 2018); more general payoff distributions (Glebkin et al., 2020); wealth effects in preferences (Glebkin et al., 2023); heterogeneous oligopoly trade (Kacperczyk et al., 2023); the trade-off between speculation and hedging (Lee and Kyle, 2018); and information externalities (Pavan et al., 2022).

trading side of our model is a special case of the latter paper, which we extend by allowing endogenous information acquisition. Our main results show that, by considering large-market limits of strategic equilibria, prices can aggregate dispersed endogenous information, in contrast to Grossman and Stiglitz (1980).

While we study what makes financial markets efficient, Lee and Kyle (2018) and Kyle (1989) are concerned with what makes them perfectly competitive. Their definition of perfect competition is more stringent than ours: it requires investors to trade the same amount in the large-market limit of a strategic market as they would in NREE. Lee and Kyle (2018) show that strategic markets become perfectly competitive if and only if the number of traders grows to infinity *and* the incentives to speculate vanish relative to risk-sharing motives. Similarly, in a model with noise traders in place of risk-sharing needs, Kyle (1989) shows that a strategic model can only become perfectly competitive if the number of strategic traders grows while the amount of noise also explodes. Viewed through their lens, our finding of fully-informative prices can be seen as an instance of imperfect competition vanishing slowly. Importantly, our efficiency result does not require noise to increase as the market grows.

Because the behavior of price informativeness is remarkably different in our strategic-trading model compared to NREE, our work has implications for various strands of the literature, theoretical and empirical. On the theory side, many papers in the NREE tradition examine how policy changes and structural shifts might impact market efficiency (e.g., Vives, 2017, Dávila and Parlatore, 2021, Buss and Sundaresan, 2023). Our result suggest some of their conclusions may depend on the competitive environment. In fact, Kacperczyk et al. (2023) use an oligopoly model to show that equilibrium price informativeness does depend critically on the distribution of market power.

Empirically, there has been a long history of efforts to identify informativeness from other market data. One common method is to proxy informativeness by price variation, particularly its idiosyncratic or firm-specific component (Roll, 1988; Campbell et al., 2023). Whereas this practice my backfire in NREE models (Dávila and Parlatore, 2023), we show it is justified in our strategic-trading setting. A second method to identify informativeness is to consider how well prices forecast future fundamentals (Kothari and Sloan, 1992; Bai et al., 2016; Kacperczyk et al., 2021; Dávila and Parlatore, 2025). In this forecasting method, a key assumption is that the "fundamental" to be learned is precisely a real cash flow (i.e., not inclusive of future prices); by contrast, we do not need to restrict "fundamentals" to be cash flows alone, because our approach does not employ cash flow data. This forecasting method also cannot easily accommodate both time-series and cross-sectional variation without additional assumptions; by contrast, our approach does allow arbitrary time-series and cross-sectional dynamics.

Finally, we complement the research on information aggregation in strategic environments with common value, such as auctions (Atakan and Ekmekci, 2023) and voting (Martinelli, 2006). Like us, these papers obtain conditions for aggregate levels of acquired information to be high even when individual incentives for learning vanish. In the finance literature, three main approaches have also proved successful in circumventing the Grossman-Stiglitz paradox: adding a source of private value for information (Vives, 2011, 2014); preventing traders from conditioning trade on prices (Dubey et al., 1987; Hellwig, 1982; Milgrom, 1981); or explicitly modeling the price-formation process (Jackson, 1991). Our paper falls closer to the third tradition, despite adopting very different assumptions.

2 Model: Strategic Trading with Information Acquisition

We study a static strategic trading model, in the style of Kyle (1985) with endogenous information-acquisition. After information is acquired, there is an ex-post equilibrium that is a special case of the results in Lambert et al. (2018). Information-acquisition is decided ex-ante, subject to a cost for signal precision. In this ex-ante stage, we will study symmetric equilibria, unless otherwise specified.

A single risky asset with random payoff v is trader in a market with n risk-neutral strategic traders. In this market, there are also a (competitive) and uninformed market maker and unmodeled noise traders that submit demand u. We assume that noise trader demand and fundamentals are independently normally distributed with zero mean (as a normalization): $v \sim \text{Normal}(0, \sigma_v)$, $u \sim \text{Normal}(0, \sigma_u)$, and $u \perp v$.

Before trading, traders have an opportunity to collect information about v, which we will describe shortly. After information collection, strategic trader i has access to a signal about the security value, $\theta_i = v + \varepsilon_i$, where ε_i is the noise in the signal. We assume the signals $\theta := (\theta_1, \ldots, \theta_n)$ are jointly normally distributed, with noises which are mutually independent of each other and fundamentals: $\varepsilon_i \perp \varepsilon_j$ and $\varepsilon_i \perp v$. Finally, let σ_i be the variance of θ_i, and note that $\sigma_i \geq \sigma_v$.

The trading protocol is as follows. Each trader submits a market order d_i, and noise trader demand u realizes. The market-maker observes aggregate demand $D = \sum_i d_i + u$ and sets a break-even price given the information in the order flow:

$$p = \mathbb{E}[v|D]. \tag{1}$$

If trader i demands d_i units of the asset and market prices are p, her ex-post profit is:

$$d_i\,(v-p)\,.$$

Trader i chooses d_i to maximize expected profits, conditional on her signal θ_i. Let $\tilde{V}_i$ be the maximized conditional expected profits in equilibrium, i.e.,

$$\tilde{V}_i := \max_{d_i} \mathbb{E}\left[d_i\,(v-p)\mid\theta_i\right] \tag{2}$$

A *linear trading equilibrium* is an equilibrium in which $d_i = \alpha_i\theta_i$ and $p = \beta D$, for some constants α_i and β. Lambert et al. (2018) solve for the unique linear trading equilibrium in this setting, which we describe in the next section.

Prior to trading, information is gathered subject to a cost. For convenience, we will let traders choose a normalization of information precision, z_i. $z_i \in [0,1]$ is defines as follows, using the signal variance σ_i:

$$\sigma_i = \frac{1}{2}\sigma_v\left(z_i^{-1}+1\right) \tag{3}$$

If $z_i = 0$, then $\sigma_i = +\infty$, representing a completely uninformative signal. If $z_i = 1$, then $\sigma_i = \sigma_v$, implying that the signal contains no error at all. Indeed, the precision of an agent's signal is a monotonic function of z_i. The trader solves the following information-gathering problem maximizing her ex-ante expected profits, net of information costs:

$$\max_{z_i\in[0,1]} \mathbb{E}\left[\tilde{V}_i\right] - c(z_i), \tag{4}$$

where $c(z) : [0,1] \mapsto \mathbb{R}_+$ is a continuously differentiable and strictly convex function satisfying $c(0) = 0$, $c'(0) \geq 0$, and $c'(1-) = +\infty$. The latter condition is not necessary for any result, but simplifies the exposition by guaranteeing an interior solution. An *information equilibrium* is a collection $\boldsymbol{z} := (z_i)_{i=1}^n$ such that z_i solves (4) for each i.

2.1 Linear trading equilibrium

For any information choice decision, $\boldsymbol{z}$, we can use the results in Lambert et al. (2018) to compute the unique linear trading equilibrium.

Lemma 1. *(Lambert et al., 2018) There is a unique linear trading equilibrium, which satisfies*

$d_i = \alpha_i \theta_i$ and $p = \beta D$, with $\alpha := (\alpha_i)_{i=1}^n$ and β given by

$$\beta = \sqrt{\frac{A^\top \Sigma_{diag} A}{\sigma_u}} \qquad \text{and} \qquad \alpha = \frac{1}{\beta} A,$$

where

$$\begin{aligned} A &= \Lambda^{-1} \Sigma_{\theta v} \\ \Lambda &= 2\Sigma_{diag} - \sigma_v I + \sigma_v \mathbf{1}\mathbf{1}^\top \end{aligned}$$

and where

$$\begin{aligned} \Sigma_{diag} = Cov(\varepsilon, \varepsilon^\top) &= \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \\ \Sigma_{\theta v} = Cov(\theta, v) &= \sigma_v \mathbf{1} \end{aligned}$$

Armed with Lemma 1, we can write expected trader profits explicitly. Obtaining a convenient expression for this ex-ante indirect utility function is critical to solving the information-gathering problem (4) analytically. The proofs of Lemma 2 and all subsequent theoretical results of this section are contained in Appendix A.

Lemma 2. *In the unique linear trading equilibrium, we may write ex-ante expected profits*

$$V_i := \mathbb{E}[\tilde{V}_i] = \sqrt{\frac{\sigma_v \sigma_u}{2}} \frac{z_i + z_i^2}{(1 + \boldsymbol{z} \cdot \mathbf{1}) \sqrt{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}},$$

where $\boldsymbol{z} := (z_i)_{i=1}^n$ denotes the vector of normalized signal precisions, i.e., $z_i = \left(1 + 2\frac{\sigma_i - \sigma_v}{\sigma_v}\right)^{-1}$.

The indirect utility V_i captures the following, opposing, economic forces. On the one hand, the individual trader benefits from acquiring information, as they can use that information to submit informed market orders, which are not perfectly detected by the market maker due to noise trader demand. Therefore, there is a direct, positive effect of z_i on V_i. On the other hand, there is strategic substitution in information acquisition: the higher the level of private information, the harder it is to disguise informed order flow as uninformed demand, reducing the profitability of private information and, as a consequence, information acquisition incentives. This is evident in the denominator V_i. More deeply, the strategic substitutability in information collection traces back to strategic substitutability in trading behavior (other traders' aggressive trades magnifies the price impact faced by trader i), as discussed in Hellwig and Veldkamp (2009). The degree of strategic substitutability matters for our results. Intuitively, if agents view information-acquisition as only moderately substitutable, then traders will continue collecting it even if all other traders are well-informed, leading to a large quantity of aggregate information.

Before proceeding, let us define the *aggregate information content* as $\sum_{i=1}^{n} z_i = \boldsymbol{z} \cdot \mathbf{1}$. Many substantive properties will depend on whether this aggregate information is large or small. For example, we derive price informativeness below as a strictly increasing function of $\boldsymbol{z} \cdot \mathbf{1}$. In addition, other statistics like excess volatility, volume, and liquidity depend critically on aggregate information.

2.2 Information equilibrium

Next, we describe the information equilibrium, where traders choose $z_i \in [0,1]$ to solve (4). Recall from Lemma 2 the individual trader's ex-ante value $V_i = V(z_i; \boldsymbol{z})$, which we write as a function of z_i and $\boldsymbol{z}$.

Here, we are interested in a symmetric equilibrium, in which $\boldsymbol{z} = z_n^* \mathbf{1}$ for some scalar z_n^*. The subscript n in this case indexes the size of the market, rather than the nth trader. Thus, nz_n^* is the "aggregate information" in the economy. Later, we will be interested in a large market, so we ask what happens to nz_n^* as n grows large. In particular, we denote the situation where $\lim_{n\to\infty} nz_n^* = +\infty$ as "full information" because the pooled version of agents' information yields an infinitely precise signal. Before taking $n \to \infty$, let us write down the symmetric equilibrium conditions. As it will be very important in the remainder of the analysis, let us define

$$\chi := c'(0),$$

the marginal cost of information at the prior—i.e. the cost of the first unit of precision.

Differentiate V with respect to z_i to get

$$\frac{d}{dz_i} V(z_i; \boldsymbol{z}) = \frac{\sqrt{\frac{\sigma_v \sigma_u}{2}}}{(1 + \boldsymbol{z} \cdot \mathbf{1})\sqrt{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}} \left[1 + 2z_i - z_i(1 + z_i) \left(\frac{1}{1 + \boldsymbol{z} \cdot \mathbf{1}} + \frac{\frac{1}{2}(1 + 2z_i)}{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2} \right) \right]$$

Define the function f_n, $n \in \mathbb{N}$, as:

$$f_n(z) := \frac{d}{dz_i} V(z; z\mathbf{1})$$

Then, $f_n(z_n^*)$ is the marginal benefit of information of all n-traders in a trading equilibrium with symmetric information. At an information equilibrium, traders must be on their first-order condition, so z_n^* must satisfy

$$f_n(z_n^*) \leq c'(z_n^*), \quad \text{with equality if} \quad z_n^* > 0. \tag{5}$$

Notice that $\lim_{z \to 0} f_n(z) = +\infty$, so that $\chi = c'(0) < \infty$ guarantees $z_n^* > 0$ for any economy of size n. On the other hand, if $\chi = \infty$, while individual traders' FOCs need not hold with equality, condition (5) still trivially holds with equality, i.e., $f_n(0) = c'(0) = \infty$.

Before studying the large-n limiting economy, let us show that such an asymptotic analysis is, in some sense, necessary to obtain large information-gathering (hence, price informativeness and other related results). To do this, we will verify that, as noise vanishes and, with it, the benefits of private information, the aggregate amount of acquired information collapses to zero.

Proposition 1. *For any n, aggregate information vanishes if noise vanishes, $\lim_{\sigma_u \to 0} n z_n^* \to 0$.*

The remainder of this section studies the large-n limit and shows that information collection remains possible in the limit. This includes the possibility that aggregate information can be arbitrarily large, in contrast to Proposition 1. But first, we recognize the general result that "large" individual information-collection by itself never occurs. Instead, aggregate information efficiency must come from the aggregation of many pieces of small information.

Proposition 2. *As $n \to \infty$, individual information-gathering in a symmetric equilibrium vanishes: $z_n^* \to 0$.*

Proposition 2 results from the strategic substitutability in information-gathering. As previously argued, as the aggregate level of information converges to infinity, it becomes impossible for individual traders to disguise their demand among noise, which leads the benefit of information acquisition to go to zero. Despite the fact that individuals collect vanishingly-small information in a large economy, our main result shows that aggregate information can be arbitrarily large.

Theorem 1. *As $n \to \infty$, aggregate information nz_n^* satisfies the following:*

1. *If $\chi = \infty$, then $nz_n^* \to 0$.*
2. *If $\chi \in (0, \infty)$, then $nz_n^* \to Z^*$, where Z^* is the unique positive solution to $Z(1+Z)^2 = \chi^{-2} \frac{\sigma_v \sigma_u}{2}$.*
3. *If $\chi = 0$, then $nz_n^* \to \infty$.*

Theorem 1 delivers the key result that any level of aggregate information is achievable in the large market equilibrium. It shows that aggregate information is an increasing function of the following sufficient statistic:

$$\Gamma := \frac{\sqrt{\sigma_v \sigma_u}}{\chi}. \tag{6}$$

By moving χ from 0 to ∞, the statistic Γ spans all possibilities, and so similarly Z^* can take any value. Note that the equilibrium level of aggregate information is independent of the details of information cost, $c(\cdot)$, up to the marginal cost of information around the prior, χ. This has two major implications. The first one is substantive: it is possible to generate a very high level of price informativeness even when it is arbitrarily expensive to produce a precise signal, as long as the first piece of information is relatively cheap. The second one is technical: as information costs are difficult to observe, this detail-free characterization of the level of information is advantageous, compared to expressions that rely heavily on the shape of $c(\cdot)$—which arise in the case of NREE models. Theorem 1 adds to the literature by establishing an equilibrium with arbitrarily large information, and completely characterizing informativeness as a function of χ.

It is instructive to understand why $\chi = 0$ is necessary and sufficient for an arbitrarily informative equilibrium. Although the details in Theorem 1 naturally depend on the simple environment we study, the possibility of full information is more general. First, the necessity of $\chi = 0$ is straightforward: as the level of aggregate information goes to infinity, information swamps the effect of noise and traders cannot hide their informed trades anymore, leading the value of information to decline to zero. Therefore, the only way that aggregate information can be obtained in the limit is if it is always worthy to acquire a small bit of information, even when the profits of doing so approach zero—implying that the cost of a small piece of information must be arbitrarily small. On its turn, sufficiency follows from a very general property of trading equilibria: the marginal value of the first piece of information remains positive for any level of aggregate information. Under this property and sufficient continuity of the marginal value of information, one can obtain that any sequence of symmetric-information equilibria converges to an unbounded level of acquired information.[2]

Fixed information costs. For comparison, Appendix B studies the same model with a fixed information cost. With fixed costs, an equilibrium with a large quantity of aggregate information *can only emerge if the entire cost function vanishes*. This stands in stark contrast to our result here that only the marginal cost at zero matters. The reason: with fixed costs, the number of informed traders stabilizes as $n \to \infty$, and so the quantity of aggregate information stabilizes too.

