Liquidity and Information Acquisition in Large Markets*

Paymon Khorrami† João Thereze‡

March 9, 2026

Abstract

We study a financial market with free entry of strategic traders and information production. Despite the market being large and competitive, microstructure frictions shape equilibrium outcomes: information acquisition is more attractive when liquidity is higher. Unlike conventional frameworks for competition, which abstract from this mechanism, our model features an increasing relationship between retail order flow volatility and price informativeness, with implications for market design. Equilibrium information aggregation is determined by a sufficient statistic, which depends on the information technology only through the marginal cost of information at the prior. In contrast to Grossman and Stiglitz (1980), equilibrium prices may successfully aggregate information, and we provide necessary and sufficient conditions for full revelation.

*We thank Christian Opp, Cecilia Parlatore, Shri Santosh, Savi Sundaresan, 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

How does liquidity affect information aggregation in competitive markets? Even in modern financial markets that are highly competitive and often quite liquid, market depth varies substantially across assets and over time (Amihud, 2002; Amihud and Noh, 2021; Chordia et al., 2005; Næs et al., 2011). To the extent that liquidity affects incentives to trade, it modulates how valuable private information is to traders. But it is not clear that liquidity matters for traders in highly competitive markets, when prices are approximately taken as given. Whether or not liquidity affects trading incentives, and information-acquisition as a consequence, may thus matter a great deal for information aggregation and price efficiency.

In this paper, we present a simple framework for studying how liquidity and price informativeness are jointly determined in a large market. For that, we study free-entry and information acquisition in a strategic trading model (Kyle, 1985; Lambert et al., 2018). Because strategic traders internalize price-impact, this setting allows us to investigate the effects of liquidity on trading behavior. At the same time, because free entry leads to a large market with infinite players, equilibrium is competitive in that strategic traders have no market power. The key takeaway is that liquidity, a friction associated with the microstructure of thin markets, has equilibrium effects in large markets. This liquidity effect generates substantively different predictions from competitive Noisy Rational Expectation Equilibrium (NREE) models, the workhorse framework for studying information in large markets. In particular, the correlation between the volatility of retail order flow and informativeness is opposite in our model compared to NREE, with policy implications for market design.

Model and Results. Our baseline model starts with 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 symmetric equilibrium with free-entry, that is, when the number of traders is determined by a zero ex-ante profits condition. In our setting, the free-entry equilibrium is the large market limit ($n \to \infty$) of this model, in which individual price impacts vanish, resembling perfectly competitive markets. Although our baseline model displays risk-neutral investors who submit market-orders to a competitive market-maker, we show that our main takeaways obtain when traders are risk-averse and submit limit orders, as long as there is free-entry of uninformed traders.

In contrast with most of the literature, we allow the information technology to be smooth, implying that traders may decide to enter to acquire a small amount of information. Concretely, 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 informativeness is an increasing function of the following sufficient statistic:

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

Our first 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" even when information production is costly.

We proceed to characterize the large-market equilibrium for all values of Γ. As Γ falls, information aggregation declines monotonically and ranges the full spectrum of price informativeness—from completely uninformative to fully revealing. 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. We then use our characterization of informativeness to obtain closed-form solutions for prices, liquidity, price-volatility and trading volume.

The second main result concerns the relationship between liquidity and informativeness in our model. Formally, it establishes that when the volatility of retail order flow increases, so do liquidity and informativeness. This result is in stark contrast with NREE, where the same increase in non-fundamental volatility leads to higher *lower* informativeness. The distinction between these two comparative statics can be explained by understanding the three channels through which noise volatility affects equilibrium, one of which is exclusive of strategic trading models. First, the direct effect of an increase in non-fundamental volatility is to mechanically decrease price informativeness. This is because, for the same amount of private information, the signal-to-noise ratio in the market must decrease. Second, this direct decrease in informativeness leads to a higher value of information for traders: because prices are wrong more often, arbitrage opportunities have become more profitable. Both the direct effect and this indirect profitability effect are present in NREE as well as in our model. The net effect of these two channels is always less information aggregation. However, our model also possesses a liquidity effect that is absent in NREE models: when noise increases, the market becomes deeper and market-makers are less wary of adverse selection, improving liquidity. The increase in liquidity boosts the profitability of trading, increasing incentives to acquire information even further. In this presence of this liquidity effect, informativeness always rises in noise volatility.

The result above has implications for several markets. For example, markets for retail order flows are highly segmented. In particular, a large share of retail orders flow to intermediaries called wholesalers, who may execute those orders outside of large exchanges. The justification is that wholesalers are willing to provide discounts to retail orders, which are largely uninformed, to the benefit of retail traders. Our model suggests caution: by removing retail order flow from the exchange, incentives for information production fall, which may lead to a decline in price informativeness. To the extent that passive investing also reduces retail noise in exchanges, the same argument could follow, providing yet another reason why passive investment worsens information aggregation.

We conclude by providing a mapping between our framework and market observables. We show one can use our model to identify 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 low price-informativeness (Dávila and Parlatore, 2018), 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 trading side of our model is a special case of the latter paper, which we extend by allowing endogenous information acquisition. One of our main results is 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

[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).

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 values, such as auctions (Atakan and Ekmekci, 2023) and voting (Martinelli, 2006). Like these papers, we 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 with multiple potentially informed traders and costly information acquisition. Information-acquisition is decided ex-ante, subject to a cost for signal precision. The baseline model below features risk-neutral traders who submit market orders to a competitive market-maker, following Kyle (1985) and Lambert et al. (2018). Neither risk-neutrality nor the specific trading protocol are fundamental. In Section 5, we discuss how our results remain the same if trading occurs through limit-orders instead of market-orders, and how they extend qualitatively to the case of risk-aversion, as long as there is free entry of uninformed traders.

A single risky asset with random payoff v is traded in a market with n risk-neutral strategic traders. In this market, there is also a representative uninformed market maker and unmodeled noise traders that submit demand u. We assume that noise trader demand and fundamentals are independently normally distributed with mean normalized to zero: $v \sim \text{Normal}(0, \tau_v^{-1})$, $u \sim \text{Normal}(0, \tau_u^{-1})$, and $u \perp v$.

Before trading, investors have an opportunity to collect information about v in the form of a signal about the security value. Formally, strategic trader i has access to a signal $s_i = v + \varepsilon_i$, where ε_i is the error in the signal, which is normally distributed with variance τ_i^{-1}. At a cost, the investor can choose the variance of the error ε_i. We assume that trader's errors are independent of each other and fundamentals: $\varepsilon_i \perp \varepsilon_j$ for any pair of traders (i, j), and $\varepsilon_i \perp v$.

The trading protocol follows Kyle (1985). Each trader submits a market order d_i, and noise trader demand u realizes. The market-maker observes aggregate demand $Y = \sum_i d_i + u$ and sets a break-even price given the information in the order flow:

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

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

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

Prior to trading, information is gathered subject to a cost. For convenience, we will let traders choose a normalization of information precision, $z_i \in [0, 1]$, which is defined as follows, using the precision of trader i's error, τ_i^{-1}:

$$z_i := \frac{\tau_i}{\tau_i + \tau_v}. \tag{2}$$

If $z_i = 0$, then $\tau_i = 0$, and trader i's signal is entirely uninformative. If $z_i = 1$, then $\tau_i = \infty$, and the signal is perfectly informative. Indeed, the precision of an agent's signal is a monotonic function of z_i.

Given a collection of precisions, $\mathbf{z} = (z_i)_{i=1}^n$, trader i chooses d_i to maximize expected profits, conditional on her signal realization s_i. Because traders are strategic, they anticipate that their order will change prices, given by a function $\mathbf{p}(\cdot, \mathbf{z})$ of the aggregate demand. Moreover, traders take as given the demand functions of other traders, $\mathbf{d}_j(\cdot, \mathbf{z})$, which maps trader j's private information to a market order. We let $\tilde{V}_i$ represent i's profits given a collection of precisions conditional on a signal realization. The trader's problem is then:

$$\tilde{V}_i(s_i|\mathbf{z}) := \max_{d_i} \mathbb{E}\left[d_i \left(v - \mathbf{p}\Big(\sum_{j\neq i} \mathbf{d}_j(s_j,\mathbf{z}) + d_i + u,\ \mathbf{z}\Big)\right) \middle| s_i \right]. \tag{3}$$

Before the trading stage, trader i chooses how much information to collect while taking all other traders' precisions as given. That choice trades-off the benefits of information in the subsequent trading equilibrium against the cost of information acquisition:

$$\max_{z\in[0,1]} \mathbb{E}\left[\tilde{V}_i(s_i|z,\mathbf{z}_{-i})\right] - c(z), \tag{4}$$

where $c : [0,1] \to \mathbb{R}_+$ is a continuously differentiable and strictly convex function satisfying $c(0) = 0$, $c'(0) \geq 0$, and $c'(1-) = \infty$. The latter condition is simply for ease of exposition. Moreover, $\mathbf{z}_{-i}$ represents the vector of precisions of all other traders. We are now ready to define equilibrium.

Definition 1. *An equilibrium is a collection of precisions, $\mathbf{z}$, demand functions $(\mathbf{d}_i)_{i=1}^n$, with $\mathbf{d}_i : \mathbb{R} \times [0,1]^n \to \mathbb{R}$, and a price function $\mathbf{p} : \mathbb{R} \to \mathbb{R}$, such that:*

1. *For any collection of precisions $\tilde{\mathbf{z}}$, $\left\{(\mathbf{d}_i(\cdot,\tilde{\mathbf{z}}))_{i=1}^n, \mathbf{p}(\cdot,\tilde{\mathbf{z}})\right\}$ is a linear trading equilibrium. That is:*
   - *(a) Traders optimize: for each trader i, given prices, other traders' demand functions, and signal realization s_i, demand $\mathbf{d}_i(s_i,\tilde{\mathbf{z}})$ solves (3).*
   - *(b) Market-makers break even: $\mathbf{p}(Y,\tilde{\mathbf{z}}) = \mathbb{E}[v|Y]$,*
   - *(c) Demand and prices are linear: there exist constants $\lambda(\tilde{\mathbf{z}}),\beta_i(\tilde{\mathbf{z}})$, such that: $\mathbf{p}(Y,\tilde{\mathbf{z}}) = \lambda(\tilde{\mathbf{z}})Y$, and $\mathbf{d}_i(s_i,\tilde{\mathbf{z}}) = \beta_i(\tilde{\mathbf{z}})s_i$.*
2. *$\mathbf{z}$ is an information equilibrium. That is, z_i solves (4) given $\mathbf{z}_{-i}$.*

Regarding the notation above, we emphasize that $\mathbf{d}_i$ and $\mathbf{p}$ are used to as demand and pricing *functions*, which are defined for an arbitrary vector of precisions $\mathbf{z}$, whereas d_i and p are the equilibrium demand and price (not functions but random variables). While this notation is rather cumbersome in principle, most of the analysis below does not need it. Indeed, we show below that our notation substantially simplifies in terms of a few sufficient summary statistics.