How fast does aggregate information grow? In the results above, what matters for the speed of aggregate information accumulation is the rate at which information-collection vanishes at the individual level. We characterize this rate of convergence in Appendix A.6. This then directly tells us how fast aggregate information nz_n^* explodes as the market grows. For example, we show that quadratic information costs imply aggregate information explodes at the same rate as $n^{3/5}$. A takeaway from this analysis is that, if $\chi = 0$, full information can obtain in a large economy even if noise vanishes as the economy grows.[3]

3 Liquidity, Efficiency, Volatility, and Volume

This section accomplishes two main goals. First, it establishes how important market quantities—liquidity, price informativeness, price volatility and trading volume—co-move in our model, which is helpful to understand the implications of our framework to the functioning of markets. Second, it contrasts the behavior of those variables in our large-limit equilibrium with their behavior in NREE models.

[2]To see that this property is sufficient, let $\frac{d}{dz}V_n(z;Z)$ denote the individual trader marginal value function in a symmetric equilibrium with n traders, where z is the trader's private information choice and Z is the aggregate information. For any compact set, assume the sequence $(\frac{d}{dz}V_n)$ is uniformly convergent and each of the functions is continuous. Finally, assume $\inf_n \frac{d}{dz}V_n(0;Z) > 0$, for any finite Z.

If (a subsequence of) the aggregate level of information converges to a finite constant, i.e., $nz_n \to Z^* < \infty$, and thus $z_n \to 0$, we obtain the following contradiction: $0 < \lim_{n\to\infty} \frac{d}{dz}V_n(0;Z^*) = \lim_{n\to\infty} \frac{d}{dz}V_n(z_n;nz_n) = \lim_{n\to\infty} c'(z_n) = c'(0) = 0$, where the first limit holds by uniform convergence of $\frac{d}{dz}V_n$, and the second equality follows since $c'(0) = 0$ implies that $z_n > 0$ along any sequence.

[3]In particular, a model with $\chi = 0$ has full information emerging as $n \to \infty$ even if noise $\sqrt{\sigma_u}$ vanishes at any rate slower than $n^{-3/2}$. This sharply contrasts with the case $\chi \in (0,\infty)$, in which case vanishing noise implies vanishing information. For related reasons, García and Urošević (2013) and Kovalenkov and Vives (2014) resort to the study of large-market limits where noise *grows large with the market* rather than vanishing or remaining constant.

We start by deriving a set of results for prices, demand, and other objects that hold for any symmetric equilibrium z_n^*.

Lemma 3. *Let z_n^* denote a symmetric equilibrium information acquisition in the economy with n strategic traders. Then,*

$$\begin{aligned}
\beta_n &= \sqrt{\frac{\sigma_v}{2\sigma_u}}\sqrt{\frac{nz_n^*(1+z_n^*)}{(1+nz_n^*)^2}} \\
D_n &= \sqrt{\frac{2\sigma_u}{\sigma_v}}\sqrt{\frac{nz_n^*}{1+z_n^*}}\Big(\frac{1}{n}\sum_{i=1}^n \theta_i\Big) + \sqrt{\sigma_u}e_u \\
p_n &= \frac{nz_n^*}{1+nz_n^*}\Big(\frac{1}{n}\sum_{i=1}^n \theta_i\Big) + \sqrt{\frac{\sigma_v}{2}}\frac{\sqrt{nz_n^*(1+z_n^*)}}{1+nz_n^*}e_u,
\end{aligned}$$

where $e_u = u/\sqrt{\sigma_u} \sim Normal(0,1)$. Let $Z^* := \lim_{n\to\infty} nz_n^* \in [0,\infty]$ be the large-n limit of aggregate information. Then, we have that

$$\frac{1}{n}\sum_{i=1}^n \theta_i \to v + \sqrt{\frac{\sigma_v}{2Z^*}}e_\theta,$$

in distribution, where $e_\theta \sim Normal(0,1)$ is independent of v and u. Consequently, the large-n limiting equilibrium objects are, in distribution,

$$\begin{aligned}
\lim_{n\to\infty} \beta_n &= \sqrt{\frac{\sigma_v}{2\sigma_u}}\frac{\sqrt{Z^*}}{1+Z^*} \\
\lim_{n\to\infty} D_n &= \sqrt{2\sigma_u}\Big[\sqrt{Z^*}\frac{v}{\sqrt{\sigma_v}} + \frac{e_\theta + e_u}{\sqrt{2}}\Big] \\
\lim_{n\to\infty} p_n &= \sqrt{\sigma_v}\frac{\sqrt{Z^*}}{1+Z^*}\Big[\sqrt{Z^*}\frac{v}{\sqrt{\sigma_v}} + \frac{e_\theta + e_u}{\sqrt{2}}\Big].
\end{aligned}$$

In Appendix B.1, we also prove an analogous version of Lemma 3 for an asymmetric equilibrium where each trader chooses to be fully-informed with probability π_n^*, which is the type of equilibrium that arises with fixed information costs. The results are identical.

We now define the following key measures. For simplicity, we define them directly in the large-n economy, as all of our analysis will be done there.

Definition 1. Liquidity $\mathcal{L}$ *is the inverse of the price impact of demand, i.e.,*

$$\mathcal{L} := \frac{1}{\lim_{n\to\infty}\beta_n}.$$

Definition 2. Price informativeness $\mathcal{I}$ *is the precision of the fundamental given the price, i.e.,*

$$\mathcal{I} := \frac{1}{Var\big[v \mid \lim_{n\to\infty} p_n\big]}$$

Definition 3. Price volatility $\mathcal{V}$ *is the unconditional variance of the price, i.e.,*

$$\mathcal{V} := Var\big[\lim_{n\to\infty} p_n\big]$$

Definition 4. Trading volume $\mathcal{D}$ *is the expected absolute equilibrium demand, in dollars:*

$$\mathcal{D} := \mathbb{E}\big[\big|\lim_{n\to\infty} p_n D_n\big|\big]$$

By combining the results of Lemma 3 with the definitions above, and using properties of the bivariate normal distribution, we immediately obtain the next proposition.

Proposition 3. *In the large-n limit of symmetric equilibria, the measures of liquidity, price informativeness, price volatility, and trading volume are given by the following:*

$$\begin{aligned}
\textit{(liquidity)} \quad & \mathcal{L} = \sqrt{\frac{2\sigma_u}{\sigma_v}} \frac{1+Z^*}{\sqrt{Z^*}} \\
\textit{(informativeness)} \quad & \mathcal{I} = \frac{1+Z^*}{\sigma_v} \\
\textit{(volatility)} \quad & \mathcal{V} = \sigma_v \frac{Z^*}{1+Z^*} \\
\textit{(volume)} \quad & \mathcal{D} = \sqrt{2\sigma_u \sigma_v Z^*}
\end{aligned}$$

We use the results above to characterize markets across various specifications for information costs, noise, and fundamental uncertainty.

The magic of markets. In a full-information economy, $Z^* = \infty$, the results of Proposition 3 specialize to $\mathcal{L} = \infty$ (perfect liquidity), $\mathcal{I} = \infty$ (fully informative prices), $\mathcal{V} = \sigma_v$ (all information embedded in prices), and $\mathcal{D} = \infty$ (infinite volume). These results hold regardless of the amount of fundamental uncertainty σ_v and the amount of noise σ_u, even if they are arbitrarily small. We refer to our collection of results, which arise when $\chi = 0$ characterizes information costs, as restoring the "magic of markets." If $\chi = 0$, markets generically aggregate dispersed information, despite information being endogenous.

Given the novelty of this "magical markets" outcome, let us pause to highlight the differences from perfectly competitive noisy rational expectations equilibria (NREEs).

Appendix C studies a version of such an NREE with endogenous information, as in Verrecchia (1982), and shows that $\mathcal{L} = \infty$ (perfect liquidity), $\mathcal{I} = \infty$ (fully informative prices), and $\mathcal{V} = \sigma_v$ (all information embedded in prices) only arise in some limiting cases. In particular, the magical markets outcomes only arise in the risk-neutral limit, or if $\sigma_v \to 0$, or if $\sigma_u \to 0$ (and additionally the information cost function has $c'(0) = 0$). That is, no information cost leads to generically magical markets in an NREE.[4]

Markets under partial efficiency. We continue the analysis by next considering the partial efficiency case in which $\chi > 0$. We investigate the inter-relationships between the various market-based measures through a series of exercises. To do this, we will use the expressions for the measures in Proposition 3, along with Theorem 1's characterization of Z^* as a monotonic function of the sufficient statistic $\Gamma := \frac{\sqrt{\sigma_v \sigma_u}}{\chi}$.

First, Figure 1 illustrates how the market measures vary with $\chi = c'(0)$. As χ increases, holding all else equal, aggregate information Z^* falls. The patterns in the other market-based measures follow: liquidity $\mathcal{L}$ is U-shaped in χ, while informativeness $\mathcal{I}$, volatility $\mathcal{V}$, and volume $\mathcal{D}$ are strictly decreasing.

For instance, liquidity $\mathcal{L}$ is U-shaped in χ for the following reason. High information-acquisition (low χ) implies a minimal amount of adverse selection because noise traders are impounding a minimal impact on the price relative to informed traders. Market makers know this and set the price appropriately, leading to liquid markets. Low information-acquisition (high χ) also means minimal adverse selection and liquid markets, because the market maker is sure that most trades are uninformed.

Let us compare these results to a competitive NREE with endogenous information. Additional details on this NREE are contained in Appendix C, including our calculation of corresponding market measures for price informativeness $\mathcal{I}$, volatility $\mathcal{V}$, and liquidity $\mathcal{L}$ (in the NREE, volume is not well-defined). Note that the NREE additionally has a risk aversion parameter γ that we set to 3 always.

[4] These limiting results are already known in one way or another. For instance, regarding price informativeness in competitive NREEs: either equilibrium ceases to exist for σ_u small enough (Grossman and Stiglitz, 1980), or equilibrium informativeness is bounded above for all σ_u (Verrecchia, 1982). Verrecchia (1982) says, for instance, "As V [equivalent to our σ_u] approaches infinity, price communicates no information despite traders' corresponding increased information acquisition activities. As V approaches zero, only the most risk tolerant of traders will continue to acquire information because of the increased informativeness of price; at some point the private incentives to acquire information are eliminated, which implies the nonexistence of a competitive equilibrium. Therefore, the informativeness of price is bounded away from infinity even as noise goes to zero" (p. 1425). The reason we, unlike Verrecchia (1982), find that price informativeness can be unbounded in an NREE as $\sigma_u \to 0$ is that we allow information costs to satisfy $c'(0) = 0$. In that case, it is true that information-collection vanishes as $\sigma_u \to 0$, but the noise in the price vanishes faster, allowing fully-revealing prices asymptotically.

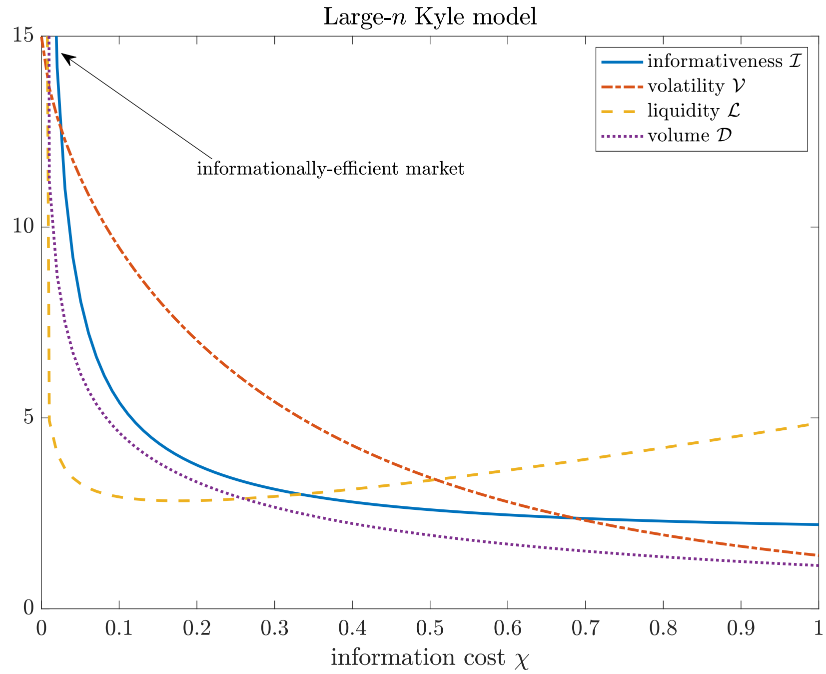

Figure 1: Measures as a function of $\chi = c'(0)$. The volatility measure is scaled to fit on the same scale as the other measures. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$.

Figure 2 displays the NREE for two different information cost functions:

$$\text{NREE, left panel:} \quad c(z) = \frac{\kappa}{2}\Big(\frac{z}{1-z}\Big)^2$$

$$\text{NREE, right panel:} \quad c(z) = \chi\Big(\frac{z}{1-z}\Big)$$

The left panel thus uses a cost function that satisfies $c'(0) = 0$ for any κ, and yet "magical markets" do not emerge, except as $\kappa \to 0$. To be clear, this is substantially more extreme than our strategic-trading model with $\chi \to 0$, because κ scales *the entire cost function*; essentially, taking $\kappa \to 0$ reduces all information costs to zero. The right panel uses a cost function that satisfies $c'(0) = \chi$, hence serving as a reasonable comparison to the strategic-trading model with partial efficiency (again, this comparison is not perfect because the parameter χ in the strategic-trading model is only the marginal cost at $z = 0$, whereas χ in the NREE scales the entire information technology). Despite these differences in the left and right panels' cost functions, the results are qualitatively identical: in an NREE, price informativeness and liquidity are falling in the information cost, while volatility is rising except in highly efficient markets. Comparing this to the strategic-trading model in Figure 1, the behavior of liquidity and volatility is *exactly the opposite in an NREE*, except in the case of highly efficient markets (i.e., when information costs are low). Thus, one potential "test" of our model, distinguishing it from an NREE, is to

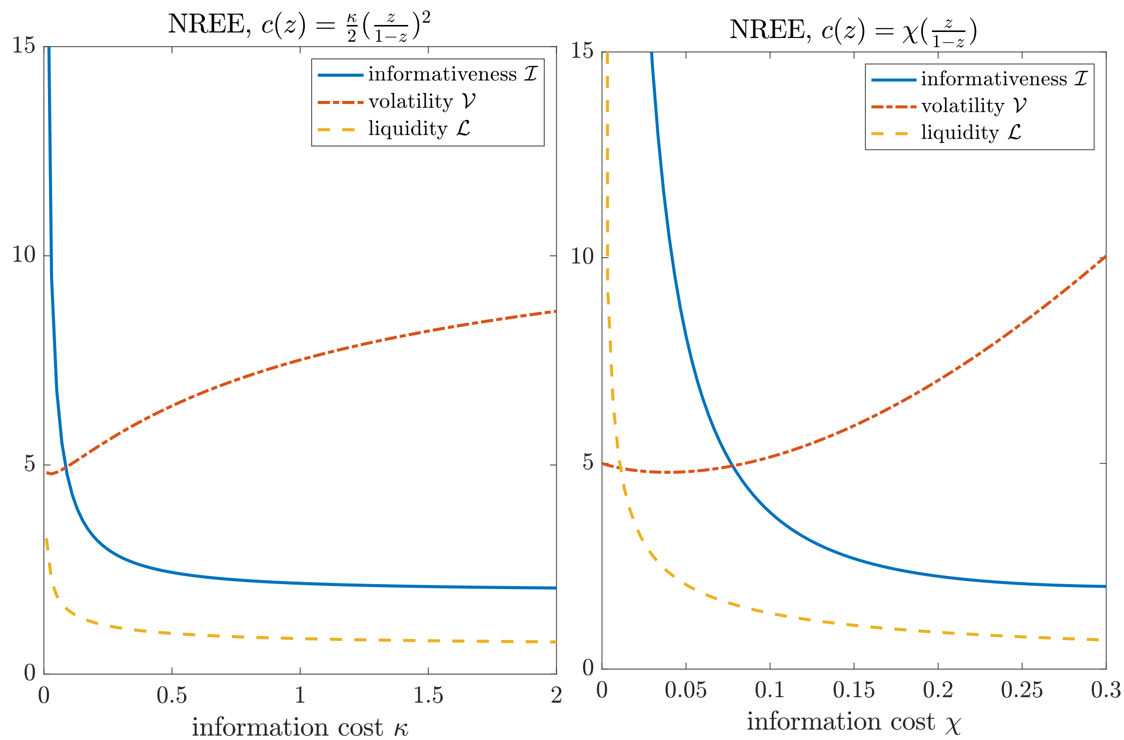

Figure 2: Measures as a function of information costs in an NREE. In the left panel, information costs are parameterized by κ. In the right panel, information costs are parameterized by χ. The volatility measure is scaled to fit on the same scale as the other measures. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, $\gamma = 3$ (risk aversion). Note that equilibrium in the right panel requires $\chi\gamma < 1$, where γ is risk aversion, which is why the upper limit of χ is 0.3. See Appendix C for more details on the NREE.

identify a shock to information technology and observe the response of volatility and (if it is available) a price-impact measure of liquidity.

Price informativeness and volatility. We now focus on price informativeness, because it is a natural measure of market efficiency. Given that informativeness is not directly observable, how might one infer it from market data?

A common thread in the various empirical literatures is that price variation, particularly its idiosyncratic or firm-specific component, proxies well for price informativeness. Dávila and Parlatore (2023) challenge this approach, arguing that price volatility and price informativeness are non-monotonically related in a class of NREEs. We show next that considering strategic trading reestablishes the common empirical wisdom: price volatility $\mathcal{V}$ is positively correlated with informativeness $\mathcal{I}$ under some conditions.[5]

[5]Technically, our price informativeness measure differs from the one used in Dávila and Parlatore (2023), but this is immaterial. They first define an "unbiased signal about the innovation to the asset payoff contained in the price" (p. 554). In our large-n strategic-trading model this signal would be $\tilde{p} := v + \sqrt{\frac{\sigma_v}{Z^*}}\left(\frac{e_\theta + e_u}{\sqrt{2}}\right)$. (In their notation, this is denoted by π.) Based on this price-based signal, they then

In our strategic-trading model, the relation between $\mathcal{I}$ and $\mathcal{V}$ is straightforward. Combining the expressions for the two objects, we obtain

$$\mathcal{V} = \sigma_v - \frac{1}{\mathcal{I}} \tag{7}$$

Since σ_v is independently present in equation (7), let us ignore it for now as a potential driver, instead considering only variation caused by shifts in σ_u or χ. Ignoring σ_v, equation (7) immediately implies that volatility is strictly increasing in price informativeness when the shifter is either σ_u or χ or a combination of the two. Dávila and Parlatore (2023) perform an analogous exercise in an NREE and, in contrast to our model, report a non-monotonic $\mathcal{V}$-$\mathcal{I}$ relation that we will reproduce shortly (see their Propositions 2-3).[6]

Figure 3 displays the volatility-informativeness relation in our model (left panel) and an NREE (middle and right panels). The left panel documents the result proved above: whether the driver is σ_u or χ, the $\mathcal{V}$-$\mathcal{I}$ relation is a unique strictly increasing curve. The middle panel, which is an NREE with exogenous information a la Dávila and Parlatore (2023), displays the non-monotonicity inherent to NREEs. (To make this plot, note that we have separately shifted noise σ_u and signal precision z^*, which is exogenously given for that model.)

There are two key differences between our strategic-trading model and an exogenous-information NREE. One difference is the competitive environment (strategic trading versus perfect competition), and the other difference is the information-acquisition (endogenous versus exogenous). To pinpoint the source of discrepancy in the $\mathcal{V}$-$\mathcal{I}$ relations, we add endogenous information to the NREE. The right panel of Figure 3 shows the result: a very similar non-monotonic pattern emerges in an endogenous-information NREE.[7] In

define informativeness as

$$\tilde{\mathcal{I}} := \mathrm{Var}[\tilde{p} \mid v]^{-1} = \frac{Z^*}{\sigma_v}$$

(In their notation, this is denoted by τ_π.) Notice that $\tilde{\mathcal{I}}$ is one-to-one with our measure $\mathcal{I} = \frac{1+Z^*}{\sigma_v}$. Thus, the discrepancy in our informativeness measures cannot explain the opposite volatility-informativeness relation we find relative to Dávila and Parlatore (2023).

[6]In unreported results, we examine how $\mathcal{V}$ and $\mathcal{I}$ co-move when σ_v is the shifter. We show that the two measures always move inversely, across both the strategic-trading and NREE models. Hence, any discrepancies between these classes of models must come from the other drivers, namely noise and information technology. Section 4 allows both σ_u and σ_v to be time-varying and shows how to extract price informativeness under such conditions.

[7]To create the third panel of Figure 3, we have used the information cost function $c(z) = \frac{\kappa}{2}(\frac{z}{1-z})^2$, which satisfies $c'(0) = 0$. In earlier exercises involving the NREE, we also explored the alternative cost function $c(z) = \chi(\frac{z}{1-z})$, which satisfies $c'(0) = \chi$. The reason for picking this particular cost function here is purely technical and aesthetic: if $c'(0) > 0$ in an NREE, then price informativeness $\mathcal{I}$ as a function of σ_u will be bounded. In that case, the non-monotonicity in the $\mathcal{V}$-$\mathcal{I}$ relation would not easily be observed graphically,

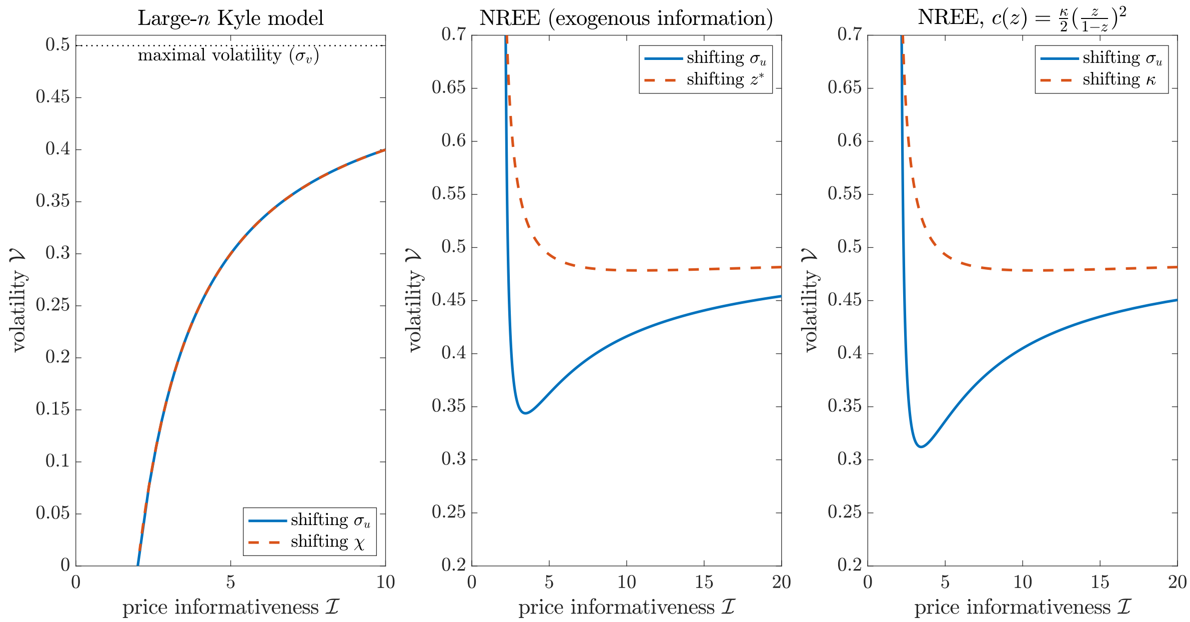

Figure 3: Volatility-informativeness relation. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, $\chi = 0.1$. In the NREE, risk aversion is $\gamma = 3$ and the information cost parameter is $\kappa = 1$. For the NREE with exogenous information, the signal precision z^* is held fixed whenever σ_u varies. See Appendix C for more details on the NREE.

summary, our $\mathcal{V}$-$\mathcal{I}$ relation is substantively different from similar results in the NREE literature, and the source of discrepancy is the nature of competition in trade.

To explore this differential competition in more detail, let us now dive into one important comparative static behind Figure 3, namely the two environments' responses to noise σ_u. Figure 4 plots the various market measures by varying σ_u in the strategic-trading model (left panel) versus the NREE (right panel).

A critical observation from Figure 4 is that *price informativeness behaves oppositely* across the two environments. In an NREE, more noise increases information collection but not sufficiently to offset the direct effect of noise introduced into the price (see Verrecchia, 1982, Corollary 3). In our model, noise similarly raises incentives to collect information, but it also scales strategic incentives to trade more aggressively on information. The second force, unique to strategic-trading models, increases trader profits, which then further encourages more information-acquisition ex-ante. A way to see this second force mathematically is to notice that the price p_n in Lemma 3 is independent of σ_u, holding information fixed; as more noise traders arrive, strategic traders trade aggressively enough to keep the information content in prices fixed. Because of this aggressive trading, the net effect of noise on prices comes solely through information, meaning $\mathcal{I}$ is increasing in σ_u via Z^*.

since the solid blue curve would simply be truncated at its nadir.

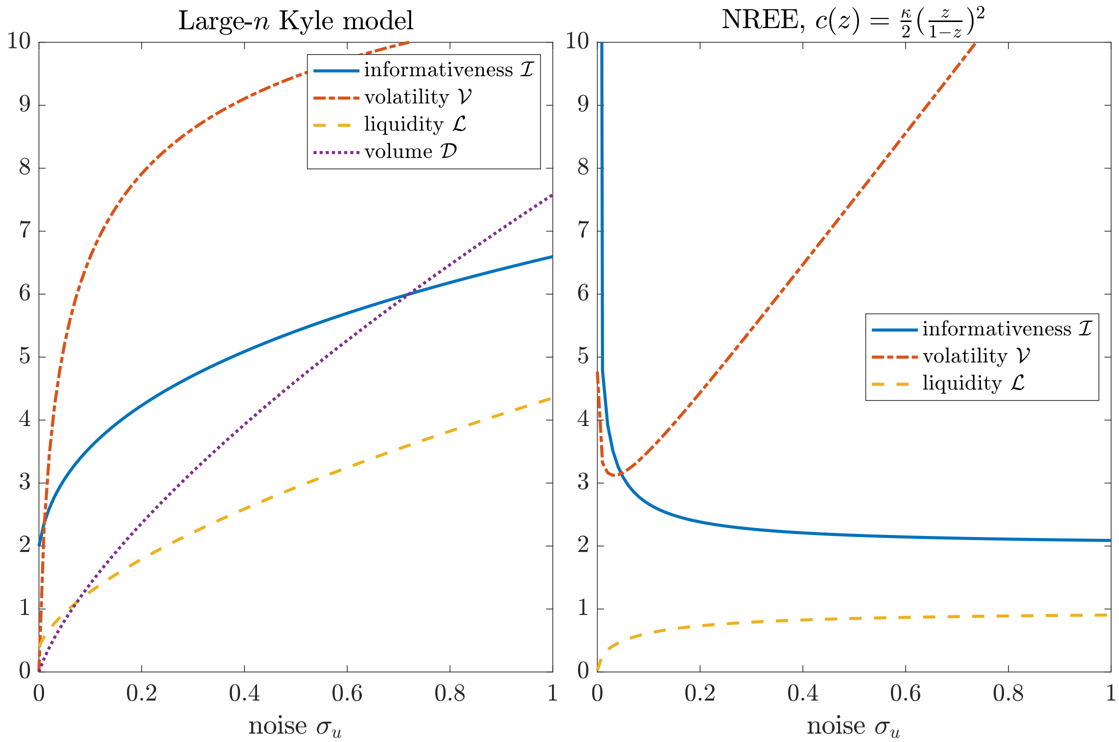

Figure 4: Measures as a function of σ_u. The volatility measure is scaled to fit on the same scale as the other measures. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, $\chi = 0.1$. In the NREE, risk aversion is $\gamma = 3$ and the information cost parameter is $\kappa = 1$. See Appendix C for more details on the NREE.

Now, consider price volatility $\mathcal{V}$ in Figure 4. In our model, higher σ_u unambiguously raises information Z^* and thus unambiguously raises $\mathcal{V}$. In an NREE, higher σ_u can have an ambiguous effect on $\mathcal{V}$, but we know that it must increase volatility at high enough levels of noise. At such high levels of noise, further increases in σ_u raise volatility $\mathcal{V}$ in both the Kyle and NREE models but push informativeness $\mathcal{I}$ in different directions. By contrast, at low levels of noise, volatility responds oppositely to σ_u in the Kyle and NREE models, meaning the volatility-informativeness relations coincide between the environments only because both volatility and informativeness behave oppositely.

4 Identifying Price Informativeness

The discussion on the relationship between price volatility and informativeness, at the end of the last section, reveals an important tension. On one hand, as shown in Figure 3, if changes in price volatility are due to non-fundamental noise, we can identify from these changes whether prices have become less or more informative. On the other hand, if changes to volatility arise from movements in fundamental uncertainty, one could derive the wrong conclusions: indeed, when σ_v increases, our model predicts a *decrease* in informativeness. A natural question, then, is how can one use observable data to identify price informativeness in our model? In this section, we provide an answer to that question that differs from the extant literature.

Disentangling changes in fundamental and non-fundamental uncertainty is essential to measure changes in informativeness. In what follows, we will demonstrate how to decompose variation in widely available data into these underlying, unobservable shocks. To formalize this approach, we consider an econometrician who can observe a dataset consisting of time series of trading volume $\mathcal{D}$ and price volatility $\mathcal{V}$. The goal of the econometrician is to estimate, from the shifts in the two available variables, shocks to fundamental and non-fundamental volatility, as well as changes in price-informativeness—i.e., variation in σ_v, σ_u and $\mathcal{I}$, respectively.

A first observation is that price volatility, trading volume, and liquidity are tightly connected by the shape of the pricing function. Indeed, the expressions in Proposition 3 imply

$$\mathcal{L} = \frac{\mathcal{D}}{\mathcal{V}}.$$

Given data on volatility and volume, we can use our model to infer liquidity, even though it may be hard to measure directly.

Going forward, it is thus equivalent to assume that the dataset contains observations of liquidity and volatility, rather than volume and volatility. Formally, we will define the set of observables to be $\{\mathcal{L}_t, \mathcal{V}_t\}_{t\in\{0,1,\ldots,T\}}$. Our results will rely on two identifying assumptions, imposed in sequence. The first assumption allows us to describe exactly the evolution of the variables of interest as a function of the evolution of observables, up to the level of aggregate precision in the market, which is unobservable.

Assumption 1. *Information costs—χ—do not change throughout the dataset.*

Assumption 1 allows us to attribute any changes in observables to variation in fundamental and non-fundamental noise. It states that while beliefs about fundamentals and non-fundamental demand may change at a high frequency, the same does not hold for information costs, which tend to move only slowly.