2.1 Linear trading equilibrium

For any information choice decision, $\mathbf{z}$, we can compute the equilibrium of the associated trading game. Although the result is a special case of Lambert et al. (2018), it is useful to express the equilibrium using the economic quantities we explore later. To that end, we define two critical statistics. Define relative price informativeness as the fraction of the fundamental variance left after observing prices in an equilibrium:[2]

$$I(\mathbf{z}) := 1 - \frac{\mathrm{Var}[v|p,\mathbf{z}]}{\mathrm{Var}[v]}.$$

As $I(\mathbf{z}) \to 1$, prices become perfectly informative, whereas $I(\mathbf{z}) \to 0$ corresponds to zero information revealed by prices. Moreover, we measure liquidity as the inverse of the price impact of a marginal trade, which is the inverse of Kyle's lambda:

$$L(\mathbf{z}) := \frac{1}{\lambda(\mathbf{z})}.$$

Our first result characterizes the unique linear trading equilibrium through these two variables and trader i's value function.

Lemma 1. *Fix any* $\mathbf{z}$. *There is a unique linear trading equilibrium. In that equilibrium, informativeness, liquidity, and trader i's value are given by*

$$\begin{aligned}
I(\mathbf{z}) &= \frac{\mathbf{z}\cdot\mathbf{1}}{1+\mathbf{z}\cdot\mathbf{1}} \\
L(\mathbf{z}) &= \sqrt{\frac{2\tau_v}{\tau_u}}\frac{1+\mathbf{z}\cdot\mathbf{1}}{\sqrt{\mathbf{z}\cdot\mathbf{1}+\mathbf{z}\cdot\mathbf{z}}} \\
V_i(\mathbf{z}) := \mathbb{E}[\tilde{V}_i(s_i,\mathbf{z})] &= \frac{1}{2\tau_v}(1-I(\mathbf{z}))^2 L(\mathbf{z})\ (z_i+z_i^2)
\end{aligned}$$

Define the *aggregate information* in the market as the sum of all normalized precisions, $\mathbf{z}\cdot\mathbf{1}$. Lemma 1 shows that more information in the market mechanically increases price informativeness, but it naturally reduces liquidity by worsening adverse selection. Moreover, liquidity depends not only on the level, but also on the dispersion of information: when information is more equally shared among traders, the market is more liquid. Finally, the above result characterizes the equilibrium expected payoff of trader i.

Trader i's indirect utility function captures three economic forces. First, the individual trader benefits from being better informed, as she can use that information to submit

[2] Many definitions of price-informativeness have been used in the literature. We follow the usage in Grossman and Stiglitz (1980).

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. Second, a higher informativeness reduces the marginal payoff of information, because equilibrium prices are closer to the real valuation of the asset. Finally, and crucially to our paper, liquidity also affects the value of information: a deeper market allows traders to leverage their private knowledge more, as prices do not move as much against their own trades. That means that the value of information is larger in deeper markets.

2.2 Information equilibrium

Next, we describe the information equilibrium, where traders choose $z_i \in [0,1]$ to solve (4). We focus on a symmetric equilibrium, in which $\mathbf{z} = z^n \mathbf{1}$ for some scalar z^n, which depends on the number of traders in the market, n. Thus, $\mathbf{z} \cdot \mathbf{1} = nz^n$ is the "aggregate information" in the market. We will let p^n denote the random variable that represents equilibrium prices in this symmetric equilibrium with n traders. 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.

At a symmetric information equilibrium, we must have

$$\frac{dV_i(z^n \mathbf{1})}{dz_i} \leq c'(z^n), \quad \text{with equality if} \quad z^n > 0. \tag{5}$$

Therefore, the fundamental quantity determining equilibrium is the marginal value of information. From Lemma 1, one can show:

$$\frac{dV_i(z, \mathbf{z}_{-i})}{dz_i} = \frac{1}{2\tau_v}(1 - I(\mathbf{z}))^2 L(\mathbf{z})(1 + 2z) + o(\mathbf{z}) \leq \frac{1}{2\tau_v}(1 - I(\mathbf{z}))^2 L(\mathbf{z})(1 + 2z), \tag{6}$$

where $o(\cdot)$ is a negative function such that $o(\mathbf{z}) \xrightarrow{z \to 0} 0$, when $\mathbf{z}_{-i} > 0$, and, $o(z\ \mathbf{1}) \xrightarrow{z \to 0} 0$.

For any fixed size-n economy, there is an information equilibrium with positive information acquisition, as long as $\chi := c'(0) < \infty$. Indeed, notice that liquidity $L(\mathbf{z})$ grows without bound as all agents become uninformed. In that case, (6) shows that $\frac{dV_i(z\ \mathbf{1})}{dz_i} \xrightarrow{z \to 0} \infty$, so traders are always willing to pay some cost to acquire a little bit of information.

The remainder of this section studies the *competitive* limit of this market: that is, the limit as the number of traders grows large. Importantly, we aim to characterize the information collection in that limit. We start by establishing that each trader must become informationally small as the market grows.

Proposition 1. *Individual information-gathering vanishes asymptotically in a symmetric equilibrium:* $z^n \xrightarrow{n\to\infty} 0$.

Proposition 1 results from the strategic substitutability in information-gathering. If z^n does not converge to zero, then the aggregate amount of information in the market must be growing without bound. But if that is the case, it becomes impossible for individual traders to disguise their demand among noise, which drives the benefit of information acquisition to zero, contradicting z^n bounded away from zero.

We are interested in price informativeness in the competitive limit of a large market: as the number of traders, n, grows to infinity. Define competitive price informativeness as the large-market limit of price-informativeness, I.

Definition 2. *Competitive price informativeness is*[3]

$$\mathcal{I} = \lim_{n\to\infty} I(z^n\, \mathbf{1}).$$

Our first key result, below, characterizes price informativeness in large markets. Despite the fact that individuals are informationally small in the competitive limit, we show that price informativeness, and aggregate information, can become arbitrarily large.

Theorem 1. *Competitive informativeness* $\mathcal{I} \in [0,1]$ *is the unique solution to:*

$$\frac{\mathcal{I}}{(1-\mathcal{I})^3} = \Gamma, \tag{7}$$

where Γ *is defined by*

$$\Gamma := \frac{1}{2}\frac{1}{\chi^2 \tau_u \tau_v}. \tag{8}$$

Theorem 1 characterizes the asymptotic level of informativeness in the market as a function of three primitives: noise trader uncertainty, τ_u^{-1}, fundamental uncertainty, τ_v^{-1}, and the marginal cost of the first piece of information, χ. Prices naturally are more informative when the marginal cost of information at the prior, χ, is lower. Moreover, prices are more informative either when fundamental uncertainty or noise uncertainty

[3] In the proof of Theorem 1, this limit is shown to exist.

are higher. This characterization is simple, clean, and relatively detail-free with respect to the information cost function, depending only on its derivative at the prior mean, χ.

As χ varies from 0 to ∞, the statistic Γ spans the whole real line and, similarly, informativeness can take any value. 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.

A simple observation that follows from Theorem 1 is that one can obtain any level of informativeness in equilibrium, depending only on the marginal cost of information at zero. We formalize this result below:

Corollary 1. *Prices are asymptotically fully-revealing—i.e. $\mathcal{I} = 1$—if and only if $\chi = 0$.*

Remarkably, this observation is in contrast with NREE, and deserves some comments. 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. If marginal costs of information at the prior are positive, this will eventually lead any particular individual to acquire exactly zero information, which cannot lead to fully-revealing prices. On its turn, sufficiency follows from a very geenral 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.[4]

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

[4]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.

A.5. 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 a rate $n^{3/5}$. 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.[5]

3 Liquidity, Informativeness, and Noise Trading

In this section we compare the predictions of our model with its NREE counterparts. The central observation is that as noise-trading demand becomes more volatile, prince informativeness moves in opposite directions compared to NREE: while in NREE informativeness must go down, the model with strategic traders predicts more informative markets. We also show how other observables, like price volatility and volume, change in response to movements on primitives.

We start by defining a set of equilibrium variables in the large economy and deriving their closed-form solutions.

Definition 3. Liquidity $\mathcal{L}$ *is the large-market limit of small-market liquidity, i.e.,*

$$\mathcal{L} := \lim_{n\to\infty} L(z^n\,\mathbf{1}).$$

Definition 4. Price volatility $\mathcal{V}$ *is the unconditional standard deviation of the asymptotic price, i.e.,*

$$\mathcal{V} := Std\big[\lim_{n\to\infty} p^n\big]$$

Definition 5. Trading volume $\mathcal{Y}$ *is the expected asymptotic absolute equilibrium demand, in dollars:*

$$\mathcal{Y} := \mathbb{E}\big[\big|\lim_{n\to\infty} p^n Y_n\big|\big]$$

We now characterize all these quantities, as well as asymptotic prices.

Proposition 2. *In the competitive limit, prices converge in distribution to:*

[5] In particular, a model with $\chi = 0$ has full information emerging as $n \to \infty$ even if noise $\sqrt{\tau_u^{-1}}$ 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.

$$\lim_{n\to\infty} p^n = \mathcal{I}\, v + \sqrt{\tau_v^{-1}\mathcal{I}(1-\mathcal{I})}\, \eta,$$

for some $\eta \sim \mathcal{N}(0,1)$, such that $\eta \perp v$.

Liquidity, price volatility, and trading volume are given by the following:

$$\begin{aligned}
&\textit{(liquidity)} \quad \mathcal{L} = \sqrt{\frac{2\tau_v}{\tau_u}}\frac{1}{\sqrt{\mathcal{I}(1-\mathcal{I})}} \\
&\textit{(volatility)} \quad \mathcal{V} = \sqrt{\frac{\mathcal{I}}{\tau_v}} \\
&\textit{(volume)} \quad \mathcal{Y} = \sqrt{\frac{2}{\tau_u\tau_v}\frac{\mathcal{I}}{1-\mathcal{I}}}
\end{aligned}$$

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

The magic of markets. If $\chi = 0$, the results of Proposition 2 specialize to $\mathcal{L} = \infty$ (perfect liquidity), $\mathcal{I} = 1$ (fully informative prices), $\mathcal{V} = \sqrt{\tau_v^{-1}}$ (price is fundamental), and $\mathcal{Y} = \infty$ (infinite volume). These results hold regardless of the amount of fundamental uncertainty τ_v^{-1} and the amount of noise τ_u^{-1}, 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": when the marginal cost of the first piece of information is zero, markets generically aggregate dispersed information, despite information being endogenous. This is in line with the research on information aggregation in common value auctions Atakan and Ekmekci (2023), which also finds that dispersed information can be well-reflected in prices if and only if the marginal cost of information is zero.

Given the novelty of this "magical markets" outcome, let us pause to highlight the differences from perfectly competitive noisy rational expectations equilibria (NREEs). Appendix D studies a version of such an NREE with endogenous information, as in Verrecchia (1982), and shows that $\mathcal{L} = \infty$, $\mathcal{I} = 1$, and $\mathcal{V} = \sqrt{\tau_v^{-1}}$ only arise in some limiting cases. In particular, the magical markets outcomes only arise in the risk-neutral limit, or if $\tau_v^{-1} \to 0$, or if $\tau_u^{-1} \to 0$ (in addition to the information cost function having $c'(0) = 0$). That is, information costs alone cannot lead to generically magical markets in an NREE.[6]

[6] 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 τ_u^{-1} small enough (Grossman

Markets under partial efficiency. We continue the analysis by next considering the empirically-relevant case of partial efficiency: $\chi > 0$. Our main result is that, as noise grows, price informativeness grows in our model, whereas it falls in NREE.[7] This comparative statics is the core empirical implication of our model, which allows it to be tested against NREE.