Proposition 4. *Under Assumption 1, the evolution of informativeness, fundamental uncertainty,* *non-fundamental noise, and aggregate information follows:*

$$d\log\mathcal{I} = -\frac{Z}{1+2Z}d\log\mathcal{V} + \frac{1+Z}{1+2Z}d\log\mathcal{L}, \quad \textit{where} \qquad \mathcal{I}_0 = \frac{Z_0}{\mathcal{V}_0} \tag{8}$$

$$d\log\sigma_v = \frac{2Z}{1+2Z}d\log\mathcal{V} - \frac{1}{1+2Z}d\log\mathcal{L}, \quad \textit{where} \qquad \sigma_{v,0} = \frac{1+Z_0}{Z_0}\mathcal{V}_0 \tag{9}$$

$$d\log\sigma_u = \frac{1+Z}{1+2Z}d\log\mathcal{V} + \frac{2+3Z}{1+2Z}d\log\mathcal{L}, \quad \textit{where} \qquad \sigma_{u,0} = \frac{\mathcal{V}_0\mathcal{L}_0^2}{2(1+Z_0)} \tag{10}$$

$$d\log Z = \frac{1+Z}{1+2Z}d\log\mathcal{V} + \frac{1+Z}{1+2Z}d\log\mathcal{L}, \quad \textit{given} \qquad Z_0 \tag{11}$$

Proposition 4 provides a method to calculate precisely how price informativeness, uncertainty, and aggregate information vary using observable variables. Given the observable dataset $\{\mathcal{L}_t, \mathcal{V}_t\}$, the only additional object one needs to implement Proposition 4 is the initial level of aggregate information, Z_0. While Z_0 may be hard, if not impossible, to observe or estimate directly, one can perform sensitivity analyses: sample the input Z_0 from a statistical prior, compute paths for $(\mathcal{I}_t, \sigma_{v,t}, \sigma_{u,t}, Z_t)_{t\geq 0}$, and then ask whether the range of paths (one path for each Z_0) shares some common properties. Of course, the outcome depends on the data for $\{\mathcal{L}_t, \mathcal{V}_t\}$ as well as the choice of prior for Z_0, but the process is easily implementable with Proposition 4.

What are some reasonable choices for Z_0? To get a sense, consider the following back-of-the-envelope calculation. Suppose there are $n = 100$ traders, each of whom obtains information with signal-to-noise ratio of $\frac{1}{50}$ (this is $\frac{z}{2(1-z)}$ in the model). In this case, which is a case of a small market with small information-collection, total information is $Z = nz = 100 \times \frac{1}{26} \approx 4$. Using calculations of this type, one can reason toward a lower bound for Z_0.

Conversely, Z_0 could be very high. A second identifying assumption allows us to approximate the dynamics of informativeness in an information-rich market.

Assumption 2. *The level of aggregate information in the market is high—i.e. $Z \approx \infty$.*

Corollary 1. *Under Assumption 1 and Assumption 2, the evolution of informativeness, funda-*

mental uncertainty, and non-fundamental noise can be approximated as follows:

$$d \log \mathcal{I} = -\frac{1}{2} d \log \mathcal{V} + \frac{1}{2} d \log \mathcal{L} + o(Z^{-1}) \tag{12}$$

$$d \log \sigma_v = d \log \mathcal{V} + o(Z^{-1}) \tag{13}$$

$$d \log \sigma_u = \frac{1}{2} d \log \mathcal{V} + \frac{3}{2} d \log \mathcal{L} + o(Z^{-1}) \tag{14}$$

These expressions allow for immediate implementation of the empirical strategy on readily available data. Assumption 2 begs the question of what assuming high Z entails about individual traders' information level. Another back-of-the-envelope calculation is as follows: consider what level of individual signal precision one needs to obtain $\frac{Z}{1+2Z} > 0.495$ (which ensures that the approximation of $d \log \mathcal{I}$ deviates at most 1% from its exact value). To obtain this level of aggregate information $Z \approx 50$ in a market with $n = 1000$ traders, each trader's information needs only to have a signal-to-noise ratio of $\frac{1}{40}$. Even with the previous back-of-the-envelope calculation with $Z \approx 4$, the exact expression for $d \log \mathcal{I}$ in only deviates from its high-Z approximation by 5.5%. Thus, Corollary 1 gives a reasonably accurate representation of informativeness dynamics.

It is instructive to compare our results with Dávila and Parlatore (2025). That paper develops a strategy to identify the level of price informativeness by running a regression of current prices on future asset payoffs, a valid approach in a large class of models—including ours—when primitives are assumed to be constant. Our approach differs from theirs in two ways. First, we are interested in identifying time-variation in price informativeness, which arises exactly from changes in primitives: our key assumption is that movements in informativeness follow from high-frequency shifts in fundamental uncertainty and non-fundamental noise. Second, instead of using the level of prices and asset payoffs, we use two observable quantities—volatility and volume—to disentangle the effects of each shock on informativeness.

5 Conclusion

We examine information efficiency in the large-market limit of a strategic trading model. Our key theoretical results characterize explicitly the mapping between information technology and aggregate information. In particular, if the marginal cost of information is zero at the prior, then aggregate information is infinitely-large and prices are efficient. We think of this result as restoring the "magic of markets" because it captures the wellunderstood potential for prices to aggregate dispersed information, but in a setting with endogenous information-acquisition. This core result also implies our framework is better-suited to study the determinants of information efficiency than perfectly competitive models that preclude such a "magic of markets" result.

We go on to characterize additional differences between our framework and the literature on noisy rational expectations equilibria, focusing on the co-movements between conventional measures like price informativeness, liquidity, and volatility. The differences we uncover highlight a novel sense in which the assumption of perfect competition is not innocuous.

Finally, we illustrate how to identify price informativeness from the data, using one of two approaches: (i) proxying it by price volatility, which is valid assuming fundamental risk is constant; or (ii) using volatility and volume, two easily observable variables, paired with an assumption of information-rich markets, to back out the dynamics of fundamental risk and non-fundamental noise, hence the dynamics of informativeness.

References

Atakan, A. E., and Ekmekci, M. (2023). *Information aggregation in auctions with costly information* (tech. rep.). Working paper.

Bai, J., Philippon, T., and Savov, A. (2016). Have financial markets become more informative? *Journal of Financial Economics*, *122*(3), 625–654.

Buss, A., and Sundaresan, S. (2023). More risk, more information: How passive ownership can improve informational efficiency. *The Review of Financial Studies*, *36*(12), 4713–4758.

Campbell, J. Y., Lettau, M., Malkiel, B., Xu, Y., et al. (2023). Idiosyncratic equity risk two decades later. *Critical Finance Review*, *12*(1-4), 203–223.

Dávila, E., and Parlatore, C. (2021). Trading costs and informational efficiency. *The Journal of Finance*, *76*(3), 1471–1539.

Dávila, E., and Parlatore, C. (2023). Volatility and informativeness. *Journal of Financial Economics*, *147*(3), 550–572.

Dávila, E., and Parlatore, C. (2025). Identifying price informativeness. *Review of Financial Studies (Forthcoming)*.

Dubey, P., Geanakoplos, J., and Shubik, M. (1987). The revelation of information in strategic market games: A critique of rational expectations equilibrium. *Journal of Mathematical Economics*, *16*(2), 105–137.

García, D., and Urošević, B. (2013). Noise and aggregation of information in large markets. *Journal of Financial Markets*, *16*(3), 526–549.

Glebkin, S., Malamud, S., and Teguia, A. (2020). *Asset prices and liquidity with market power and non-gaussian payoffs* (tech. rep.).

Glebkin, S., Malamud, S., and Teguia, A. (2023). Strategic trading with wealth effects.

Grossman, S. (1976). On the efficiency of competitive stock markets where trades have diverse information. *The Journal of Finance*, *31*(2), 573–585.

Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393–408.

Han, J., and Kyle, A. S. (2018). Speculative equilibrium with differences in higher-order beliefs. *Management Science*, *64*(9), 4317–4332.

Hellwig, C., and Veldkamp, L. (2009). Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies*, *76*(1), 223–251.

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

Hellwig, M. F. (1982). Rational expectations equilibrium with conditioning on past prices: A mean-variance example. *Journal of Economic Theory*, *26*(2), 279–312.

Jackson, M. O. (1991). Equilibrium, price formation, and the value of private information. *The Review of Financial Studies*, *4*(1), 1–16.

Kacperczyk, M., Sundaresan, S., and Wang, T. (2021). Do foreign institutional investors improve price efficiency? *The Review of Financial Studies*, *34*(3), 1317–1367.

Kacperczyk, M. T., Nosal, J., and Sundaresan, S. (2023). Market power and price informativeness. *Available at SSRN 3137803*.

Kothari, S. P., and Sloan, R. G. (1992). Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics*, *15*(2-3), 143–171.

Kovalenkov, A., and Vives, X. (2014). Competitive rational expectations equilibria without apology. *Journal of Economic Theory*, *149*, 211–235.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315–1335.

Kyle, A. S. (1989). Informed speculation with imperfect competition. *The Review of Economic Studies*, *56*(3), 317–355.

Kyle, A. S., Obizhaeva, A. A., and Wang, Y. (2018). Smooth trading with overconfidence and market power. *The Review of Economic Studies*, *85*(1), 611–662.

Lambert, N. S., Ostrovsky, M., and Panov, M. (2018). Strategic trading in informationally complex environments. *Econometrica*, *86*(4), 1119–1157.

Lee, J., and Kyle, A. S. (2018). When are financial markets perfectly competitive?

Martinelli, C. (2006). Would rational voters acquire costly information? *Journal of Economic Theory*, *129*(1), 225–251.

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

Pavan, A., Sundaresan, S., and Vives, X. (2022). (in) efficiency in information acquisition and aggregation through prices.

Roll, R. (1988). R^2. *The Journal of Finance*, *43*(3), 541–566.

Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica: Journal of the Econometric Society*, 1415–1430.

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

Vives, X. (2014). On the possibility of informationally efficient markets. *Journal of the European Economic Association*, *12*(5), 1200–1239.

Vives, X. (2017). Endogenous public information and welfare in market games. *The Review of Economic Studies*, *84*(2), 935–963.

A Proofs and derivations

A.1 General expressions for equilibrium objects

We collect here some formulas for important equilibrium objects. We have the following general formulas from Lemma 1:

$$\begin{aligned}
\beta &= \sqrt{\frac{A^\top \Sigma_{\text{diag}} A}{\sigma_u}} \\
\alpha &= \frac{1}{\beta} A \\
D &= \alpha \cdot \theta + u \\
p &= \beta D
\end{aligned}$$

where θ is the vector of all agents' signals. And in the process of proving Lemma 2, we proved that

$$\begin{aligned}
A &= \frac{\boldsymbol{z}}{1 + \boldsymbol{z} \cdot \mathbf{1}} \\
A^\top \Sigma_{\text{diag}} A &= \frac{\sigma_v}{2} \frac{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}{(1 + \boldsymbol{z} \cdot \mathbf{1})^2}
\end{aligned}$$

For reference, note that plugging these expressions for A into the expressions for α and β delivers

$$\begin{aligned}
\beta &= \sqrt{\frac{\sigma_v}{2\sigma_u}} \frac{\sqrt{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}}{1 + \boldsymbol{z} \cdot \mathbf{1}} \\
\alpha &= \sqrt{\frac{2\sigma_u}{\sigma_v}} \frac{\boldsymbol{z}}{\sqrt{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}}
\end{aligned}$$

A.2 Indirect utility function (Proof of Lemma 2)

We now use Lemma 1 to calculate β and α explicitly. Applying the Sherman-Morrison formula to the definition of Λ, we obtain:

$$\Lambda^{-1} = \Delta^{-1} - \frac{\sigma_v}{1 + \sigma_v \sum_{k=1}^{n} \frac{1}{\delta_k}} \Delta^{-1} \mathbf{1} \mathbf{1}^\top \Delta^{-1}$$

where:

$$\Delta = \text{diag}(\delta_1, \delta_2, \ldots, \delta_n) \quad \text{and} \quad \delta_i = 2\sigma_i - \sigma_v$$

We can then compute A:

$$A = \Lambda^{-1}\Sigma_{\theta v} = \sigma_v \Delta^{-1}\mathbf{1} - \frac{\sigma_v^2}{1 + \sigma_v \sum_{k=1}^{n} \frac{1}{\delta_k}} \Delta^{-1}\mathbf{1}\mathbf{1}^\top \Delta^{-1}\mathbf{1}$$

This can be simplified as follows. The first term is:

$$\sigma_v \Delta^{-1}\mathbf{1} = \sigma_v \begin{pmatrix} \frac{1}{\delta_1} \\ \frac{1}{\delta_2} \\ \vdots \\ \frac{1}{\delta_n} \end{pmatrix}$$

For the second term, recall that $\mathbf{1}^\top \Delta^{-1}\mathbf{1} = \mathrm{tr}(\Delta^{-1}) = \sum_{i=1}^{n} \frac{1}{\delta_i}$. Thus, the second term is:

$$\frac{\sigma_v^2 \sum_{i=1}^{n} \frac{1}{\delta_i}}{1 + \sigma_v \sum_{i=1}^{n} \frac{1}{\delta_i}} \begin{pmatrix} \frac{1}{\delta_1} \\ \frac{1}{\delta_2} \\ \vdots \\ \frac{1}{\delta_n} \end{pmatrix}$$

Putting these pieces together, we have

$$A = \frac{\sigma_v}{1 + \sigma_v \sum_{i=1}^{n} \frac{1}{\delta_i}} \begin{pmatrix} \frac{1}{\delta_1} \\ \frac{1}{\delta_2} \\ \vdots \\ \frac{1}{\delta_n} \end{pmatrix}$$

We can now easily calculate $A^\top \Sigma_{\text{diag}} A$:

$$A^\top \Sigma_{\text{diag}} A = \frac{\sigma_v^2}{\left(1 + \sigma_v \sum_{i=1}^{n} \frac{1}{\delta_i}\right)^2} \sum_{i=1}^{n} \frac{\sigma_i}{\delta_i^2}$$

To obtain the formulas for A and $A^\top \Sigma_{\text{diag}} A$ in Appendix A.1, we need to write them in terms of z_i and $\boldsymbol{z} := (z_i)_{i=1}^{n}$. To do this, simply note from equation (3) that

$$\begin{aligned} \frac{\sigma_v}{\delta_i} &= \frac{\sigma_v}{2\sigma_i - \sigma_v} = z_i \\ \frac{\sigma_i \sigma_v^2}{\delta_i^2} &= \sigma_i z_i^2 = \frac{1}{2}\sigma_v \left(z_i + z_i^2\right) \end{aligned}$$

Making these replacements, we obtain the formulas in Appendix A.1.

We can finally recover the ex-ante expected profits V_i. To do this, we will use the law of iterated expectations, implying that $V_i = \mathbb{E}[\tilde{V}_i] = \mathbb{E}\big[\mathbb{E}[d_i(v-p) \mid \theta_i]\big] = \mathbb{E}\big[d_i(v-p)\big]$. Using the expressions for d_i, p, and D in Appendix A.1, we may write the ex-post realized profit

$$d_i(v-p) = \alpha_i \theta_i \Big[v - \beta \Big(\sum_{k=1}^{n} \alpha_k \theta_k + u \Big) \Big]$$

Take the prior expectation of this realized profit:

$$\begin{aligned} V_i &= \alpha_i \left\{ \mathbb{E}[v\theta_i] - \beta \sum_{k=1}^{n} \alpha_k \mathbb{E}[\theta_i \theta_k] - \beta \mathbb{E}[u\theta_i] \right\} \\ &= \alpha_i \left\{ \Big(1 - \beta \sum_{k \neq i} \alpha_k \Big) \sigma_v - \beta \alpha_i \sigma_i \right\} \end{aligned}$$

where we have used the statistical properties of the signals. Now, plug in α and β from the expressions in Appendix A.1. Substituting those objects and doing extensive algebra yields:

$$V_i = \sqrt{\frac{\sigma_v \sigma_u}{2}} \frac{z_i + z_i^2}{(1 + \boldsymbol{z} \cdot \mathbf{1}) \sqrt{\boldsymbol{z} \cdot \mathbf{1} + \|\boldsymbol{z}\|^2}}$$

A.3 Grossman-Stiglitz paradox (Proof of Proposition 1)

For clarity, let us write $z_n^*(\sigma_u)$ and $f_n(z;\sigma_u)$ to emphasize the dependence on noise. Note that $\lim_{\sigma_u \to 0} f_n(z;\sigma_u) = 0$ for any fixed $z > 0$. Note also that $f_n(z;\sigma_u)$ is uniformly continuous in z on $[\epsilon, 1]$ for any $\epsilon > 0$.

Now suppose, leading to contradiction, that $z_n^*(0+) := \lim_{\sigma_u \to 0} z_n^*(\sigma_u) > 0$. Because of the above facts on f_n, and using the conjecture $z_n^*(0+) > 0$ we have $\lim_{\sigma_u \to 0} f_n(z_n^*(\sigma_u);\sigma_u) = f_n(z_n^*(0+);0) = 0$. But this can only be consistent with the equilibrium condition (5) if $z_n^*(0+) = 0$, since $c'(z) > 0$ for all $z > 0$. This contradiction shows that $nz_n^*(0+) = z_n^*(0+) = 0$.

A.4 Vanishing individual information (Proof of Proposition 2)

First, consider the case $\chi = \infty$. Given strict convexity of c and the fact that $f_n(z) < \infty$ for all $z > 0$, the only possible solution to equation (5) is $z_n^* = 0$.

Now, suppose $\chi < \infty$, so that $z_n^* > 0$ for each n. Consider a convergent sub-sequence $(n_t)_{t\geq 0}$ such that $\lim_{t\to\infty} z_{n_t}^* = z_\infty^* \neq 0$. If so, we have $\lim_{t\to\infty} n_t z_{n_t}^* = +\infty$. Thus, we have $\lim_{t\to\infty} f_{n_t}(z_{n_t}^*) = 0$. By equation (5), we must also have $\lim_{t\to\infty} c'(z_{n_t}^*) = 0$. Since c is strictly convex, we have by the monotone convergence theorem $0 = \lim_{t\to\infty} c'(z_{n_t}^*) = c'(z_\infty^*)$. But strict convexity also implies $c'(z) > 0$ for all $z > 0$, contradicting $z_\infty^* > 0$. This argument works for any sub-sequence, and so $\lim_{n\to\infty} z_n^*$ exists and is equal to zero.

A.5 Aggregate information characterization (Proof of Theorem 1)

Claim 1 is already proved in Proposition 2, which showed that $z_n^* = 0$ for each n when $\chi = \infty$.

Next, we prove claim 3 by contradiction. By Proposition 2, we know that z_n^* is a positive sequence that converges to zero. Suppose $nz_n^* \to Z < \infty$. Then, we have $\lim_{n\to\infty} f_n(z_n^*) > 0$. On the other hand, the fact that $z_n^* \to 0$ implies by the monotone convergence theorem that $c'(z_n^*) \to \chi = 0$. But, along the sequence, we have that (5) holds with equality, which contradicts the fact that $f_n(z_n^*)$ and $c'(z_n^*)$ have different limit points.

Finally, we prove claim 2 using a similar method. By Proposition 2, we have $c'(z_n^*) \to \chi > 0$, which implies $f_n(z_n^*) \to \chi$ by equation (5). Now, let us show that $Z := \lim_{n\to\infty} nz_n^* = Z^*$. Suppose, leading to contradiction, that $Z \neq Z^*$. Taking the limit $n \to \infty$ of $f_n(z_n^*)$, using $z_n^* \to 0$ and $nz_n^* \to Z$, we obtain

$$\begin{aligned}
&\lim_{n\to\infty} f_n(z_n^*) \\
&= \lim_{n\to\infty} \frac{\sqrt{\frac{\sigma_v\sigma_u}{2}}(1+2z_n^*)}{(1+nz_n^*)\sqrt{nz_n^*(1+z_n^*)}}\left[1 - \frac{z_n^*(1+z_n^*)}{1+2z_n^*}\left(\frac{1}{1+nz_n^*} + \frac{\frac{1}{2}(1+2z_n^*)}{nz_n^*(1+z_n^*)}\right)\right] \\
&= \sqrt{\frac{\sigma_v\sigma_u}{2}}\frac{1}{(1+Z)\sqrt{Z}}
\end{aligned}$$

If $Z \neq Z^*$, then this result differs from χ, a contradiction.

A.6 Rate of convergence if $\chi = 0$

Here, we characterize the rate at which information-collection vanishes at the individual level. To formalize this question, note that there exists some $\tilde{\zeta} \geq 0$ such that $z_n^* n^{\tilde{\zeta}} \to 0$ (by Proposition 2). Let ζ be the largest such parameter, i.e., $\zeta := \sup\{\tilde{\zeta} \geq 0 : \lim_{n\to\infty} z_n^* n^{\tilde{\zeta}} = 0\}$; this ζ is the relevant convergence rate. By Theorem 1, we have that $\zeta = 1$ if $\chi \in (0, \infty)$, whereas $\zeta \in (0,1]$ if $\chi = 0$. And based on the results in the paper, this is all we can say about the rate of convergence in the $\chi = 0$ case.

To characterize ζ when $\chi = 0$, we will specialize to a class of cost functions that share the following property: $\lim_{z\to 0} \frac{c'(z)}{z^\gamma} \to \kappa$, for some $\gamma > 0$ and $\kappa > 0$. This class of functions automatically satisfies $c'(0) = 0$ and includes, for instance, all power function costs $c(z) \propto z^{1+\gamma}$, for $\gamma > 0$, or any function such that the dominant term as $z \to 0$ is such a power. In that case, we find $\zeta = \frac{3}{3+2\gamma}$, so that the rate of convergence depends on the curvature of information costs at the prior, namely $\gamma = \lim_{z\to 0} \frac{zc''(z)}{c'(z)}$.

Proposition A.1. *Let $c(z)$ be such that $\lim_{z\to 0} \frac{c'(z)}{z^\gamma} = \kappa$ for some $\gamma > 0$ and $\kappa > 0$. Then, $z_n^* \to 0$ at the rate $n^{-3/(3+2\gamma)}$, in the following sense:*

$$n^{3/(3+2\gamma)} z_n^* \to \kappa^{-2/(3+2\gamma)} \left(\frac{\sigma_v \sigma_u}{2}\right)^{1/(3+2\gamma)}.$$

Proof. By Proposition 2, we have $c'(z_n^*)/(z_n^*)^\gamma \to \kappa$. By equation (5), this implies $f_n(z_n^*)/(z_n^*)^\gamma \to \kappa$ must hold.

Let $Z \in (0,\infty)$, and let $\zeta \in (0,1)$ be an arbitrary constant. Substitute $z = n^{-\zeta} Z$ into $f_n(z)/z^\gamma$ to get

$$\frac{f_n(n^{-\zeta}Z)}{n^{-\gamma\zeta}Z^\gamma}$$
$$= \frac{\sqrt{\frac{\sigma_v\sigma_u}{2}}(n^{\gamma\zeta}Z^{-\gamma} + 2n^{-(1-\gamma)\zeta}Z^{1-\gamma})}{(1+n^{1-\zeta}Z)\sqrt{n^{1-\zeta}Z(1+n^{-\zeta}Z)}}\left[1 - \frac{n^{-\zeta}Z(1+n^{-\zeta}Z)}{1+2n^{-\zeta}Z}\left(\frac{1}{1+n^{1-\zeta}Z} + \frac{\frac{1}{2}(1+2n^{-\zeta}Z)}{n^{1-\zeta}Z(1+n^{-\zeta}Z)}\right)\right].$$

By inspection, the term in square brackets converges to 1. The leading term converges to κ if and only if $\zeta = \frac{3}{3+2\gamma}$ and $Z = [\kappa^{-2}(\sigma_v\sigma_u/2)]^{1/(3+2\gamma)}$. [Algebra: we need

$$\frac{\sigma_v\sigma_u}{2}(n^{\gamma\zeta}Z^{-\gamma} + 2n^{-(1-\gamma)\zeta}Z^{1-\gamma})^2 \sim \kappa^2(1+n^{1-\zeta}Z)^2 n^{1-\zeta}Z(1+n^{-\zeta}Z)$$
$$\frac{\sigma_v\sigma_u}{2} n^{2\gamma\zeta}Z^{-2\gamma} \sim \kappa^2 n^{3-3\zeta}Z^3$$
$$\frac{\sigma_v\sigma_u}{2} \sim \kappa^2 n^{3-3\zeta-2\gamma\zeta}Z^{3+2\gamma}$$

which delivers the result. □

Proposition A.1 tells us how fast aggregate information nz_n^* explodes as the market grows. The answer: $nz_n^* \sim (\frac{\sigma_v\sigma_u}{2\kappa^2})^{\frac{1}{3+2\gamma}} n^{\frac{2\gamma}{3+2\gamma}}$ for large n. If the information technology is highly curved (high γ), we should expect to see market information grow quickly with market size.

An additional takeaway from Proposition A.1 is that full information can obtain in a large economy even if noise vanishes as the economy grows. In particular, a model with $\chi = 0$ has full information emerging as $n \to \infty$ even if noise $\sqrt{\sigma_u}$ vanishes at any rate slower than $n^{-3/2}$. This sharply contrasts with the case $\chi \in (0, \infty)$, as discussed following Theorem 1. If $\chi \in (0, \infty)$, vanishing noise implies vanishing information.

A.7 Symmetric equilibrium (Proof of Lemma 3)

In a symmetric equilibrium, $\boldsymbol{z} = z_n^* \mathbf{1}$, and so $\boldsymbol{z} \cdot \mathbf{1} = n z_n^*$ and $\|\boldsymbol{z}\|^2 = n(z_n^*)^2$. Substituting these results into the expressions in Appendix A.1, we obtain

$$\beta_n = \sqrt{\frac{\sigma_v}{2\sigma_u}} \sqrt{\frac{n z_n^*(1 + z_n^*)}{(1 + n z_n^*)^2}}$$

$$\alpha_n = \frac{1}{n} \sqrt{\frac{2\sigma_u}{\sigma_v}} \sqrt{\frac{n z_n^*}{1 + z_n^*}} \mathbf{1}$$

Then, noting $e_u := u / \sqrt{\sigma_u} \sim \text{Normal}(0, 1)$, we have

$$D_n = \sqrt{\frac{2\sigma_u}{\sigma_v}} \sqrt{\frac{n z_n^*}{1 + z_n^*}} \frac{\theta \cdot \mathbf{1}}{n} + \sqrt{\sigma_u} e_u$$

$$p_n = \frac{n z_n^*}{1 + n z_n^*} \frac{\theta \cdot \mathbf{1}}{n} + \sqrt{\frac{\sigma_v}{2}} \frac{\sqrt{n z_n^*(1 + z_n^*)}}{1 + n z_n^*} e_u$$

We will now take the limit $n \to \infty$. First, note the fact that $\theta = v\mathbf{1} + \varepsilon$, where

$$\varepsilon \sim \text{Normal}\left(0, \frac{1}{2} \sigma_v \frac{1 - z_n^*}{z_n^*} I_n\right)$$

in a symmetric equilibrium. Therefore,

$$\frac{\varepsilon \cdot \mathbf{1}}{n} \sim \text{Normal}\left(0, \frac{1}{2} \sigma_v \frac{1 - z_n^*}{n z_n^*}\right)$$

and so using the definition $Z^* := \lim_{n \to \infty} n z_n^*$ and the result that $z_n^* \to 0$ (Proposition 2), we have by the Central Limit Theorem,

$$\frac{\theta \cdot \mathbf{1}}{n} \to v + \sqrt{\frac{\sigma_v}{2Z^*}} e_\theta$$

in distribution, where $e_\theta \sim \text{Normal}(0, 1)$. Consequently, regardless of whether Z^* is finite or infinite, we have the following limits in distribution,

$$
\begin{aligned}
D_n &\to \sqrt{2\sigma_u}\Big(\sqrt{Z^*}\frac{v}{\sqrt{\sigma_v}} + \frac{e_\theta + e_u}{\sqrt{2}}\Big) \\
p_n &\to \sqrt{\sigma_v}\frac{\sqrt{Z^*}}{1+Z^*}\Big(\sqrt{Z^*}\frac{v}{\sqrt{\sigma_v}} + \frac{e_\theta + e_u}{\sqrt{2}}\Big)
\end{aligned}
$$

This proves Lemma 3.

A.8 Identification (Proof of Proposition 4)

Start with the observable metrics:

$$
\begin{aligned}
\mathcal{V} &= \sigma_v \frac{Z}{1+Z} \\
\mathcal{L} &= \frac{\sigma_u}{\chi Z} \\
Z(1+Z)^2 &= \frac{1}{2}\frac{\sigma_u \sigma_v}{\chi^2}
\end{aligned}
$$

Differentiate these objects, assuming $d\chi = 0$, and using the chain rule to replace $d\log(1+Z) = \frac{Z}{1+Z} d\log Z$, to get

$$
\begin{aligned}
d\log \mathcal{V} &= d\log \sigma_v + \frac{1}{1+Z} d\log Z \\
d\log \mathcal{L} &= d\log \sigma_u - d\log Z \\
\frac{1+3Z}{1+Z} d\log Z &= d\log \sigma_u + d\log \sigma_v
\end{aligned}
$$

Substitute the third equation into the first two to get

$$
\begin{aligned}
d\log \mathcal{V} &= d\log \sigma_v + \frac{1}{1+3Z}\Big(d\log \sigma_u + d\log \sigma_v\Big) \\
d\log \mathcal{L} &= d\log \sigma_u - \frac{1+Z}{1+3Z}\Big(d\log \sigma_u + d\log \sigma_v\Big)
\end{aligned}
$$

Stacking this gives

$$
\begin{pmatrix} d\log \mathcal{V} \\ d\log \mathcal{L} \end{pmatrix} = \frac{1}{1+3Z}\begin{pmatrix} 1 & 2+3Z \\ 2Z & -(1+Z) \end{pmatrix}\begin{pmatrix} d\log \sigma_u \\ d\log \sigma_v \end{pmatrix}
$$

Invert this system to obtain

$$d\log\sigma_u = \frac{1+Z}{1+2Z}d\log\mathcal{V} + \frac{2+3Z}{1+2Z}d\log\mathcal{L} \tag{A.1}$$

$$d\log\sigma_v = \frac{2Z}{1+2Z}d\log\mathcal{V} - \frac{1}{1+2Z}d\log\mathcal{L} \tag{A.2}$$

Augment this with the evolution of Z, so that we have a closed dynamical system:

$$\begin{aligned} d\log Z &= \frac{1+Z}{1+3Z}\Big(d\log\sigma_u + d\log\sigma_v\Big) \\ &= \frac{1+Z}{1+2Z}\Big(d\log\mathcal{V} + d\log\mathcal{L}\Big) \end{aligned} \tag{A.3}$$

We also differentiate the price informativeness measure $\mathcal{I} = \frac{1+Z}{\sigma_v}$ to get

$$\begin{aligned} d\log\mathcal{I} &= \frac{Z}{1+Z}d\log Z - d\log\sigma_v \\ &= Zd\log\mathcal{V} - (1+Z)d\log\sigma_v \\ &= -\frac{Z}{1+2Z}d\log\mathcal{V} + \frac{1+Z}{1+2Z}d\log\mathcal{L} \end{aligned} \tag{A.4}$$

Thus, if we are given initial values $(Z_0, \sigma_{u,0}, \sigma_{v,0})$, we can simulate the entire system forward. We can simplify this problem by noting that we observe $\mathcal{V}_0$ and $\mathcal{L}_0$, meaning we can essentially observe two of these initial conditions, leaving only one unobservable. In particular, using the expressions $\mathcal{V} = \sigma_v\frac{Z}{1+Z}$ and $\mathcal{L}^2 = \frac{2\sigma_u}{\sigma_v}\frac{(1+Z)^2}{Z}$, and conjecturing a value for Z_0, we may solve for

$$\sigma_{u,0} = \frac{1}{2}\mathcal{L}_0^2\mathcal{V}_0\frac{1}{1+Z_0} \tag{A.5}$$

$$\sigma_{v,0} = \mathcal{V}_0\frac{1+Z_0}{Z_0} \tag{A.6}$$

Thus, the entire problem boils down to forming a statistical prior for Z_0.