Proposition 3. *$\mathcal{I}$ is decreasing with τ_u.*

Proof. The result is immediate from Theorem 1. Indeed, the left-hand-side of (7) is increasing in informativeness, and the right-hand-side, Γ, is decreasing in τ_u. $\square$

Proposition 3 states that as the uncertainty associated with noise trading increases, so does informativeness. This result is striking in that it is the opposite of what is obtained in NREE (Verrecchia, 1982). The reasoning behind this is that, just as in microstructure models, liquidity plays a role in trader's behavior even in large markets. In other words, the proposition shows how, even as markets become large, microstructure frictions may persist.

To be concrete, an increase in noise-demand uncertainty has three effects in equilibrium informativeness. The first effect, present in both NREE models and in strategic trading models, is direct: an increase in noise traders makes prices mechanically less informative. The second effect is that this reduction in informativeness increases the marginal value of information by creating profitable opportunities—deviations between the price and the fundamental value—which lead to an increase in information acquisition. This force, also present in NREE, is not enough to compensate for the mechanical deterioration of informativeness. It is the third force, unique to models of strategic trading, which reverses that deterioration, leading to higher informativeness. We call it the liquidity effect: the increase of noise traders deepens the market, making the trader's *marginal* trade more profitable, as prices move less against her. Because she is able to trade more aggressively, the trader is willing to acquire even more information. On

and Stiglitz, 1980), or equilibrium informativeness is bounded above for all τ_u^{-1} (Verrecchia, 1982). Verrecchia (1982) says, for instance, "As V [equivalent to our τ_u^{-1}] 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 $\tau_u^{-1} \to 0$ is that we allow information costs to satisfy $c'(0) = 0$. In that case, it is true that information-collection vanishes as $\tau_u^{-1} \to 0$, but the noise in the price vanishes faster, allowing fully-revealing prices asymptotically.

[7] For the result in NREE, see Verrecchia (1982) or Appendix D.

aggregate, prices become more informative exactly because the market becomes more liquid.

This distinction between our model and NREE is relevant for several applications. One example is the increasing routing of retail-flow to off-exchange venues. To the extent that a large part of retail volume is executed off-exchange, the uncertainty of on-exchange noise trading is reduced. NREE suggests that this segmentation brings no trade-offs, because exchange price should become more informative, and retail traders benefit from better deals with wholesalers. Our model suggests caution: the decrease in noise should deteriorate the information quality on the exchange.

An interesting consequence of the comparative statics above is that the amount of adverse selection in the market, which is proxied by informativeness, correlates positively with liquidity as noise trading volatility varies. The logic is that strategic traders acquire more information when markets are deeper, which happens when noise trading abounds. More generally, an increase in noise trading raises the amount of private information in the market, leading to more volume and volatility. This channel was noted by Admati and Pfleiderer (1988), and is consistent with empirical evidence (Collin-Dufresne and Fos, 2015).[8]

Finally, as information costs, χ, vary, liquidity first declines and then rises again. This behavior is intuitive: when information costs are 0, the market is perfectly informative and perfectly liquid. Indeed, the market-maker perfectly knows the value of the asset, v in equilibrium, and therefore sets equilibrium prices exactly equal to v. Given that amount of information, no variation in aggregate demand affects the price—the market is perfectly deep despite an abundance of private information because all this information is reflected in prices. On the other extreme, as information costs become arbitrarily large, there is no private information in the market. As a consequence of the absence of adverse selection, all demand stems from noise traders, and the market-maker can make zero profits by simply posting a price of 0. This intuitive behavior of liquidity does not necessarily hold in NREE. There, markets are never perfectly liquid. For an NREE to exhibit perfect liquidity, it would need to have perfect price informativeness, which is an impossibility. On the other extreme, of no information acquisition, price impacts are bounded and depend only on fundamental volatility, not noise trader demand. Because of that, it is often—but not always—the case that liquidity decreases monotonically with information costs in NREE models.

[8]Another explanation for this positive correlation between adverse selection and liquidity, proposed in Collin-Dufresne and Fos (2016), is that informed investors time the market to trade when liquidity is higher.

4 Identifying Price Informativeness

We now provide a proof-of-concept on how to identify informativeness and the primitives of our model from observables. To that end, we leverage the closed-form solutions obtained in Theorem 1 and Proposition 2. Although there are several methodologies suitable for identifying informativeness (Bai et al., 2016; Dávila and Parlatore, 2018; Farboodi et al., 2022), our approach overcomes a number of difficulties faced by previous work. Concretely, to identify informativeness, the existing literature needs to make assumptions about which fundamentals prices are supposed to predict, and at which time window. Moreover, because the usual empirical methods rely on accounting data, price-informativeness can only be identified at low-frequencies. Our structural approach sidesteps these issues entirely, at the cost of taking the model seriously: it uses equilibrium variables as observables, and relies on the model mapping to identify both the primitives of the model as well as price informativeness.

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 $\mathcal{Y}$ 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 τ_v^{-1}, τ_u^{-1} 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 2 imply

$$\mathcal{L} = \frac{\mathcal{Y}}{\mathcal{V}^2}.$$

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 precision, and non-fundamental precision follows:*

$$d\log\mathcal{I} = \frac{1-\mathcal{I}}{1+\mathcal{I}}\left(d\log\mathcal{L} + 2d\log\mathcal{V}\right), \tag{9}$$

$$d\log\tau_v = \frac{1-\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{L} - \frac{4\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{V}, \tag{10}$$

$$d\log\tau_u = -\frac{2+\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{L} - \frac{2}{1+\mathcal{I}}d\log\mathcal{V}. \tag{11}$$

Given an initial value for informativeness, $\mathcal{I}_0$ and given observables $\mathcal{L}_0$, $\mathcal{V}_0$, one can obtain the initial values for $\tau_{v,0}$ and $\tau_{u,0}$.

Proposition 4 provides a method to calculate precisely how price informativeness and uncertainty 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 informativeness $\mathcal{I}_0$. While $\mathcal{I}_0$ may be hard to observe or estimate directly, one can perform sensitivity analyses: sample the input $\mathcal{I}_0$ from a statistical prior, compute paths for $(\mathcal{I}_t, \tau_{v,t}, \tau_{u,t})_{t\geq 0}$, and then ask whether the range of paths (one path for each $\mathcal{I}_0$) shares 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 $\mathcal{I}_0$, but the process is easily implementable with Proposition 4.

An alternative approach is to rely on the reduced-form estimates in the literature to chooce $\mathcal{I}_0$. Remarkably, estimates of relative price informativeness tend to be low. Dávila and Parlatore (2018), for example, obtain values close to 2.5%. Our next identifying assumption allows us to approximate the dynamics of the market under the assumption that informativeness if low.

Assumption 2. *Informativeness is low: $\mathcal{I} \approx 0$*

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

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

$$d\log\mathcal{I} = d\log\mathcal{L} + 2\,d\log\mathcal{V} + o(\mathcal{I}), \tag{12}$$

$$d\log\tau_v = d\log\mathcal{L} + o(\mathcal{I}), \tag{13}$$

$$d\log\tau_u = -2\,(\,d\log\mathcal{L} + d\log\mathcal{V}) + o(\mathcal{I}). \tag{14}$$

These expressions allow for immediate implementation of the empirical strategy on readily available data. 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 Discussion

In this section we briefly discuss the key assumptions of our baseline model. First, we show that modeling the trading protocol as one in which traders submit limit- rather than market-orders is irrelevant for our results, as long as there is free-entry of uninformed speculators. In particular, Theorem 1 and Proposition 2 hold when traders submit market orders. We then consider risk-averse traders and show that the results still hold qualitatively: informativeness is determined by the same sufficient statistic, but now we need one more equation to determine the total amount of information in the market (which is determinant to the level of liquidity). We provide a characterization of these pair of equations, extending Theorem 1, and obtain the relevant market measures extending Proposition 2. We show that our central comparative statics also holds in this setting. Finally, we discuss the role of smooth against fixed costs.

Limit Orders. Consider the same setup as in our baseline model, with two modifications. First, the trading protocol is as in Kyle (1989): traders submit limit-orders, in the form of a demand schedule, and markets clear. Second, for any fixed level of informed traders, n, there is free-entry of uninformed traders. The second assumption guarantees that there is sufficient uninformed demand to arbitrage away all the public information. In particular, if one maintains the same definition of Y as the aggregate order-flow from informed and noise traders, free-entry of uninformed traders keeps the unbiasedness of prices:

$$p = \mathbb{E}[v|Y].$$

In this setup, we show that Theorem 1 and Proposition 2 hold exactly as stated. We do not formalize this claim—which is a direct consequence of Theorem 2. The key implication of this result is that the trading protocol becomes irrelevant in competitive markets: although some microstructure phenomena persist in large markets, others do not.

Risk-Aversion. Next, we maintain the assumption of limit-orders and free-entry of uninformed investors, and allow traders to be risk-averse. Traders are assumed to be CARA utility maximizers with risk-aversion γ. That is, the ex-post payoff for trader i from demanding d_i at price p is:

$$\frac{1 - e^{-\gamma(v-p)d_i}}{\gamma}.$$

We develop this model in Section B. Compared to a model with risk-neutral traders, the key difference is that the amount of the acquired information that gets revealed by prices now varies with fundamentals. When traders are risk-neutral, and as in Kyle (1985), 50% of traders' private information is leaked through prices under risk-neutrality. Under risk-aversion, this fraction is φ, which plays a central role in the equilibrium characterization. Formally, define:

Definition 6. *Price leakage φ is the fraction of all the information acquired which gets revealed by prices, i.e.,*

$$\varphi := \lim_{n \to \infty} \frac{Var[v|p, \{s_i\}_i]}{Var[v|p]}.$$

Armed with that definition we obtain the following result.

Theorem 2. *Assume $\gamma\chi < 1$. In a model with risk-averse traders, limit-orders and free-entry of uninformed traders, competitive informativeness $\mathcal{I}$ is the unique solution to:*

$$\frac{\mathcal{I}}{(1-\mathcal{I})^3} \frac{1-\varphi}{2\varphi^2} = \Gamma,$$

where $\varphi = \frac{1-\frac{\gamma\chi}{1-\mathcal{I}}}{2}$*. Moreover, liquidity satisfies:*

$$\mathcal{L} = \sqrt{\frac{\tau_v}{\tau_u(1-\varphi)}}\frac{1}{\sqrt{\mathcal{I}(1-\mathcal{I})}}.$$

Finally, $\mathcal{I}$ *is decreasing with* τ_u*.*

If $\gamma\chi \geq 1$*, there is no information acquisition in equilibrium.*

Theorem 2 shows that our results obtained in the simpler, risk-neutral model, are not knife-edge. First, it continues being true that prices become fully-revealing if and only if $\chi = 0$. In that case, the equations above look exactly as they do in the risk-neutral case. Second, when $\chi > 0$, there is an interior level of price informativeness. The total amount of the acquired information that leaks through prices, φ is now less than half of the total acquired precision. φ affects liquidity, as the total amount of adverse selection in the market changes the willingness of the competitive fringe of uninformed investors to trade against the order flow.

Finally, and central to our analysis, the result that informativeness increases with noise-trading volatility persists in this model. Interestingly, one can show that the price leakage φ decreases with that same volatility. In other words, the fraction of information acquired that is revealed goes down as noise grows, but the total amount of information goes up so sharply, that it more than compensates any decrease in φ.