B Alternative: fixed cost of information

Now, let's assume the information cost satisfies $\bar{\chi} := c(z) > 0$ for all z. There is a fixed cost of acquiring any information. The fact that there is no additional variable cost simplifies the analysis so that either $z_i = 0$ or $z_i = 1$ for each trader i (none or perfect information).

One concern with fixed costs, suggested by the analysis of Grossman and Stiglitz (1980), is that an equilibrium may fail to exist. We will show that this is not the case here. In fact, there is a well-defined equilibrium for any $\bar{\chi}$. First, we will consider symmetric equilibria, followed by asymmetric equilibria. The key statistic, which is exactly analogous to (6), is

$$\Gamma := \frac{\sqrt{\sigma_v \sigma_u}}{\bar{\chi}}. \tag{B.1}$$

Whether Γ is greater or less than 2 determines the type of equilibrium. And the level of Γ determines the level of aggregate information in a large market.

In a symmetric equilibrium, the pre-information-cost utility value of each trader, assuming all other traders collect information z_n^*, is the same as before:

$$V_n(z; z_n^*) = \sqrt{\frac{\sigma_v \sigma_u}{2}} \frac{z(1+z)}{(1+(n-1)z_n^* + z)\sqrt{(n-1)z_n^*(1+z_n^*) + z(1+z)}}.$$

Notice that $V_n(0;0) = V_n(0;1) = 0$. The key criterion for equilibrium is optimality: in an equilibrium with information collection, trader utility must satisfy

$$V_n(1;1) - V_n(0;1) \geq \bar{\chi}, \quad \text{if} \quad z_n^* = 1. \tag{B.2}$$

On the other hand, for an equilibrium with ignorance, trader utility must satisfy

$$V_n(1;0) - V_n(0;0) \leq \bar{\chi}, \quad \text{if} \quad z_n^* = 0. \tag{B.3}$$

An asymmetric equilibrium is more complicated. To set it up, let π_n^* denote the equilibrium probability a trader acquires any information. Because each trader only acquires either $z_i = 0$ or $z_i = 1$, aggregate information in the market will be equal to the total number of informed traders, which is the binomial random variable $\text{Binom}(n, \pi_n^*)$. The pre-cost value of a single trader, assuming the fraction of the other $n-1$ traders acquiring information is π_n, is

$$V_n(z; \pi_n) = \sqrt{\frac{\sigma_v \bar{\sigma}_u}{2}} \mathbb{E}\Big[\frac{z(1+z)}{(1+(n-1)\pi_n + z)\sqrt{2(n-1)\pi_n + z(1+z)}}\Big],$$

where the expectation is over the possible realizations of π_n. By rational expectations, this random variable has probability distribution $(n-1)\pi_n \sim \text{Binom}(n-1, \pi_n^*)$. In equilibria of the large-n economy, the amount of randomness in π_n vanishes, i.e., $\pi_n - \pi_n^* \to 0$

almost-surely, by the law of large numbers. In a mixed-strategy equilibrium with $\pi_n^* > 0$ being the information probability, it must be the case that

$$V_n(1;\pi_n) - \bar{\chi} = V_n(0;\pi_n) = 0, \quad \text{if} \quad \pi_n \in (0,1),$$

i.e., the trader is indifferent between acquiring information or not.

Based on this characterization, we can prove the following theorem, which is roughly speaking the fixed cost version of Theorem 1.

Theorem B.1. *An equilibrium with fixed information costs always exists in the large-n limit and satisfies the following:*

1. *If $\Gamma \leq 2$, then the equilibrium is symmetric and features $nz_n^* = 0$ for all n large enough.*
2. *If $\Gamma > 2$, then the equilibrium is asymmetric and features $n\pi_n^* \to \Pi^*$, where Π^* is the unique positive solution to $(\Pi+2)^2(\Pi+1) = \Gamma^2$.*

Proof. 1. Suppose $\Gamma \leq 2$. First, we prove that, assuming a symmetric equilibrium exists, $nz_n^* = 0$ for all n large enough. This is a direct consequence of $\lim_{n\to\infty} V_n(1;1) = 0$, implying $V_n(1) < \bar{\chi}$ for all n large enough, so (B.2) cannot hold. Second, we prove existence of a symmetric equilibrium with $z_n^* = 0$ for large n. The utility from deviating from $z_n^* = 0$ is $V(1;0) = \frac{1}{2}\sqrt{\sigma_v\sigma_u}$. Thus, $V_n(1;0) - V_n(0;0) - \bar{\chi} = \frac{1}{2}\sqrt{\sigma_v\sigma_u} - \bar{\chi} \leq 0$, so (B.3) holds, and the equilibrium is confirmed.

2. Suppose $\Gamma > 2$. As a consequence of the argument for claim 1, a symmetric equilibrium cannot occur if $\Gamma > 2$, because in such case neither (B.2) nor (B.3) can hold. Moving to asymmetric equilibria, it is easy to see that $n\pi_n^* \not\to \infty$ as $n \to \infty$. Indeed, if $n\pi_n^* \to \infty$ occurred, then $V_n(1;\pi_n^*) \to 0$, and so $V_n(1;\pi_n) \to 0$ a.s. (recall $\pi_n - \pi_n^* \to 0$ by the law of large numbers). In fact, writing $\Pi^* := \lim_{n\to\infty} n\pi_n^*$ and then evaluating $V_n(1;\pi_n) = \bar{\chi}$ in the large-n limit, we find that Π^* must solve the cubic equation

$$(\Pi+2)^2(\Pi+1) = \frac{\sigma_v\sigma_u}{\bar{\chi}^2}.$$

This has a unique positive solution $\Pi^* > 0$ if and only if $\Gamma > 2$. Thus, an asymmetric equilibrium exists if $\Gamma > 2$, but not if $\Gamma \leq 2$, and the aggregate information in this equilibrium is Π^* as $n \to \infty$. $\square$

B.1 Asymmetric equilibrium

In this section, we prove a result analogous to Lemma 3 but for the fixed information cost case covered above in Theorem B.1 (with $\Gamma > 2$). In such case, an "asymmetric equilibrium" arises where each trader chooses to be fully-informed with probability π_n^*.

Lemma B.1. *Let π_n^* denote the probability of information acquisition in an asymmetric equilibrium with n strategic traders, and let $\Pi_n := Binom(n, \pi_n^*)$ be the aggregate number of informed traders. Then,*

$$\beta_n = \sqrt{\frac{\sigma_v}{\sigma_u}\frac{\Pi_n}{(1+\Pi_n)^2}}$$
$$D_n = \sqrt{\frac{\sigma_u}{\sigma_v}}\Pi_n v + \sqrt{\sigma_u} e_u$$
$$p_n = \sqrt{\sigma_v}\Big(\frac{\Pi_n}{1+\Pi_n}\frac{v}{\sqrt{\sigma_v}} + \frac{\sqrt{\Pi_n}}{1+\Pi_n} e_u\Big)$$

where $e_u = u/\sqrt{\sigma_u} \sim Normal(0,1)$. Let $\Pi^* := \lim_{n\to\infty} n\pi_n^* \in [0,\infty]$ be the large-n limit of aggregate information. Then, we have the following limiting equilibrium objects, almost-surely,

$$\lim_{n\to\infty} \beta_n = \sqrt{\frac{\sigma_v}{\sigma_u}}\frac{\sqrt{\Pi^*}}{1+\Pi^*}$$
$$\lim_{n\to\infty} D_n = \sqrt{\sigma_u}\Big[\sqrt{\Pi^*}\frac{v}{\sqrt{\sigma_v}} + e_u\Big]$$
$$\lim_{n\to\infty} p_n = \sqrt{\sigma_v}\frac{\sqrt{\Pi^*}}{1+\Pi^*}\Big[\sqrt{\Pi^*}\frac{v}{\sqrt{\sigma_v}} + e_u\Big]$$

Proposition B.1. *In the large-n limit of asymmetric equilibria, the measures of liquidity, price informativeness, excess price volatility, and trading volume are given by the following:*

$$\text{(liquidity)} \quad \mathcal{L} = \sqrt{\frac{\sigma_u}{\sigma_v}}\frac{1+\Pi^*}{\sqrt{\Pi^*}}$$
$$\text{(informativeness)} \quad \mathcal{I} = \frac{1+\Pi^*}{\sigma_v}$$
$$\text{(volatility)} \quad \mathcal{V} = \sigma_v\frac{\Pi^*}{(1+\Pi^*)^2}$$
$$\text{(volume)} \quad \mathcal{D} = \sigma_u(1+\Pi^*)$$

Proof. Suppose in the size-n economy, each trader obtains a perfect signal with probability π_n^*, and otherwise remains uninformed. In this case, $\boldsymbol{z}$ is a vector of zeros and ones, with the probability of each non-zero entry being an independent Bernoulli draw. Then, $\boldsymbol{z} \cdot \mathbf{1} = \|z\|^2 := \Pi_n \sim \text{Binom}(n, \pi_n^*)$. Substituting this into formulas in Appendix A.1, we have

$$\beta_n = \sqrt{\frac{\sigma_v}{\sigma_u}\frac{\Pi_n}{(1+\Pi_n)^2}} \quad \text{and} \quad \alpha_n \cdot \theta = \sqrt{\frac{\sigma_u}{\sigma_v}}\Pi_n v,$$

and so

$$\begin{aligned} D_n &= \sqrt{\frac{\sigma_u}{\sigma_v}}\Pi_n v + \sqrt{\sigma_u} e_u \\ p_n &= \sqrt{\sigma_v}\Big(\frac{\Pi_n}{1+\Pi_n}\frac{v}{\sqrt{\sigma_v}} + \frac{\sqrt{\Pi_n}}{1+\Pi_n} e_u\Big) \end{aligned}$$

To take the limit $n \to \infty$, note that $\Pi_n/(n\pi_n^*)$ converges to 1 almost-surely, by the Strong Law of Large Numbers. Hence, denoting $\Pi^* := \lim_{n\to\infty} n\pi_n^*$, we have $\Pi_n \to \Pi^*$ almost-surely. Using this fact, we have

$$\begin{aligned} D_n &\to \sqrt{\sigma_u}\Big(\sqrt{\Pi^*}\frac{v}{\sqrt{\sigma_v}} + e_u\Big) \\ p_n &\to \sqrt{\sigma_v}\frac{\sqrt{\Pi^*}}{1+\Pi^*}\Big(\sqrt{\Pi^*}\frac{v}{\sqrt{\sigma_v}} + e_u\Big) \end{aligned}$$

This proves Lemma B.1.

Proposition B.1 is proved by combining the results of Lemma B.1 with the definitions of the measures, contained in Definitions 1-4. □

Comparing lem:symmetric-eqm and lem:asymmetric-eqm, we see that the limiting objects are very similar across the symmetric and asymmetric equilibria (which recall arise with variable and fixed costs, respectively). In particular, modulo some $\sqrt{2}$ constants, the formulas are identical in distribution. For instance, compare the limiting price in the two equilibrium types:

$$\begin{aligned} \text{(symmetric equilibrium)} \quad \lim_{n\to\infty} p_n &= \sqrt{\sigma_v}\frac{\sqrt{Z^*}}{1+Z^*}\Big[\sqrt{Z^*}\frac{v}{\sqrt{\sigma_v}} + \frac{e_\theta + e_u}{\sqrt{2}}\Big] \\ \text{(asymmetric equilibrium)} \quad \lim_{n\to\infty} p_n &= \sqrt{\sigma_v}\frac{\sqrt{\Pi^*}}{1+\Pi^*}\Big[\sqrt{\Pi^*}\frac{v}{\sqrt{\sigma_v}} + e_u\Big] \end{aligned}$$

Since Z^* and Π^* have similar interpretations, as aggregate information (in units of precision), the formulas are almost identical. The only difference is that the noise $\frac{e_\theta + e_u}{\sqrt{2}}$ in the symmetric equilibrium price is replaced by e_u in the asymmetric equilibrium price. But since both are standard normal random variables, the prices are identical in distribution.

For this reason, the measures in Proposition 3 and Proposition B.1, for symmetric and asymmetric equilibria, are identical.

Theorem B.1 proves that equilibrium exists generically. In competitive models like Grossman and Stiglitz (1980), equilibrium non-existence is a problem: neither a fully-informed nor a fully-uninformed equilibrium can exist. But here, we can obtain either of these extreme cases depending on the size of the statistic Γ. If $\Gamma \leq 2$, we have a fully-uninformed equilibrium. But if $\Gamma \to \infty$, we have a fully-informed equilibrium, in the sense that aggregate information becomes maximal. The characterization of the result in terms of $\Pi^* = \lim_{n\to\infty} n\pi_n^*$ is intuitive: Π^* represents the *number of informed traders* in the large-n limit, which necessarily remains finite for any fixed cost $\bar{\chi}$.

C Comparison to Competitive NREE

We compare our results to a competitive Noisy Rational Expectations Equilibrium (NREE) setting along the lines of Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980), and Verrecchia (1982). Because we will ignore strategic considerations, we now need traders to be risk averse. We assume trader i's utility function is of the CARA type:

$$\tilde{V}_i := -\exp\big(-\gamma[d_i(v-p) - c(z_i)]\big),$$

where z_i is the signal precision acquired, and $c(\cdot)$ its cost, analogous to the baseline model. In the trading stage, traders receive their signal $\theta_i = v + \varepsilon_i$ and solve

$$\max_{d_i} \mathbb{E}\left[\tilde{V}_i \mid \theta_i, p\right]$$

In the information stage, traders choose their precision z_i by solving

$$\max_{z_i} \mathbb{E}\Big[\max_{d_i} \mathbb{E}\left[\tilde{V}_i \mid \theta_i, p\right]\Big]$$

As in the baseline model, we assume signal errors ε_i are independent across traders and independent of both the fundamental v and the noise u. Furthermore, recall that the precision z_i is defined to be related to the overall signal variance $\sigma_i := \mathrm{Var}[\theta_i]$ in that

$$\sigma_i = \sigma(z_i) := \frac{1}{2}\sigma_v(z_i^{-1} + 1)$$

To keep the proper comparison to the baseline model, traders pay a cost to increase z_i, but the results would be very similar if the cost was defined over the precision of the signal error $(\sigma_i - \sigma_v)^{-1} = \frac{2}{\sigma_v}\frac{z_i}{1-z_i}$.[8] The equilibrium asset price is determined via the market clearing condition

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} d_i + u = 0$$

It is standard in the NREE literature to clear markets based on the *average* demand, by analogy to the continuum limit, rather than the total demand; equivalently, some versions explicitly model the noise u grow with the number of traders (i.e., u is the noise per capita).[9] Finally, conjecture (and later verify) that the equilibrium pricing function takes the form

$$p = \beta_v v + \beta_u u,$$

for some β_v and β_u.

In this setup, joint normality of (v, u, θ_i) implies the distribution of (θ_i, p) is

$$(\theta_i, p) \sim \text{Normal}\left(0, \tilde{\Sigma}_i\right),$$
$$\text{where} \quad \tilde{\Sigma}_i := \begin{pmatrix} \sigma_i & \beta_v\sigma_v \\ \beta_v\sigma_v & \beta_v^2\sigma_v + \beta_u^2\sigma_u \end{pmatrix}$$

[8]Note that Verrecchia (1982) defines a cost function $\tilde{c}(s)$ over the error precision $s(z) := \frac{2}{\sigma_v}\frac{z}{1-z}$. Given our chosen cost function $c(z)$, this is implemented simply by putting $\tilde{c}(s) := c(z^{-1}(s)) = c(\frac{\sigma_v s}{2+\sigma_v s})$. However, notice that $\tilde{c}'(0) = \frac{1}{2}\sigma_v c'(0)$, so the critical object—the marginal cost at zero precision—is invariant to this transformation.