Fixed information costs. For comparison, Appendix C studies a version of our baseline model with fixed information costs. In that case, it is well-understood that an equilibrium with a large quantity of aggregate information *can only emerge if the entire cost function vanishes* (Kyle, 1989; Vives, 2014). This stands in stark contrast to our result with smooth costs. The intuition for this result is that with fixed costs, the number of informed traders stabilizes as the market becomes large, and so the quantity of aggregate information stabilizes too. The remaining qualitative results hold: there is a quite similar characterization of price-informativeness and liquidity, and price-informativeness is again increasing in noise-trader volatility—i.e., decreasing in τ_u.

6 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 well-understood 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 volatility and volume, two easily observable variables, to back out the dynamics of fundamental risk and non-fundamental noise, hence the dynamics of informativeness.

References

Admati, A. R., and Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *The review of financial studies*, *1*(1), 3–40.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31–56.

Amihud, Y., and Noh, J. (2021). Illiquidity and stock returns ii: Cross-section and time-series effects. *The Review of Financial Studies*, *34*(4), 2101–2123.

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.

Bernhardt, D., and Taub, B. (2006). Kyle v. kyle ('85 v.'89). *Annals of Finance*, *2*(1), 23–38.

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.

Chordia, T., Sarkar, A., and Subrahmanyam, A. (2005). An empirical analysis of stock and bond market liquidity. *The Review of Financial Studies*, *18*(1), 85–129.

Collin-Dufresne, P., and Fos, V. (2015). Do prices reveal the presence of informed trading? *The Journal of Finance*, *70*(4), 1555–1582.

Collin-Dufresne, P., and Fos, V. (2016). Insider trading, stochastic liquidity, and equilibrium prices. *Econometrica*, *84*(4), 1441–1475.

Dávila, E., and Parlatore, C. (2018). *Identifying price informativeness* (tech. rep.). National Bureau of Economic Research.

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.

Farboodi, M., Matray, A., Veldkamp, L., and Venkateswaran, V. (2022). Where has all the data gone? *The Review of Financial Studies*, *35*(7), 3101–3138.

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, 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.

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2011). Stock market liquidity and the business cycle. *The Journal of Finance*, *66*(1), 139–176.

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 Proof of Lemma 1

The existence and uniqueness of a linear equilibrium is a direct consequence of Theorem 1 in Lambert et al. (2018).

Because throughout this proof $\mathbf{z}$ is fixed, we call, with a slight abuse of notation $\mathbf{p} = \mathbf{p}(\cdot, \mathbf{z})$, and, for each trader i, $\mathbf{d}_i = \mathbf{d}_i(\cdot, \mathbf{z})$—that is, we drop $\mathbf{z}$ from the arguments of there functions.

To characterize the linear equilibrium, notice that, under the equilibrium strategies, the pricing function from the point of view of trader i is:

$$\mathbf{p}(Y) = \lambda d_i + \lambda \sum_{j \neq i} \beta_j s_j + \lambda u = \lambda d_i + \lambda (n-1) v + \lambda \sum_{j \neq i} \varepsilon_j + \lambda u.$$

Taking that into account, the first-order-condition for trader i determines:

$$d_i = \frac{1}{2\lambda} \left(1 - \lambda \sum_{j \neq i} \beta_j \right) \mathbb{E}[v | s_i] = \frac{1}{2\lambda} \left(1 - \lambda \sum_{j \neq i} \beta_j \right) \frac{\tau_i}{\tau_v + \tau_i} s_i,$$

where the second equality follows from Bayesian updating given joint normality of (v, s_i) and $\varepsilon_i \perp v$. Recall that in a linear equilibrium, $d_i = \beta_i s_i$. Define $B = \sum_j \beta_j$, and $\kappa_i = \frac{\tau_i}{\tau_i + \tau_v}$. The equation above is then equivalent to:

$$\beta_i = \frac{1}{2\lambda} \left(1 - (B - \beta_i) \right) \kappa_i, \tag{A.1}$$

which allows us to solve for β_i as a function of λ and B:

$$\beta_i = \underbrace{\frac{\kappa_i}{2 - \kappa_i}}_{z_i} \frac{1 - \lambda B}{\lambda}.$$

By summing both sides up and solving for B:

$$B = \frac{\mathbf{z} \cdot \mathbf{1}}{1 + \mathbf{z} \cdot \mathbf{1}} \frac{1}{\lambda}, \tag{A.2}$$

and plugging into the expression of β_i:

$$\beta_i = \frac{z_i}{1 + \mathbf{z} \cdot \mathbf{1}} \frac{1}{\lambda}. \tag{A.3}$$

We now obtain an expression for λ by using the market-maker's breakeven condition. To that goal, define the unbiased signal from aggregate demand:

$$\tilde{Y} = \frac{Y}{B} = v + \underbrace{\frac{\sum_j \beta_j \varepsilon_j + u}{B}}_{\eta},$$

with $\eta \perp v$ normally distributed with precision

$$\tau_\eta := \text{Var}^{-1}[\eta] = \frac{B^2}{\sum_j \beta_j^2 \tau_j^{-1} + \tau_u^{-1}}.$$

Then, by standard Bayesian updating for joint normal variables:

$$\mathbb{E}[v|Y] = \mathbb{E}[v|\tilde{Y}] = \frac{\frac{B^2}{\sum_j \beta_j^2 \tau_j^{-1} + \tau_u^{-1}}}{\frac{B^2}{\sum_j \beta_j^2 \tau_j^{-1} + \tau_u^{-1}} + \tau_v} \tilde{Y} = \frac{B}{B^2 + \tau_v \left(\sum_j \beta_j^2 \tau_j^{-1} + \tau_u^{-1} \right)} Y,$$

where the first equality holds because Y and $\tilde{Y}$ contain the same information.

Because $\mathbb{E}[v|Y] = p = \lambda Y$, we obtain an expression for λ:

$$\lambda = \frac{B}{B^2 + \tau_v \left(\sum_j \beta_j^2 \tau_j^{-1} + \tau_u^{-1} \right)}.$$

Towards a closed-form solution for λ, we will obtain simplifications for parts of the expression above. First, using $\tau_i^{-1} = \frac{1}{2\tau_v} \frac{1-z_i}{z_i}$, and A.3:

$$\sum_j \beta_j^2 \tau_j^{-1} = \frac{1}{2} \frac{\mathbf{z} \cdot \mathbf{1} - \mathbf{z} \cdot \mathbf{z}}{(\mathbf{z} \cdot \mathbf{1})^2 \lambda^2} \tau_v^{-1}.$$

Plugging the equation above and equation A.2 in the expression for λ, one can solve for λ to obtain:

$$\frac{1}{\mathcal{L}} = \lambda = \sqrt{\frac{\tau_u}{2\tau_v}} \frac{\sqrt{\mathbf{z} \cdot \mathbf{1} + \mathbf{z} \cdot \mathbf{z}}}{1 + \mathbf{z} \cdot \mathbf{1}},$$

which established our result on liquidity, and $\beta_i = \sqrt{\frac{2\tau_v}{\tau_u}} \frac{z_i}{\sqrt{\mathbf{z} \cdot \mathbf{1} + \mathbf{z} \cdot \mathbf{z}}}$.

We now characterize price informativeness. Because prices contain the same information as Y, which contains the same information as $\tilde{Y}$, $\text{Var}[v|p] = \text{Var}[v|\tilde{Y}]$. By standard Bayesian updating for jointly normal variables:

$$\text{Var}^{-1}[v|\tilde{Y}] = \tau_v^{-1} + \tau_\eta^{-1} = \tau_v(1 + \mathbf{z} \cdot \mathbf{1}).$$

But then:

$$\mathcal{I} = 1 - \frac{\text{Var}[v|\tilde{Y}]}{\text{Var}[v]} = \frac{\mathbf{z}\cdot\mathbf{1}}{1+\mathbf{z}\cdot\mathbf{1}},$$

as we wanted to prove.

All that is left is to calculate the ex-ante utility. For that, note that $\mathbb{E}[p|s_i] = \lambda\mathbb{E}[Y|s_i] = \lambda\sum_{j\neq i}\beta_j\kappa_i s_i + \lambda\beta_i s_i$. Moreover, by equation A.1, $2\lambda\beta_i = (1-\lambda\sum_{j\neq i}\beta_j)\kappa_i$. Thus:

$$\mathbb{E}[\mathbf{d}_i(s_i)(v-\mathbf{p}(Y))|s_i] = \beta_i s_i^2\,(2\lambda\beta_i - \lambda\beta_i) = \lambda\beta_i^2 s_i^2.$$

Taking ex-ante utility and recalling that, ex-ante, s_i has mean zero, we have $\mathbb{E}[s_i^2] = \text{Var}[s_i] = \tau_v^{-1} + \tau_i^{-1}$, which delivers, after algebra:

$$V_i(z_i,\mathbf{z}) = \lambda\beta_i^2(\tau_v^{-1}+\tau_i^{-1}) = \sqrt{\frac{1}{2\tau_v\tau_u}}\frac{1}{1+\mathbf{z}\cdot\mathbf{1}}\frac{z_i+z_i^2}{\sqrt{\mathbf{z}\cdot\mathbf{1}+\mathbf{z}\cdot\mathbf{z}}},$$

which can be rewritten as in the statement of the proposition. This concludes the proof. ■

A.2 Symmetric equilibrium (Statement and Proof of Lemma A.1)

In this section we characterize the symmetric information equilibrium for any finite n, and consider their asymptotic limits.

Lemma A.1. *Let z^n denote a symmetric equilibrium information acquisition in the economy with n strategic traders, in which $p^n = \lambda^n Y$ and $d^n = \beta^n s_i$ represent prices and (symmetric) demand on the equilibrium path. Then,*

$$\lambda^n = \sqrt{\frac{\tau_u}{2\tau_v}}\sqrt{\frac{nz^n(1+z^n)}{(1+nz^n)^2}}$$

$$Y^n = \sqrt{\frac{2\tau_v}{\tau_u}}\sqrt{\frac{nz^n}{1+z^n}}\left(\frac{1}{n}\sum_{i=1}^n s_i\right) + \sqrt{\tau_u^{-1}}e_u$$

$$p^n = \frac{nz^n}{1+nz^n}\left(\frac{1}{n}\sum_{i=1}^n s_i\right) + \sqrt{\frac{\tau_v^{-1}}{2}}\frac{\sqrt{nz^n(1+z^n)}}{1+nz^n}e_u,$$

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

$$\frac{1}{n}\sum_{i=1}^{n} s_i \to v + \sqrt{\frac{\tau_v^{-1}}{2Z^*}}e_s,$$

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

$$\begin{aligned}
\lim_{n\to\infty} \lambda^n &= \sqrt{\frac{\tau_v^{-1}}{2\tau_u^{-1}}}\frac{\sqrt{Z^*}}{1+Z^*} \\
\lim_{n\to\infty} Y^n &= \sqrt{2\tau_u^{-1}}\left[\sqrt{Z^*}\frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s+e_u}{\sqrt{2}}\right] \\
\lim_{n\to\infty} p^n &= \sqrt{\tau_v^{-1}}\frac{\sqrt{Z^*}}{1+Z^*}\left[\sqrt{Z^*}\frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s+e_u}{\sqrt{2}}\right].
\end{aligned}$$

In Appendix C.1, we also prove an analogous version of Lemma A.1 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.