[9]One could consider an alternative economy in which the noise is given by $u_n \sim \text{Normal}(0, \sigma_u/n^2)$. In this specification, analogously to our baseline model, write the market clearing condition in levels:

$$\sum_{i=1}^{n} d_i + n u_n = 0$$

The aggregate noise $n u_n$ has constant size σ_u by construction. Dividing the market clearing condition by n and taking the limit, we have

$$\lim_{n\to\infty} \left(\frac{1}{n}\sum_{i=1}^{n} d_i + u_n\right) = 0$$

By the law of large numbers, the result is $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} d_i = 0$ almost-surely, i.e., the asymptotic economy is noise-less. Therefore, a way to analyze a competitive model analogous to our baseline model is to study the small-noise limit *after* having derived the equilibrium for each σ_u.

Furthermore, the conditional distribution of v given (θ_i, p) is

$$
\begin{aligned}
&(v \mid \theta_i, p) \sim \text{Normal}(\tilde{\mu}_{vi}, \tilde{\sigma}_{vi}),\\
\text{where} \quad &\tilde{\mu}_{vi} := \frac{\sigma_v}{(\sigma_i - \sigma_v)\beta_v^2\sigma_v + \sigma_i\beta_u^2\sigma_u}\Big(\beta_u^2\sigma_u\theta_i + (\sigma_i - \sigma_v)\beta_v p\Big)\\
&\tilde{\sigma}_{vi} := \sigma_v - \frac{\sigma_v^2}{(\sigma_i - \sigma_v)\beta_v^2\sigma_v + \sigma_i\beta_u^2\sigma_u}\begin{pmatrix}1\\ \beta_v\end{pmatrix}^{\top}\begin{pmatrix}\beta_v^2\sigma_v + \beta_u^2\sigma_u & -\beta_v\sigma_v\\ -\beta_v\sigma_v & \sigma_i\end{pmatrix}\begin{pmatrix}1\\ \beta_v\end{pmatrix}\\
&\quad = \sigma_v\frac{(\sigma_i - \sigma_v)\beta_u^2\sigma_u}{(\sigma_i - \sigma_v)\beta_v^2\sigma_v + \sigma_i\beta_u^2\sigma_u}
\end{aligned}
$$

Then, the standard CARA solution for asset demand is

$$
d_i = \frac{\tilde{\mu}_{vi} - p}{\gamma\tilde{\sigma}_{vi}}
$$

Before continuing, note that if agents collect a symmetric amount of information, $z_i = z^*$ for all i so that $\sigma_i = \sigma^*$, then market clearing implies by the law of large numbers

$$
p = \frac{\sigma_v}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u}\Big(\beta_u^2\sigma_u v + (\sigma^* - \sigma_v)\beta_v p\Big) + \gamma\sigma_v\frac{(\sigma^* - \sigma_v)\beta_u^2\sigma_u}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u}u
$$

Rearranging this expression, we have

$$
\begin{aligned}
p &= \frac{1}{1 - \frac{\sigma_v(\sigma^* - \sigma_v)\beta_v}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u}}\Big[\frac{\sigma_v\beta_u^2\sigma_u}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u}v + \gamma\sigma_v\frac{(\sigma^* - \sigma_v)\beta_u^2\sigma_u}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u}u\Big]\\
&= \frac{1}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u - \sigma_v(\sigma^* - \sigma_v)\beta_v}\Big[\sigma_v\beta_u^2\sigma_u v + \gamma\sigma_v(\sigma^* - \sigma_v)\beta_u^2\sigma_u u\Big]
\end{aligned}
$$

Matching the loadings on v and u with β_v and β_u, respectively, we obtain the system of two equations in the two unknowns (β_v, β_u):

$$
\begin{aligned}
\beta_v &= \frac{\sigma_v\beta_u^2\sigma_u}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u - \sigma_v(\sigma^* - \sigma_v)\beta_v}\\
\beta_u &= \frac{\gamma\sigma_v(\sigma^* - \sigma_v)\beta_u^2\sigma_u}{(\sigma^* - \sigma_v)\beta_v^2\sigma_v + \sigma^*\beta_u^2\sigma_u - \sigma_v(\sigma^* - \sigma_v)\beta_v}
\end{aligned}
$$

The solution is

$$\beta_v = \frac{\sigma_v[1 + \gamma^2\sigma_u(\sigma^* - \sigma_v)]}{\sigma_v + \gamma^2\sigma^*\sigma_u(\sigma^* - \sigma_v)}$$
$$\beta_u = \gamma(\sigma^* - \sigma_v)\beta_v$$

Thus, we have solved (β_v, β_u) in terms of the symmetric equilibrium information-acquisition z^*, or equivalently σ^*. These results so far also match the limiting economy of Hellwig (1980) for the case of identical risk aversions.

C.1 Equilibrium information-acquisition

Let us now drop i subscripts everywhere, and use the notation $\sigma(z) := \frac{1}{2}\sigma_v(z^{-1} + 1)$ for the individual-specific signal variance. To solve the information choice, a la Verrecchia (1982), first compute the ex-ante utility

$$V(z) := \mathbb{E}\Big[\max_d \mathbb{E}\big[\tilde{V} \mid \theta, p\big]\Big]$$
$$= -\mathbb{E}\Big[\exp\Big(-\frac{1}{2}\frac{(\tilde{\mu}_v - p)^2}{\tilde{\sigma}_v} + \gamma c(z)\Big)\Big]$$

To compute this unconditional expectation, note that $\tilde{\mu}_v$ is linear in (θ, p), which has a normal distribution. The joint density is of (θ, p) is

$$\varphi(\theta, p) = \frac{1}{2\pi}K_0^{1/2}\exp\Big[-\frac{1}{2}\big(k_\theta\theta^2 + 2k_{\theta p}\theta p + k_p p^2\big)\Big]$$
$$K_0 := \frac{1}{\det(\tilde{\Sigma})} = \frac{1}{(\sigma(z) - \sigma_v)\beta_v^2\sigma_v + \sigma(z)\beta_u^2\sigma_u}$$
$$k_\theta = \tilde{\Sigma}_{11}^{-1} = K_0(\beta_v^2\sigma_v + \beta_u^2\sigma_u)$$
$$k_{\theta p} = \tilde{\Sigma}_{12}^{-1} = \tilde{\Sigma}_{21}^{-1} = -K_0\beta_v\sigma_v$$
$$k_p = \tilde{\Sigma}_{22}^{-1} = K_0\sigma(z)$$

Furthermore, write $-\frac{1}{2}\frac{(\tilde{\mu}_v - p)^2}{\tilde{\sigma}_v} = a(b_\theta\theta + b_p p)^2$ where we define

$$a := -\frac{1}{2\tilde{\sigma}_v} = -\frac{1}{2}K_0^{-1}\frac{1}{(\sigma(z) - \sigma_v)\beta_u^2\sigma_v\sigma_u}$$
$$b_\theta := K_0\beta_u^2\sigma_u\sigma_v$$
$$b_p := K_0[(\sigma(z) - \sigma_v)\beta_v\sigma_v - K_0^{-1}]$$

Using this expression and the normal density, compute

$$\mathbb{E}\Big[\exp\Big(-\frac{1}{2}\frac{(\tilde{\mu}_v - p)^2}{\tilde{\sigma}_v}\Big)\Big] = \mathbb{E}[e^{a(b_\theta\theta + b_p p)^2}]$$
$$= \frac{1}{2\pi}K_0^{1/2}\iint e^{a(b_\theta\theta+b_p p)^2}e^{-\frac{1}{2}(k_\theta\theta^2+2k_{\theta p}\theta p+k_p p^2)}\mathrm{d}\theta\mathrm{d}p$$
$$= \frac{1}{2\pi}K_0^{1/2}\iint \exp\Big[-\frac{1}{2}\binom{\theta}{p}^\top \Omega^{-1}\binom{\theta}{p}\Big]\mathrm{d}\theta\mathrm{d}p,$$

where

$$\Omega^{-1} := \begin{pmatrix} k_\theta - 2ab_\theta^2 & k_{\theta p} - 2ab_\theta b_p \\ k_{\theta p} - 2ab_\theta b_p & k_p - 2ab_p^2 \end{pmatrix}$$

Since the integral above involves a normal kernel, we obtain

$$V = -\exp[\gamma c(z)]\frac{K_0^{1/2}}{\det(\Omega^{-1})^{1/2}}$$

After a lengthy amount of algebra, we obtain

$$\Omega^{-1} = \begin{pmatrix} (\sigma(z)-\sigma_v)^{-1} & -(\sigma(z)-\sigma_v)^{-1} \\ -(\sigma(z)-\sigma_v)^{-1} & \sigma_v^{-1} + (\sigma(z)-\sigma_v)^{-1} + \frac{(1-\beta_v)^2}{\beta_u^2\sigma_u} \end{pmatrix}$$

Therefore,[10]

$$\frac{K_0}{\det(\Omega^{-1})} = \Big(\sigma_v^{-1} + (\sigma(z)-\sigma_v)^{-1} + \frac{\beta_v^2}{\beta_u^2\sigma_u}\Big)^{-1}\Big(\beta_u^2\sigma_u + (1-\beta_v)^2\sigma_v\Big)^{-1}.$$

It is equivalent to maximize $-\log(-V)$, so we solve

$$\max_z -\gamma c(z) + \frac{1}{2}\log\Big(\sigma_v^{-1} + (\sigma(z)-\sigma_v)^{-1} + \frac{\beta_v^2}{\beta_u^2\sigma_u}\Big) + \frac{1}{2}\log\Big(\beta_u^2\sigma_u + (1-\beta_v)^2\sigma_v\Big)$$

[10] Note that Verrecchia (1982) makes an algebra mistake in deriving his equation (7). Indeed, in deriving his ex-ante value function, he obtains (after translating into our notation)

$$\frac{K_0}{\det(\Omega^{-1})} = \Big(\sigma_v^{-1} + (\sigma(z)-\sigma_v)^{-1} + \frac{\beta_v^2}{\beta_u^2\sigma_u}\Big)^{-1}\frac{1}{\beta_u^2\sigma_u}.$$

One can check that the mistake originates in his appendix, where he uses (his notation) $a_3 + a_6 = h_0 + s$ rather than the correct expression $a_3 + a_6 = h_0 + s + \frac{(1-\beta)^2}{\gamma^2 V}$. That said, this mistake is inconsequential, because it is multiplicatively separable from the information choice, as can be seen in the subsequent derivations.

Notice that the final term is irrelevant, given z is absent. Also, at this point let us make the replacement $(\sigma(z)-\sigma_v)^{-1} = 2\sigma_v^{-1}\frac{z}{1-z}$. The FOC for z is

$$\gamma(1-z)^2 c'(z)\Big[1+2\frac{z}{1-z}+\frac{\beta_v^2\sigma_v}{\beta_u^2\sigma_u}\Big] \quad \begin{cases} \geq 1, & \text{if } z=0; \\ =1, & \text{if } z\in(0,1); \\ \leq 1, & \text{if } z=1. \end{cases} \tag{C.1}$$

In a symmetric equilibrium where agents choose z^*, we may use the expressions for $\beta_v/\beta_u = 2\gamma^{-1}\sigma_v^{-1}z^*/(1-z^*)$ to rewrite (C.1) as the equilibrium condition

$$\gamma(1-z^*)^2 c'(z^*)\Big[1+2\big(\frac{z^*}{1-z^*}\big)+\frac{4}{\gamma^2\sigma_u\sigma_v}\big(\frac{z^*}{1-z^*}\big)^2\Big] \quad \begin{cases} \geq 1, & \text{if } z^*=0; \\ =1, & \text{if } z^*\in(0,1); \\ \leq 1, & \text{if } z^*=1. \end{cases} \tag{C.2}$$

again with equality if $z^*>0$. Under some conditions on the cost function, Verrecchia (1982) proves that a solution exists to this condition, hence an NREE exists with endogenous information acquisition. Notice that $\gamma c'(0)<1$ suffices to ensure that the solution necessarily satisfies $z^*>0$.

Let us, for reference, also write the pricing function in terms of z^* rather than the variance σ^*. We have

$$p = \beta_v v + \beta_u u \tag{C.3}$$

$$\text{where} \quad \beta_v = \frac{1+\frac{\gamma^2}{2}\sigma_v\sigma_u\frac{1-z^*}{z^*}}{1+\frac{\gamma^2}{4}\sigma_v\sigma_u\frac{1-z^*}{z^*}\frac{1+z^*}{z^*}}$$

$$\beta_u = \frac{\gamma\sigma_v}{2}\frac{1-z^*}{z^*}\beta_v$$

Next, we will examine equilibrium information and its pricing consequences in two limits, $\sigma_u\to 0$ and $\gamma\to 0$. These are relevant for a comparison to our baseline economy, which has bounded aggregate noise (hence vanishing per-capita noise in the large-n limit) and has risk-neutral agents. As we will show, the small-noise limit bears a closer analogy to our main results, in the sense that a fully-revealing equilibrium can emerge if $c'(0)=0$ but not if $c'(0)>0$. By contrast, the risk-neutral limit always features a fully-revealing equilibrium.

C.2 Small-noise limit

As a first result, notice that as noise vanishes ($\sigma_u \to 0$), the left-hand-side of (C.2) blows up for any $z^* > 0$. Hence, a requirement is $z^* \to 0$ as $\sigma_u \to 0$, assuming an equilibrium exists for each σ_u along this sequence.[11] This proves right away that, in a competitive model, individual information collection requires the noise to be growing with the number of traders, such that per-capita noise remains non-trivial. This is why the literature universally adopts this setup.

What happens to the price in this small-noise limit? Although information-collection vanishes individually, so does per-capita noise, and so it is theoretically possible that the price becomes informative. In other words, it becomes a delicate balance of limits. Let us consider the case $0 < \gamma c'(0) < 1$, which ensures that $z^* > 0$ along the sequence of equilibria (if $\gamma c'(0) \geq 1$, then z^* would just collapse to zero for any small enough noise). We will consider the case $c'(0) = 0$ afterward.

If $0 < \gamma c'(0) < 1$, inspection of condition (C.2) shows that $\sigma_u/(z^*)^2$ must not explode as $\sigma_u \to 0$. We can also show that $\sigma_u/(z^*)^2$ cannot be vanishing, because if it did, the left-hand-side of (C.2) would necessarily exceed 1 for all small enough σ_u (given we know $z^* > 0$ along the sequence). Hence, $\sigma_u/(z^*)^2$ has a finite, non-zero limit. Now, returning to the equilibrium price in (C.3), and using this fact, we find that β_v and $\sqrt{\sigma_u}\beta_u$ both converge to a finite constants. The limiting value for β_v is necessarily positive but less than 1. The limiting value for $\sqrt{\sigma_u}\beta_u$ is also necessarily positive and finite. Therefore, we have proven that the small-noise limit, if such equilibrium exists, must still have a noisy price. In particular, prices can never be fully revealing.

If $c'(0) = 0$, the analysis is more delicate. Suppose $z^* \sim A\sigma_u^{\zeta/2}$ for some $A > 0$ and $\zeta > 0$. The left-hand-side of (C.2) behaves asymptotically like

$$\frac{4A^2}{\gamma\sigma_v} c'(A\sigma_u^{\zeta/2})\sigma_u^{\zeta-1} + o(\sigma_u)$$

For this expression to remain non-trivial, hence remain consistent with (C.2), we see that $\zeta < 1$ is required. In that case, $\beta_v \to 1$ and $\sqrt{\sigma_u}\beta_u \to 0$, so that $p \to v$. Thus, assuming an equilibrium exists along this sequence, the economy approaches one with a fully-revealing price and nevertheless traders acquire information all along the sequence.

To develop an understanding, let's assume that $c(z) = \frac{\kappa}{1+\rho} z^{1+\rho}$ for some $\rho > 0$. Then,

[11] We conjecture this existence is in fact true. Studying the equilibrium condition (C.2), all that is required is for z^* to vanish at the order of $\sigma_u^{-1/2}$ if $c'(0) > 0$, or potentially at any slower rate if $c'(0) = 0$.