A.2.1 Proof of Lemma A.1

In a symmetric equilibrium, $\mathbf{z} = z^n\mathbf{1}$, and so $\mathbf{z}\cdot\mathbf{1} = nz^n$ and $\mathbf{z}\cdot\mathbf{z} = n(z^n)^2$. Substituting these results into the expressions obtained in the proof of Lemma 1 for λ^n and β^n

$$\begin{aligned}
\lambda^n &= \sqrt{\frac{\tau_u}{2\tau_v}}\sqrt{\frac{nz^n(1+z^n)}{(1+nz^n)^2}}, \\
\beta^n &= \frac{1}{n}\sqrt{\frac{2\tau_v}{\tau_u}}\sqrt{\frac{nz^n}{1+z^n}}.
\end{aligned}$$

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

$$\begin{aligned}
Y^n &= \sqrt{\frac{2\tau_v}{\tau_u}}\sqrt{\frac{nz^n}{1+z^n}}\frac{\sum_i s_i}{n} + \sqrt{\tau_u^{-1}}e_u, \\
p^n &= \frac{nz^n}{1+nz^n}\frac{\sum_i s_i}{n} + \sqrt{\frac{\tau_v^{-1}}{2}}\frac{\sqrt{nz^n(1+z^n)}}{1+nz^n}e_u.
\end{aligned}$$

We will now take the limit $n \to \infty$. First, note the fact that $s = v + \varepsilon_i$ in a symmetric equilibrium, where

$$\varepsilon \sim \mathcal{N}\Big(0, \frac{1}{2}\tau_v^{-1}\frac{1-z^n}{z^n}\Big).$$

Therefore,

$$\frac{\sum_i \varepsilon_i}{n} \sim \text{Normal}\Big(0, \frac{1}{2}\tau_v^{-1}\frac{1-z^n}{nz^n}\Big)$$

and so using the definition $Z^* := \lim_{n\to\infty} nz^n$ and the result that $z^n \to 0$ (Proposition 1), we have by the Central Limit Theorem,

$$\frac{\sum_i s_i}{n} \to v + \sqrt{\frac{\tau_v^{-1}}{2Z^*}}e_s$$

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

$$Y_n \to \sqrt{2\tau_u^{-1}}\left(\sqrt{Z^*}\frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s + e_u}{\sqrt{2}}\right)$$

$$p^n \to \sqrt{\tau_v^{-1}}\frac{\sqrt{Z^*}}{1+Z^*}\left(\sqrt{Z^*}\frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s + e_u}{\sqrt{2}}\right)$$

This proves Lemma A.1. ■

A.3 Vanishing individual information (Proof of Proposition 1)

First, consider the case $\chi = \infty$. Given strict convexity of c and the fact that $\frac{dV_i(z\ \mathbf{1})}{dz_i} < \infty$ 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 $\lim_{t\to\infty} z^{n_t} = z^* > 0$. In that case, we have $\lim_{t\to\infty} n_t z^{n_t} = +\infty$. Thus, we have $\lim_{t\to\infty} \frac{dV_i(z^{n_t}\ \mathbf{1})}{dz_i} = 0$. 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 $0 = \lim_{t\to\infty} c'(z^{n_t}) = c'(z^*)$. But strict convexity also implies $c'(z) > 0$ for all $z > 0$, contradicting $z^* > 0$. This argument works for any sub-sequence, and so $\lim_{n\to\infty} z^n$ exists and is equal to zero. ■

A.4 Aggregate information characterization (Proof of Theorem 1)

The case of $\chi = \infty$ (and therefore $\Gamma = 0$) is already proved in Proposition 1, which showed that $z^n = 0$ for each n.

Next, we prove the case $\chi = 0$ (and therefore $\Gamma = \infty$). By Proposition 1, we know that z^n is a positive sequence that converges to zero. Assume a bounded subsequence of z^n exists. Consider a subsequence such that $n_t z^{n_t} \to Z < \infty$—the existence of which is a consequence of the existence of a bounded subsequence. Then, we have $\lim_{t\to\infty} \frac{dV_i(z^{n_t}\,\mathbf{1})}{dz_i} > 0$. On the other hand, the fact that $z^{n_t} \to 0$ implies by the monotone convergence theorem that $c'(z^{n_t}) \to \chi = 0$. But, along that subsequence, we have that (5) holds with equality, which contradicts the fact that $\frac{dV_i(z^{n_t}\,\mathbf{1})}{dz_i}$ and $c'(z^{n_t})$ have different limit points. Therefore, each subsequence of z_n must be unbounded, and $nz^n \to \infty$, which proves that $\mathcal{I} \to 1$.

Finally, we prove the claim for $\chi \in (0, \infty)$—and so $\Gamma \in (0, \infty)$. By Proposition 1, we have $c'(z^n) \to \chi > 0$, which implies $\frac{dV_i(z^{n_t}\,\mathbf{1})}{dz_i} \to \chi$ for each subsequence $(n_t)_{t\geq 0}$, by equation (5). If $z^{n_t} \to 0$ and $n_t z^{n_t} \to \infty$, $\frac{dV_i(z^{n_t}\,\mathbf{1})}{dz_i} \to 0$, which contradicts the equality, which means that each subsequence is bounded, and therefore the sequence $n\, z^n$ is itself bounded. Take any converging subsequence $(n_t)_{t\geq 0}$ and define $Z := \lim_{t\to\infty} n_t z^{n_t}$. We then obtain:

$$\begin{aligned}
\chi &= \lim_{t\to\infty} \frac{dV_i(z^{n_t}\,\mathbf{1})}{dz_i} \\
&= \lim_{t\to\infty} \sqrt{\frac{1}{2\tau_v\tau_u}} \frac{1}{1+n_t z^{n_t}} \frac{1}{\sqrt{n_t z^{n_t} + n_t (z^{n_t})^2}} \\
&= \sqrt{\frac{1}{2\tau_v\tau_u}} \frac{1}{1+Z} \frac{1}{\sqrt{Z}},
\end{aligned}$$

which can be reorganized to obtain: $Z(1+Z)^2 = \Gamma$. Because there is a unique positive root Z, for this equation, call it Z^*, we conclude that all converging subsequences converge to Z^*. Because nz^n was shown to be bounded, in the first place, then $nz^n \to Z^*$. To conclude, it suffices to show $Z^* = \frac{\mathcal{I}}{1-\mathcal{I}}$, so that we obtain the equality in the statement of the theorem.

To that end, recall from A.1 that limiting prices satisfy:

$$p := \lim_{n\to\infty} p^n = \sqrt{\tau_v^{-1}} \frac{\sqrt{Z^*}}{1+Z^*} \left[\sqrt{Z^*} \frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s + e_u}{\sqrt{2}} \right].$$

Define $\tilde{p}$ as the signal extracted from the price that is unbiased for v. That is, $\tilde{p} =$

$p\frac{1+Z^*}{Z^*} = v + \zeta$, where $\zeta \perp v$ is normally distributed with precision: $\tau_\zeta = Z^* \tau_v$. Therefore,

$$\text{Var}[v|\tilde{p}] = \text{Var}[v|p] = \frac{1}{Z^*\tau_v + \tau_v},$$

and therefore

$$\mathcal{I} = 1 - \frac{\text{Var}[v|p]}{\text{Var}[v]} = \frac{Z^*}{1+Z^*}.$$

Reorganizing this equality, we obtain $Z^* = \frac{\mathcal{I}}{1-\mathcal{I}}$, proving the theorem. ■

A.5 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 1). 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:*

$$n^{3/(3+2\gamma)} z^n \to \kappa^{-2/(3+2\gamma)} \left(\frac{\tau_v^{-1} \tau_u^{-1}}{2} \right)^{1/(3+2\gamma)}.$$

Proof. By Proposition 1, 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

$$\frac{f_n(n^{-\zeta}Z)}{n^{-\gamma\zeta}Z^\gamma}$$
$$= \frac{\sqrt{\frac{\tau_v^{-1}\tau_u^{-1}}{2}}(n^{\gamma\zeta}Z^{-\gamma} + 2n^{-(1-\gamma)\zeta}Z^{1-\gamma})}{(1+n^{1-\zeta}Z)\sqrt{n^{1-\zeta}Z(1+n^{-\zeta}Z)}}\Big[1 - \frac{n^{-\zeta}Z(1+n^{-\zeta}Z)}{1+2n^{-\zeta}Z}\Big(\frac{1}{1+n^{1-\zeta}Z} + \frac{\frac{1}{2}(1+2n^{-\zeta}Z)}{n^{1-\zeta}Z(1+n^{-\zeta}Z)}\Big)\Big].$$

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 $Z = [\kappa^{-2}(\tau_v^{-1}\tau_u^{-1}/2)]^{1/(3+2\gamma)}$. [Algebra: we need

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

which delivers the result. □

Proposition A.1 tells us how fast aggregate information nz^n explodes as the market grows. The answer: $nz^n \sim (\frac{\tau_v^{-1}\tau_u^{-1}}{2\kappa^2})^{\frac{1}{3+2\gamma}} n^{\frac{2\gamma}{3+2\gamma}}$ 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 $\sqrt{\tau_u^{-1}}$ 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.6 Identification (Proof of Proposition 4)

We start with the observable metrics for $\mathcal{V}$ and $\mathcal{L}$ from Proposition 2, and the equation for informativeness in Theorem 1. Taking logs on both side of these equations, recalling

χ is assumed to be fixed and using total differentiation we obtain:

$$\begin{aligned}
d\log\mathcal{V} &= \frac{1}{2}\left(d\log\mathcal{I} - d\log\tau_v\right),\\
d\log\mathcal{L} &= \frac{1}{2}\left(d\log\tau_v - d\log\tau_u - \log\mathcal{I} - \log\left(1-\mathcal{I}\right)\right),\\
d\log\mathcal{I} - 3d\log\left(1-\mathcal{I}\right) &= -d\log\tau_u - d\log\tau_v.
\end{aligned}$$

By the chain rule, we rewrite $d\log(1-\mathcal{I}) = -\frac{\mathcal{I}}{1-\mathcal{I}}d\log\mathcal{I}$. Substituting that in the equations above, and solving for $d\log\tau_v, d\log\tau_u, d\log\mathcal{I}$ as a function of $d\log\mathcal{V}$ and $d\log\mathcal{L}$, we conclude:

$$d\log\mathcal{I} = \frac{1-\mathcal{I}}{1+\mathcal{I}}\left(d\log\mathcal{L} + 2d\log\mathcal{V}\right), \tag{A.4}$$

$$d\log\tau_v = \frac{1-\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{L} - \frac{4\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{V}, \tag{A.5}$$

$$d\log\tau_u = -\frac{2+\mathcal{I}}{1+\mathcal{I}}d\log\mathcal{L} - \frac{2}{1+\mathcal{I}}d\log\mathcal{V}. \tag{A.6}$$

Given an initial value for $\mathcal{I}_0$, and initial observables $\mathcal{L}_0, \mathcal{V}_0$, we can use the characterization of equilibrium to calculate $\tau_{v,0}$ and $\tau_{u,0}$. So given $\mathcal{I}_0$ we can obtain the complete dynamics of the system. ■