$c'(z) = \kappa z^\rho$, so

$$\frac{4A^2}{\gamma\sigma_v} c'(A\sigma_u^{\zeta/2})\sigma_u^{\zeta-1} = \frac{4A^{2+\rho}\kappa}{\gamma\sigma_v}\sigma_u^{\zeta-1+\zeta\rho/2}$$

This expression needs to remain non-zero and finite, so we need $\zeta = (1+\rho/2)^{-1}$, which is appropriately below 1. This shows that there exists a sequence $z^*(\sigma_u)$ vanishing at rate $\zeta = (1+\rho/2)^{-1}$ such that the equilibrium condition (C.2) holds with equality for every σ_u small enough. Consequently, equilibria with information-acquisition exist for all σ_u small enough, and these converge to a fully-revealing equilibrium.

C.3 Risk-neutral limit

Second, consider what happens if agents become asymptotically risk-neutral, $\gamma \to 0$. Going back to the equilibrium price, and assuming any information is gathered at all (i.e., z^* converges to a non-zero limit), this makes the price become fully-revealing, i.e., noise-free with $\beta_u \to 0$ and $\beta_v \to 1$. However, the left-hand-side of equilibrium condition (C.2) explodes unless $z^* \to 0$. Thus, a requirement is $z^* \to 0$ as $\gamma \to 0$.

How fast does $z^* \to 0$? As a first possibility, assume $z^* = 0$ for γ sufficiently close to zero, i.e., information vanishes before risk aversion does. This cannot be an equilibrium, because for γ sufficiently small, equilibrium condition (C.2) will also be smaller than 1, a contradiction.

As a second possibility, suppose $z^* \sim A\gamma^\zeta$ as $\gamma \to 0$, with some $A > 0$ and some $\zeta > 0$. Note that the left-hand-side of equilibrium condition (C.2) asymptotically looks like

$$\frac{4A^2}{\sigma_u\sigma_v} c'(0)\gamma^{2\zeta-1} + o(\gamma)$$

If $\zeta > 1/2$, then this expression vanishes in the $\gamma \to 0$ limit, which from (C.2) implies $z^* = 1$ eventually, contradicting the fact that $z^* \to 0$. Therefore, the appropriate rate of convergence for z^* is $\zeta \leq 1/2$. There are two cases. If $c'(0) > 0$, then clearly $\zeta = 1/2$ is required. If $c'(0) = 0$, then $\zeta < 1/2$ is required, with the exact rate of convergence determined by c'' near zero. In either case, the fact that $\zeta \leq 1/2$ means that $\beta_v \to 1$ and $\beta_u \to 0$ as $\gamma \to 0$. Therefore, we have proven that as γ vanishes, equilibrium becomes fully-revealing, if it exists.

C.4 Price informativeness, volatility, and liquidity

Analogous to the baseline model, define the following measures of price informativeness, volatility, and liquidity:

$$\mathcal{I} := \mathrm{Var}[v \mid p]^{-1}$$
$$\mathcal{V} := \mathrm{Var}[p]$$
$$\mathcal{L} := [\partial p / \partial u]^{-1}$$

The informativeness and volatility measures are exactly analogous to our baseline model. The liquidity measure deserves discussion. Note that $\mathcal{L}^{-1}$ is the price response to an exogenous increase in demand, whereas it was the loading of price on aggregate demand in our baseline model. The definition is written this way for the NREE because "aggregate demand" is not well-defined here—it is zero by market clearing. But a noise shock represents an exogenous increase in asset demand that must be absorbed by the informed traders, and hence generates a price impact. As an alternative way to justify our definition of $\mathcal{L}^{-1}$, recall that price impact in the baseline model comes out exactly equal to the price loading on the noise shock u. That is, the result for price impact in our baseline model coincides with $\partial p / \partial u$ here. By contrast, volume is not well-defined in the competitive economy, so we do not examine it.

Using expression (C.3) and properties of the joint normal distribution, compute

$$\mathcal{I} = \sigma_v^{-1}\Big(\frac{\beta_v^2 \sigma_v + \beta_u^2 \sigma_u}{\beta_u^2 \sigma_u}\Big)$$
$$\mathcal{V} = \beta_u^2 \sigma_u + \beta_v^2 \sigma_v$$
$$\mathcal{L} = \beta_u^{-1}$$

Recall that in the limits $\sigma_u \to 0$ or $\gamma \to 0$, the equilibrium can become fully-revealing (for $\sigma_u \to 0$, recall this additionally required $c'(0) = 0$). The consequences on these measures is intuitive. For the risk-neutral limit, taking $\gamma \to 0$ with $z^* \sim \gamma^{\zeta}$ and $\zeta \leq 1/2$, notice that $\mathcal{I} \to +\infty$ and $\mathcal{V} \to \sigma_v$. For the small-noise limit, taking $\sigma_u \to 0$ with $z^* \sim \sigma_u^{\zeta}$ and $\zeta < 1/2$, notice that $\mathcal{I} \to +\infty$ and $\mathcal{V} \to \sigma_v$. Unlike our strategic-trading model, these limiting cases are the only ones where the fully-efficient "magical markets" outcomes arise in an NREE.

C.5 Two examples: zero and positive marginal cost at zero

We will evaluate the equilibrium for two different cost functions, the first representing a zero marginal cost (i.e., $c'(0) = 0$) and the second having positive marginal cost at zero (i.e., $c'(0) > 0$).

Example 1 (quadratic cost). Consider the cost function

$$c(z) = \frac{\kappa}{2}\Big(\frac{z}{1-z}\Big)^2. \tag{C.4}$$

This function satisfies $c'(0) = 0$ and therefore is capable of generating infinite price informativeness in our baseline Kyle model in the text. To provide an interpretation, recall that the chosen precision over the signal error ε_i satisfies $s_i = \frac{2}{\sigma_v}\frac{z_i}{1-z_i}$. Hence, the function (C.4) represents quadratic costs over this signal error precision.

Evaluating the equilibrium condition (C.2) for this case, we have

$$\gamma\kappa\Big[\frac{z^*}{1-z^*} + 2\Big(\frac{z^*}{1-z^*}\Big)^2 + \frac{4}{\gamma^2\sigma_u\sigma_v}\Big(\frac{z^*}{1-z^*}\Big)^3\Big] = 1 \tag{C.5}$$

Thus, $z^*/(1-z^*)$ solves a cubic equation. There is exactly one root satisfying $z^* \in (0,1)$.

Example 2 (linear cost). Now, consider the cost function

$$c(z) = \chi\Big(\frac{z}{1-z}\Big), \tag{C.6}$$

where we assume $\gamma\chi < 1$. This function satisfies $c'(0) = \chi > 0$ and therefore is comparable to the case analyzed for our baseline Kyle model in the text. (One has to be careful to make comparisons by varying χ, however, because here χ modulates the entire cost function in addition to the marginal cost at the prior $c'(0)$.) This also means that the information costs are "larger" here than for the quadratic case above. Indeed, given that the chosen precision over the signal error ε_i satisfies $s_i = \frac{2}{\sigma_v}\frac{z_i}{1-z_i}$, the function (C.6) represents linear costs over this signal error precision.

Evaluating the equilibrium condition (C.2) for this case, we have

$$\gamma\chi\Big[1 + 2\Big(\frac{z^*}{1-z^*}\Big) + \frac{4}{\gamma^2\sigma_u\sigma_v}\Big(\frac{z^*}{1-z^*}\Big)^2\Big] = 1 \tag{C.7}$$

Thus, $z^*/(1-z^*)$ solves a quadratic equation. Assuming $\gamma\chi < 1$, there exists a unique positive solution z^*. (If $\gamma\chi > 1$, then there is no positive solution, so $z^* = 0$ is the

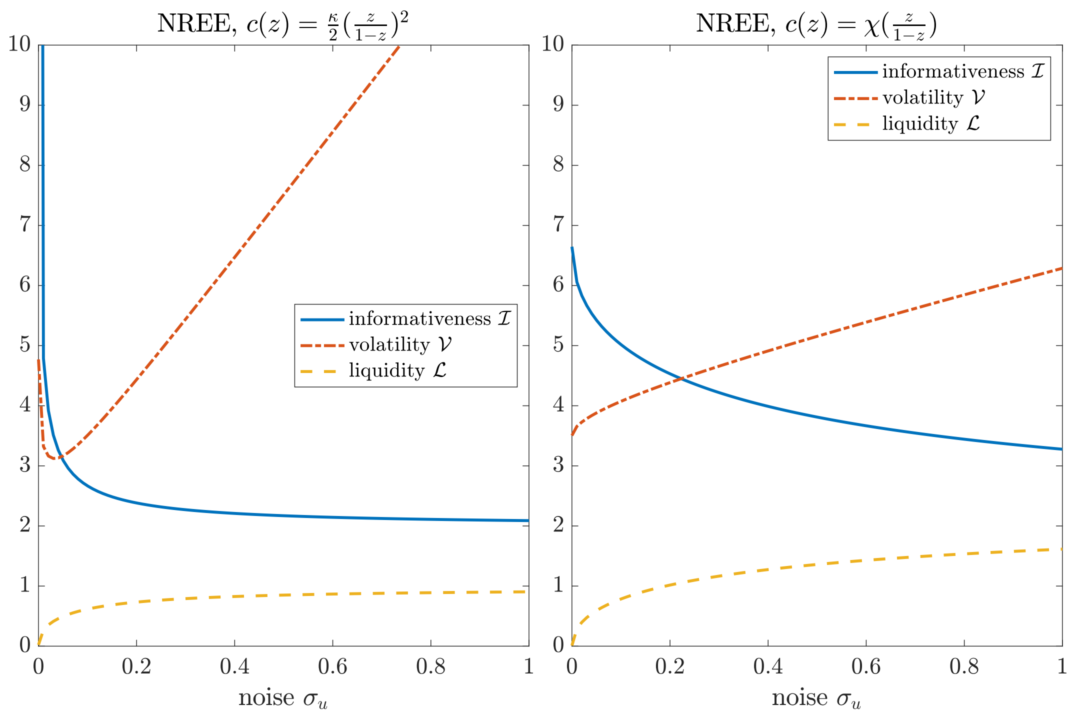

Figure C.1: Measures as a function of σ_u. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.

equilibrium.) This unique solution is

$$\frac{z^*}{1-z^*} = \frac{\gamma\sigma_u\sigma_v}{4\chi}\Big[-\gamma\chi + \sqrt{(\gamma\chi)^2 + (1-\gamma\chi)\frac{4\chi}{\gamma\sigma_u\sigma_v}}\Big] \tag{C.8}$$

Market measures in the examples. Figures C.1-C.3 display informativeness, volatility, and liquidity in the two examples. We vary σ_u, σ_v, and γ one at a time in the figures. Broadly speaking, the examples generate similar behavior qualitatively. Furthermore, note that the figures confirm the limiting theoretical analysis from earlier. As $\gamma \to 0$ or $\sigma_v \to 0$, the equilibrium becomes fully-revealing, with $\mathcal{I} \to +\infty$ and $\mathcal{V} \to \sigma_v$. As $\sigma_u \to 0$, the equilibrium becomes fully-revealing only if $c'(0) = 0$ but not if $c'(0) > 0$.

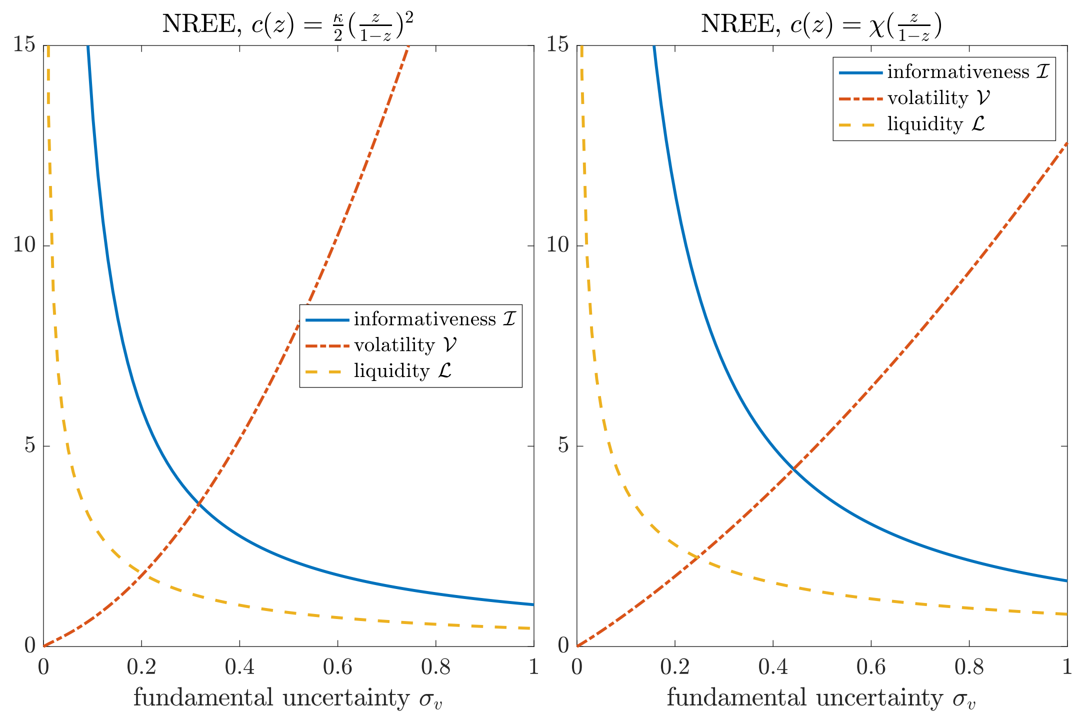

Figure C.2: Measures as a function of σ_v. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.

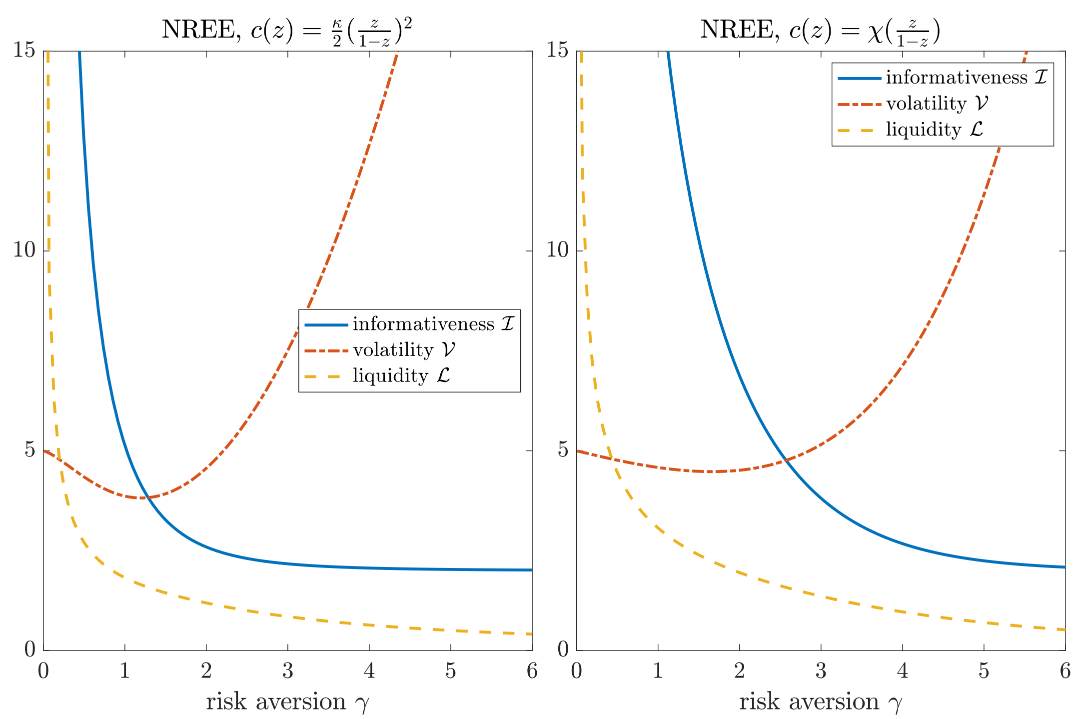

Figure C.3: Measures as a function of γ. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\sigma_u = 0.5$, $\sigma_v = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.