B A model of Limit-Orders and Risk-Aversion

We maintain the assumption that there is one risky asset with payoff $v \sim \mathcal{N}(0, \tau_v^{-1})$, noise trader demand is $u \sim \mathcal{N}(0, \tau_u^{-1})$, and traders observe signal $s_i = v + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, \tau_i^{-1})$. Moreover, $u \perp v \perp \varepsilon_i$ for all i, and $\varepsilon_i \perp \varepsilon_j$, for all i, j. Trader i's expected payoff of buying d units of the asset at price p is:

$$u_i(x) = \frac{1 - \mathbb{E}\left[e^{-\gamma(v-p)d}\right]}{\gamma}.$$

Besides the payoff difference, there are two other changes relative to the baseline model. First, the trader submits a demand schedule, specifying for each possible price p the amount $d_i(p)$ of the asset he demands, taking into account market-clearing. Second, there is a competitive fringe—also understood as free-entry—of uninformed traders. This assumption is often used in the literature (Bernhardt and Taub, 2006; Vives, 2014), and it leads to a simplification of the model: market clearing is substituted by the unbiasedness of prices. Formally, given aggregate demand $Y = \sum_i d_i + u$, for all informed traders i:

$$p = \mathbb{E}[v|Y].$$

B.1 Proof of Theorem 2

We start by assuming a linear equilibrium for a fixed number of informed entrants, n. That is, in this equilibrium each trader submits demand

$$d_i = B_i s_i + C_i Y \tag{B.1}$$

And prices satisfy

$$p = \lambda Y \tag{B.2}$$

We start by noticing that the definition of Y and the aggregation of individual demand leads to:

$$Y = \sum_j B_j s_j + \underbrace{C}_{\equiv \sum_j C_j} Y + u, \iff Y = \frac{1}{1-C}\left(\sum_j B_j s_j + u\right).$$

Notice that in this model $\dfrac{\partial p}{\partial d_i} = \lambda \frac{1}{1-\sum_{j\neq i} C_j}$, for any i. By the mean-variance representation of the individual problem, we must have:

$$d_i = \frac{\mathbb{E}[v|p,s_i] - p}{\gamma \mathrm{Var}[v|p,s_i] + \lambda \frac{1}{1-\sum_{j\neq i} C_j}} \tag{B.3}$$

We will now obtain the Gaussian filtering. For that, define, for any player i, the auxiliary variable W_i which orthogonalizes the aggregate order flow and their signal:

$$W_i = \frac{(1-C)Y - B_i s_i}{\underbrace{B - B_i}_{\equiv \sum_{j\neq i} B_j}} = v + \underbrace{\frac{\sum_j B_j \varepsilon_j + u}{B - B_i}}_{\equiv \eta_i}. \tag{B.4}$$

We then have that $\eta_i \perp (v, \varepsilon_i)$, and

$$\tau_{\eta_i} \equiv \mathrm{Var}^{-1}[\eta_i] = \frac{(B - B_i)^2}{\sum_{j\neq i} B_j^2 \tau_j^{-1} + \tau_u^{-1}}.$$

This reformulation allows us to think of the posterior precision of v as linear in the precisions of the orthogonal signals. That is:

$$\kappa_i \equiv \text{Var}^{-1}[v|p,s_i] = \tau_v + \tau_i + \tau_{\eta_i}, \tag{B.5}$$

and

$$\mathbb{E}[v|p,s_i] = \frac{\tau_i}{\kappa_i}s_i + \frac{\tau_{\eta_i}}{\kappa_i}W_i = \frac{\tau_i\,(B-B_i) - \tau_{\eta_i}B_i}{\kappa_i(B-B_i)}s_i + \frac{\tau_{\eta_i}(1-C)}{\kappa_i(B-B_i)}Y. \tag{B.6}$$

Then, we can plug B.5 and B.6 in the mean-variance expression for the demand of agent i, B.3, and equate that to the linear assumption in B.1. Matching coefficients in those expressions we get:

$$B_i\left(\lambda\frac{1}{1-\sum_{j\neq i}C_j}\kappa_i + \gamma\right) + \frac{B_i}{B-B_i}\tau_{\eta_i} = \tau_i, \tag{B.7}$$

and

$$C_i\left(\lambda\frac{1}{1-\sum_{j\neq i}C_j}\kappa_i + \gamma\right) = \frac{1-C}{B-B_i}\tau_{\eta_i} - \lambda\kappa_i \tag{B.8}$$

Note that because τ_{η_i} depends only on B's and κ_i depends only on τ_{η_i}, the system of equations in B.7 solves for $B's$ as a function of λ. With that information, one can then solve B.8 as a function of λ. So we just need to close the equilibrium by learning λ. That is obtained by recalling that $p = \mathbb{E}[v|Y]$.

We start by making Y unbiased for v. That is, define:

$$\tilde{Y} \equiv \frac{1-C}{B}Y = v + \underbrace{\frac{\sum_j B_j\varepsilon_j + u}{B}}_{\eta}.$$

We have that $\tau_\eta = \dfrac{B^2}{\underbrace{\sum_j B_j^2\tau_j^{-1} + \tau_u^{-1}}_{\equiv D}}$.

Thus:

$$\lambda Y \equiv p = \frac{B^2/D}{\tau_v + \frac{B^2}{D}}\frac{(1-C)}{B}Y \iff \lambda = \frac{B(1-C)}{\tau_v D + B^2}. \tag{B.9}$$

Finally, we can specialize to the case of interest, in which trader 1 has precision τ_1 and all other traders have precision τ^* and are, therefore, symmetric. In that case, the equilibrium equations become:

$$B_1\left(\lambda\frac{1}{1-(n-1)C^*}\kappa_1+\gamma\right)+\frac{B_1}{(n-1)B^*}\tau_{\eta_1}=\tau_1, \tag{B.10}$$

$$B^*\left(\lambda\frac{1}{1-(n-2)C^*-C_1}\kappa^*+\gamma\right)+\frac{B^*}{(n-2)B^*+B_1}\tau_{\eta^*}=\tau^*, \tag{B.11}$$

With $\kappa_1=\tau_v+\tau_1+\tau_{\eta_1}$, $\kappa^*=\tau_v+\tau^*+\tau_{\eta^*}$, and:

$$\tau_{\eta_1}=\frac{[(n-1)B^*]^2}{\frac{[(n-1)B^*]^2}{(n-1)\tau^*}+\frac{1}{\tau_u}}, \qquad \tau_{\eta^*}=\frac{\left(\frac{n-2}{n-1}(n-1)B^*+B_1\right)^2}{\frac{B_1^2}{\tau_1}+\frac{n-2}{n-1}\frac{[(n-1)B^*]^2}{(n-1)\tau^*}+\frac{1}{\tau_u}} \tag{B.12}$$

and $\lambda=\frac{B(1-C)}{\tau_v D+B^2}$, with $D=\frac{B_1^2}{\tau_1}+\frac{[(n-1)B^*]^2}{(n-1)\tau^*}+\frac{1}{\tau_u}$. We also have equations for C, but those will play a minor role in what follows.

We now calculate the indirect payoff of trader 1 in a linear equilibrium.

Before submitting a demand schedule, the trader observes his signal s_i. Therefore, his demand optimization must solve:

$$\begin{aligned} d_i(p) &\in \arg\min_x \mathbb{E}_v\left[e^{-\gamma(v-p)x}|(p,s_i)\right] \\ &= \arg\max_x \mathbb{E}[(v-p)x|p,s_i]-\frac{\gamma}{2}x^2\mathrm{Var}[v|p,s_i] \end{aligned}$$

The first-order-condition to this problem satisfies:

$$d_i(p)=\frac{\mathbb{E}[v-p|p,s_i]}{\gamma\mathrm{Var}[v|p,s_i]+\frac{dp}{dd_i}}.$$

In a linear equilibrium, it will be the case that $\frac{dp}{dd_i}=\gamma\lambda_i$ for some constant. Therefore, the ex-ante indirect utility of trader i is:

$$\begin{aligned} \gamma V_i := 1-\mathbb{E}\left[\mathbb{E}_v[e^{-\gamma(v-p)d_i(p)}|p,s_i]\right] &= 1-\mathbb{E}\left[e^{-\gamma\left(\mathbb{E}[(v-p)d_i(p)|p,s_i]-\frac{\gamma}{2}x_i^2(p)\mathrm{Var}[v|p,s_i]\right)}\right] \\ &= 1-\mathbb{E}\left[e^{-\gamma^2\frac{1}{2}(\mathrm{Var}[v|p,s_i]+2\lambda_i)d_i(p)^2}\right], \end{aligned}$$

where the last observation follows from the first-order-condition for $d_i(p)$. Under a linear equilibrium, $d_i(p)$ will be an ex-ante mean zero gaussian random variable. Because $\mathrm{Var}[v|p,s_i]$ is a constant across realizations of p and s_i, the term in the exponent is a constant multiplying a normal random variable. Therefore, up to normalizing $d_i(p)$, the expectation is the m.g.f. of a chi-squared distribution, and therefore:

$$\gamma V_i = 1 - \left(1 + \gamma^2 \left(\mathrm{Var}[v|p,s_i] + 2\,\lambda_i\right) \mathrm{Var}[d_i(p)]\right)^{-1/2}.$$

Finally, recall that

$$d_i(p) = \frac{\mathbb{E}[v-p|p,s_i]}{\gamma(\mathrm{Var}[v|p,s_i]+\lambda_i)},$$

and therefore

$$\mathrm{Var}[d_i(p)] = \frac{\mathrm{Var}[\mathbb{E}[v-p|p,s_i]]}{\gamma^2(\mathrm{Var}[v|p,s_i]+\lambda_i)^2} = \frac{\mathrm{Var}[v-p] - \mathrm{Var}\,[v|p,s_i]}{\gamma^2(\mathrm{Var}[v|p,s_i]+\lambda_i)^2},$$

where the second equality uses the law of total variance and the fact that $\mathrm{Var}[v|p,s_i]$ is constant in realizations of (p,s_i). We thus conclude that:

$$V_i = 1 - \left(\frac{\mathrm{Var}[v-p] + \frac{\mathrm{Var}[v-p]+\lambda_i}{\mathrm{Var}[v|p,s_i]+\lambda_i}\lambda_i}{\mathrm{Var}[v|p,s_i]+\lambda_i}\right)^{-1/2}.$$

Noting that, $\lambda_i = \frac{1}{\gamma}\lambda\frac{1}{1-\sum_{j\neq i} C_j}$:

$$V_i = \frac{1}{\gamma}\left(1 - \left(\frac{\mathrm{Var}[v-p] + \frac{\mathrm{Var}[v-p]+\lambda\frac{1}{\gamma}\frac{1}{1-(n-1)C^*}}{\mathrm{Var}[v|p,s_i]+\lambda\frac{1}{\gamma}\frac{1}{1-(n-1)C^*}}\lambda\frac{1}{\gamma}\frac{1}{1-(n-1)C^*}}{\mathrm{Var}[v|p,s_i]+\lambda\frac{1}{\gamma}\frac{1}{1-(n-1)C^*}}\right)^{-1/2}\right).$$

We now use our price equation in the semi-symmetric equilibrium:

$$p - v = \lambda Y - v = \frac{\lambda}{1-C}\left[Bv + \sum_j B_j\varepsilon_j + u\right] - v = \frac{(\lambda B - (1-C))}{1-C}v + \frac{\lambda}{1-C}B\eta$$

Thus,

$$\mathrm{Var}[p-v] = \left[\frac{(\lambda B-(1-C))}{1-C}\right]^2 \tau_v^{-1} + \left[\frac{\lambda}{1-C}B\right]^2 \frac{D}{B^2}.$$

By substituting the formula for λ from (B.9) above, we obtain:

$$\kappa_p \equiv \text{Var}^{-1}[v-p] = \tau_v + \frac{B^2}{D}.$$

Moreover, we know $\text{Var}[v|p,s_1] = \kappa_1^{-1}$. By defining $R \equiv \lambda \frac{1}{\gamma}\frac{1}{1-(n-1)C^*}$, we obtain:

$$V_1(\tau_1, \tau^* \mathbf{1}) = \frac{1}{\gamma}\left(1 - \left(\frac{\kappa_p + \kappa_1\kappa_p R}{\kappa_1 + \frac{\kappa_1+\kappa_1\kappa_p R}{\kappa_p+\kappa_1\kappa_p R}\kappa_1\kappa_p R}\right)^{1/2}\right)$$

By unbiasedness of prices, it must be the case that $V_1 = 0$ when $\tau_1 = 0$. We want to find the derivative of V_1 with respect to τ_1, for any fixed level of aggregate information $T \equiv (n-1)\tau^*$, when n is large. That derivative is the highest constant ζ_T such that:

$$\underbrace{\frac{1}{2\gamma}\left(1-(1-\gamma V_1(\tau_1,\tau^*\mathbf{1}))^2\right)}_{\approx V_1'(0,\tau^*\mathbf{1})\tau_1 \text{ for small } \tau_1} = \frac{1}{2\gamma}\frac{\kappa_1-\kappa_p}{\kappa_1}\frac{1+\frac{\kappa_1 R}{1+\kappa_1 R}}{1+\kappa_1 R\frac{1+\kappa_p R}{1+\kappa_1 R}} \geq \zeta_T\tau_1 + O(\tau_1^2),$$

Notice that $\frac{B^2}{D} = \tau_\eta$ is the informativeness of the aggregate order flow. So κ_p is the precision of v conditional on the price.

Notice that:

$$\kappa_1 - \kappa_p = \tau_1 + (\tau_{\eta_i} - \tau_\eta) = \tau_1 + \left(\frac{[(n-1)B^*]^2}{\frac{[(n-1)B^*]^2}{T}+\frac{1}{\tau_u}} - \frac{((n-1)B^*+B_1)^2}{\frac{B_1^2}{\tau_1}+\frac{[(n-1)B^*]^2}{T}+\frac{1}{\tau_u}}\right) \tag{B.13}$$

By equations B.11 and B.10, we can obtain upper bounds for B^* and B_1. Indeed:

$$(n-1)B^* \leq \frac{T}{\gamma},$$

and

$$B_1 \leq \frac{\tau_1}{\gamma}.$$

Define the variables $\bar{B} = (n-1)B^*$ and $\bar{C} = (n-1)C^*$.

Recall that $\tau_\eta = \frac{(\bar{B}+B_1)^2}{\frac{B_1^2}{\tau_1}+\frac{\bar{B}^2}{T}+\frac{1}{\tau_u}}$. The inequality above shows that, for small τ_1, $\tau_\eta \approx \frac{\bar{B}^2}{\frac{\bar{B}^2}{T}+\frac{1}{\tau_u}}$. Notice, at the same time, that for small τ_1, $\tau_{\eta_1} \approx \tau_\eta$.

Moreover, for large enough n, one can easily see that $\tau_{\eta^*} \approx \tau_\eta$. We now use B.8 to prove $C_1 \approx 0$ and, $\bar{C} \approx C$ when n is large.

$$C_1 \left(\frac{B(1-C)}{\tau_v D + B^2} \frac{1}{1-\bar{C}} \kappa_1 + \gamma \right) = \frac{1-C}{\bar{B}} \tau_{\eta_i} - \frac{B(1-C)}{\tau_v D + B^2} \kappa_1 \tag{B.14}$$

$$C^* \left(\frac{B(1-C)}{\tau_v D + B^2} \frac{1}{1 - \frac{(n-2)}{(n-1)} \bar{C} - C_1} \kappa^* + \gamma \right) = \frac{1-C}{\frac{n-2}{n-1} \bar{B} + B_1} \tau_{\eta^*} - \frac{B(1-C)}{\tau_v D + B^2} \kappa^* \tag{B.15}$$

Because when τ_1 is small, $\tau_\eta \approx \frac{\bar{B}^2}{\frac{\bar{B}^2}{T} + \frac{1}{\tau_u}} = \frac{\bar{B}^2}{D}$, we have $\tau_{\eta_i} \approx \tau_\eta$, and $\kappa_i \approx \tau_v + \frac{B^2}{D} = \frac{\tau_v D + B^2}{D}$.

Plugging these in B.14

$$C_1 \left(\frac{B(1-C)}{\tau_v D + B^2} \frac{1}{1-\bar{C}} \kappa_1 + \gamma \right) \approx \frac{(1-C)\bar{B}}{D} - \frac{\bar{B}(1-C)}{D} = 0,$$

and therefore $C_1 \approx 0$, and $\bar{C} \approx C$. Thus, when τ_1 is small and n is large, we have from B.11:

$$\bar{B} \left(\lambda \frac{1}{1-C} \kappa_p + \gamma + \frac{1}{\bar{B}} \tau_\eta \right) \approx T,$$

and from B.10:

$$B_1 \left(\lambda \frac{1}{1-C} \kappa_p + \gamma + \frac{1}{\bar{B}} \tau_\eta \right) \approx \tau_1.$$

Therefore, B_1 is approximately linear with τ_1 for large n and low τ_1. Formally:

$$B_1 \approx \frac{\bar{B}}{T} \tau_1.$$

We can then use B.16 to write:

$$\kappa_1 - \kappa_p \approx \tau_1 + \left(\frac{\bar{B}^2}{\frac{\bar{B}^2}{T} + \frac{1}{\tau_u}} - \frac{(\bar{B} + \frac{\bar{B}}{T}\tau_1)^2}{\frac{\bar{B}^2}{T^2}\tau_1 + \frac{\bar{B}^2}{T} + \frac{1}{\tau_u}} \right) \tag{B.16}$$

By rewriting this equation, we obtain, for small τ_1:

$$\kappa_1 - \kappa_p \approx \left(1 + \frac{\tau_\eta}{T} \left(\frac{\tau_\eta}{T} - 2 \right) \right) \tau_1 = \left(1 - \frac{\tau_\eta}{T} \right)^2 \tau_1.$$

Moreover, for fixed T, large n and τ_1 close to zero we know that $\kappa_1 = \kappa_p = \tau_v + \frac{\bar{B}^2}{D}$.

We also know

$$R\kappa_p = \frac{\lambda}{\gamma}\frac{1}{1-C}\left(\tau_v + \frac{\bar{B}^2}{D}\right) = \frac{1}{\gamma}\frac{\bar{B}}{\tau_v D + \bar{B}^2}\left(\frac{\tau_v D + \bar{B}^2}{D}\right) = \frac{1}{\gamma}\frac{\bar{B}}{D}.$$

Therefore, the derivative of trader 1's value at $\tau_1 = 0$ is:

$$V_1'(0, \tau^*\mathbf{1}) = \frac{1}{2\gamma}\frac{1}{\tau_v + \tau_\eta}\frac{1 + \frac{2}{\gamma}\frac{\tau_\eta}{B}}{1 + \frac{2}{\gamma}\frac{\tau_\eta}{B} + [\frac{1}{\gamma}\frac{\tau_\eta}{B}]^2}\left(1 - \frac{\tau_\eta}{T}\right)^2, \quad \text{where} \quad \tau_\eta = \frac{B^2}{\frac{B^2}{T} + \frac{1}{\tau_u}}.$$

Equilibrium is defined by a pair of equations: $V_1'(0, \tau^*\mathbf{1}) = \frac{\chi}{2\tau_v}$, and, by summing all the matching coefficient equations for B:[9]

$$T \approx \bar{B}\left(\lambda\frac{1}{1-C}\kappa_p + \gamma + \frac{1}{\bar{B}}\tau_\eta\right) = B\left(2\frac{B}{D} + \gamma\right).$$

Instead of solving for B, which is the aggregate demand-signal sensitivity, and T, which is the total amount of precision, we use the change-of-variables

$$Z := \frac{\frac{B^2}{\tau_v}}{\frac{B^2}{T} + \frac{1}{\tau_u}} \quad \text{and} \quad \varphi := \frac{\frac{B^2}{T}}{\frac{B^2}{T} + \frac{1}{\tau_u}}.$$

It is useful to observe that Z is $\frac{\tau_\eta}{\tau_v}$, that is, Z is the precision of the signal in the aggregate demand—and thus in the price—about v, normalized by the pecision of v. We then get $\mathcal{I} = \frac{Z}{1+Z}$. Moreover, $\varphi = \frac{\tau_\eta}{T}$ is the percentage of the aggregate market precision that leaks through prices.

With these variables, we have the relationships $T = \tau_v\frac{Z}{\varphi}$ and $B^2 = \frac{Z}{1-\varphi}\frac{\tau_v}{\tau_u}$. Plugging this into the cubic equation for B, we recover

$$1 - 2\varphi = \frac{\gamma\varphi}{\sqrt{1-\varphi}}\frac{1}{\sqrt{Z\tau_v\tau_u}} \tag{B.17}$$

$$(1-2\varphi)\frac{\sqrt{1-\varphi}}{\sqrt{\varphi}} = \frac{\gamma}{\sqrt{T\tau_u}}$$

[9] The adjustment for χ comes from observing that V_1 was written as a function of τ_1 and not z_1. Noticing that $\frac{dz_1}{d\tau_1}(0) = \frac{1}{2\tau v}$ gives us the adjustment.

Plugging the change-of-variables into the information FOC $U_1'(0) = c'(0) = \frac{\chi}{2\tau_v}$, we get

$$\frac{1}{2\gamma\tau_v}\frac{1-2\varphi}{1+Z} = \frac{\chi}{2\tau_v}$$

or written differently

$$\frac{1}{2\tau_v}\frac{1}{\sqrt{Z}(1+Z)}\frac{\varphi}{\sqrt{1-\varphi}}\frac{1}{\sqrt{\tau_v\tau_u}} = \frac{\chi}{2\tau_v}$$

We want to keep this form in hand to develop intuition later (because the left-hand-side is exactly $U_1'(0)$. However, we also want to simplify this equation to obtain a cleaner expression. After some algebra, we obtain the expression

$$Z(1+Z)^2 = \Gamma\frac{2\varphi^2}{1-\varphi}, \quad \text{where} \quad \Gamma := \frac{1}{2\chi^2\tau_v\tau_u} \tag{B.18}$$

From these two equations, we can easily take the limit $\gamma \to 0$ to get

$$\lim_{\gamma\to 0}\varphi = \frac{1}{2}$$
$$(\lim_{\gamma\to 0} Z)(1+\lim_{\gamma\to 0} Z)^2 = \Gamma$$

These equations are identical to the market order model. But in general since $\varphi \leq \frac{1}{2}$, strictly so when $\gamma > 0$, equation (B.18) shows that $Z(1+Z)^2 \leq \frac{1}{2}\Gamma$, meaning that Z is always smaller in this model than the baseline model. We can also manipulate equations (B.17) and (B.18) to rewrite them as

$$\varphi = \frac{1-\frac{\gamma\chi}{1-\mathcal{I}}}{2}$$
$$\frac{\mathcal{I}}{1-\mathcal{I}}\frac{1-\varphi}{2\varphi^2} = \Gamma$$

From the first equation, we clearly see a requirement for equilibrium is $\gamma\chi \leq 1$. Plugging the first into the second, we need only to solve a single nonlinear equation for Z.

$$Z(1+Z)^2 = \frac{1}{2}\Gamma\frac{(1-\gamma\chi(1+Z))^2}{1+\gamma\chi(1+Z)} \tag{B.19}$$

Because the left-hand-side is increasing in Z and the right-hand-side is decreasing in Z, this equation has at most one solution. Setting $Z = 0$, the left-hand-side is lower than the right-hand-side (or equal if $\Gamma = 0$). Setting $Z = \frac{1}{\gamma\chi} - 1$, we have that the left-hand-side is positive larger. Thus a solution exists as long as $\gamma\chi < 1$.

To show the comparative statics, rewrite the expression as:

$$Z(1+Z)^2 \frac{1+\gamma\chi(1+Z)}{(1-\gamma\chi(1+Z))^2} = \frac{1}{2}\Gamma,$$

then the right-hand-side is decreasing in τ_u and the left-hand-side is increasing in Z, thus Z decreases in τ_u, and so does $\mathcal{I}$.

Finally, the expression from liquidity follows from $\frac{dp}{du} = \frac{\lambda}{1-C}$. ■

C A model with fixed costs 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 (8), is

$$\Gamma := \frac{1}{2\chi\tau_u\tau_v}. \tag{C.1}$$

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:

$$V_n(z;z^n) = \sqrt{\frac{\tau_v^{-1}\tau_u^{-1}}{2}} \frac{z(1+z)}{(1+(n-1)z^n+z)\sqrt{(n-1)z^n(1+z^n)+z(1+z)}}.$$

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

$$V_n(1;1) - V_n(0;1) \geq \bar{\chi}, \quad \text{if} \quad z^n = 1. \tag{C.2}$$

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

$$V_n(1;0) - V_n(0;0) \leq \bar{\chi}, \quad \text{if} \quad z^n = 0. \tag{C.3}$$

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

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

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 C.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 (C.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 $V(1;0) = \frac{1}{2}\sqrt{\tau_v^{-1}\tau_u^{-1}}$. Thus, $V_n(1;0) - V_n(0;0) - \bar{\chi} = \frac{1}{2}\sqrt{\tau_v^{-1}\tau_u^{-1}} - \bar{\chi} \leq 0$, so (C.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 (C.2) nor (C.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

$$(\Pi + 2)^2(\Pi + 1) = \frac{\tau_v^{-1}\tau_u^{-1}}{\bar{\chi}^2}.$$

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 $n \to \infty$. $\square$

C.1 Asymmetric equilibrium

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

Lemma C.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,*

$$\beta_n = \sqrt{\frac{\tau_v^{-1}}{\tau_u^{-1}} \frac{\Pi_n}{(1+\Pi_n)^2}}$$

$$Y_n = \sqrt{\frac{\tau_u^{-1}}{\tau_v^{-1}}\Pi_n} v + \sqrt{\tau_u^{-1}} e_u$$

$$p^n = \sqrt{\tau_v^{-1}} \Big(\frac{\Pi_n}{1+\Pi_n} \frac{v}{\sqrt{\tau_v^{-1}}} + \frac{\sqrt{\Pi_n}}{1+\Pi_n} e_u \Big)$$

where $e_u = u/\sqrt{\tau_u^{-1}} \sim Normal(0,1)$. 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,*

$$\lim_{n\to\infty} \beta_n = \sqrt{\frac{\tau_v^{-1}}{\tau_u^{-1}}} \frac{\sqrt{\Pi^*}}{1+\Pi^*}$$
$$\lim_{n\to\infty} Y_n = \sqrt{\tau_u^{-1}} \Big[\sqrt{\Pi^*} \frac{v}{\sqrt{\tau_v^{-1}}} + e_u \Big]$$
$$\lim_{n\to\infty} p^n = \sqrt{\tau_v^{-1}} \frac{\sqrt{\Pi^*}}{1+\Pi^*} \Big[\sqrt{\Pi^*} \frac{v}{\sqrt{\tau_v^{-1}}} + e_u \Big]$$

Proposition C.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:*

$$\text{(liquidity)} \quad \mathcal{L} = \sqrt{\frac{\tau_v}{\tau_u}} \frac{1+\Pi^*}{\sqrt{\Pi^*}}$$
$$\text{(informativeness)} \quad \mathcal{I} = \frac{\Pi^*}{1+\Pi^*}$$
$$\text{(volatility)} \quad \mathcal{V} = \tau_v^{-1} \frac{\Pi^*}{(1+\Pi^*)}$$
$$\text{(volume)} \quad \mathcal{Y} = \sqrt{\tau_u^{-1} \tau_v^{-1} \Pi^*}$$

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

$$\lambda^n = \sqrt{\frac{\tau_v^{-1}}{\tau_u^{-1}} \frac{\Pi_n}{(1+\Pi_n)^2}} \quad \text{and} \quad \beta^n s = \sqrt{\frac{\tau_u^{-1}}{\tau_v^{-1}}} \Pi_n v,$$

and so

$$Y_n = \sqrt{\frac{\tau_u^{-1}}{\tau_v^{-1}}} \Pi_n v + \sqrt{\tau_u^{-1}} e_u$$
$$p^n = \sqrt{\tau_v^{-1}} \Big(\frac{\Pi_n}{1+\Pi_n} \frac{v}{\sqrt{\tau_v^{-1}}} + \frac{\sqrt{\Pi_n}}{1+\Pi_n} e_u \Big)$$

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

$$Y_n \to \sqrt{\tau_u^{-1}}\left(\sqrt{\Pi^*}\frac{v}{\sqrt{\tau_v^{-1}}} + e_u\right)$$
$$p^n \to \sqrt{\tau_v^{-1}}\frac{\sqrt{\Pi^*}}{1+\Pi^*}\left(\sqrt{\Pi^*}\frac{v}{\sqrt{\tau_v^{-1}}} + e_u\right)$$

This proves Lemma C.1.

Proposition C.1 is proved by combining the results of Lemma C.1 with the definitions of the measures, contained in Definitions 3-5. □

Comparing the result above with Theorem **??**, 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:

$$\text{(symmetric equilibrium)} \quad \lim_{n\to\infty} p^n = \sqrt{\tau_v^{-1}}\frac{\sqrt{Z^*}}{1+Z^*}\left[\sqrt{Z^*}\frac{v}{\sqrt{\tau_v^{-1}}} + \frac{e_s+e_u}{\sqrt{2}}\right]$$
$$\text{(asymmetric equilibrium)} \quad \lim_{n\to\infty} p^n = \sqrt{\tau_v^{-1}}\frac{\sqrt{\Pi^*}}{1+\Pi^*}\left[\sqrt{\Pi^*}\frac{v}{\sqrt{\tau_v^{-1}}} + e_u\right]$$

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 $\frac{e_s+e_u}{\sqrt{2}}$ 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 2 and Proposition C.1, for symmetric and asymmetric equilibria, are identical.

Theorem C.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 $\bar{\chi}$.

D 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 $s_i = v + \varepsilon_i$ and solve

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

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

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

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 $\sigma_i := \mathrm{Var}[s_i]$ in that

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

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 $(\sigma_i - \tau_v^{-1})^{-1} = \frac{2}{\tau_v^{-1}} \frac{z_i}{1-z_i}$.[10] 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

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

versions explicitly model the noise u grow with the number of traders (i.e., u is the noise per capita).[11] 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, s_i) implies the distribution of (s_i, p) is

$$(s_i, p) \sim \text{Normal}\left(0, \tilde{\Sigma}_i\right),$$
$$\text{where} \quad \tilde{\Sigma}_i := \begin{pmatrix} \sigma_i & \beta_v \tau_v^{-1} \\ \beta_v \tau_v^{-1} & \beta_v^2 \tau_v^{-1} + \beta_u^2 \tau_u^{-1} \end{pmatrix}$$

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

$$(v \mid s_i, p) \sim \text{Normal}(\tilde{\mu}_{vi}, \tilde{\sigma}_{vi}),$$
$$\text{where} \quad \tilde{\mu}_{vi} := \frac{\tau_v^{-1}}{(\sigma_i - \tau_v^{-1})\beta_v^2 \tau_v^{-1} + \sigma_i \beta_u^2 \tau_u^{-1}} \left(\beta_u^2 \tau_u^{-1} s_i + (\sigma_i - \tau_v^{-1})\beta_v p\right)$$
$$\tilde{\sigma}_{vi} := \tau_v^{-1} - \frac{\tau_v^{-1}}{(\sigma_i - \tau_v^{-1})\beta_v^2 \tau_v^{-1} + \sigma_i \beta_u^2 \tau_u^{-1}} \begin{pmatrix} 1 \\ \beta_v \end{pmatrix}^\top \begin{pmatrix} \beta_v^2 \tau_v^{-1} + \beta_u^2 \tau_u^{-1} & -\beta_v \tau_v^{-1} \\ -\beta_v \tau_v^{-1} & \sigma_i \end{pmatrix} \begin{pmatrix} 1 \\ \beta_v \end{pmatrix}$$
$$= \tau_v^{-1} \frac{(\sigma_i - \tau_v^{-1})\beta_u^2 \tau_u^{-1}}{(\sigma_i - \tau_v^{-1})\beta_v^2 \tau_v^{-1} + \sigma_i \beta_u^2 \tau_u^{-1}}$$

Then, the standard CARA solution for asset demand is

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

[11] One could consider an alternative economy in which the noise is given by $u_n \sim \text{Normal}(0, \tau_u^{-1}/n^2)$. 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 τ_u^{-1} 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 τ_u^{-1}.

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

$$p = \frac{\tau_v^{-1}}{(\sigma^* - \tau_v^{-1})\beta_v^2\tau_v^{-1} + \sigma^*\beta_u^2\tau_u^{-1}}\Big(\beta_u^2\tau_u^{-1}v + (\sigma^* - \tau_v^{-1})\beta_v p\Big) + \gamma\tau_v^{-1}\frac{(\sigma^* - \tau_v^{-1})\beta_u^2\tau_u^{-1}}{(\sigma^* - \tau_v^{-1})\beta_v^2\tau_v^{-1} + \sigma^*\beta_u^2\tau_u^{-1}}u$$

Rearranging this expression, we have

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

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):

$$\begin{aligned} \beta_v &= \frac{\tau_v^{-1}\beta_u^2\tau_u^{-1}}{(\sigma^* - \tau_v^{-1})\beta_v^2\tau_v^{-1} + \sigma^*\beta_u^2\tau_u^{-1} - \tau_v^{-1}(\sigma^* - \tau_v^{-1})\beta_v} \\ \beta_u &= \frac{\gamma\tau_v^{-1}(\sigma^* - \tau_v^{-1})\beta_u^2\tau_u^{-1}}{(\sigma^* - \tau_v^{-1})\beta_v^2\tau_v^{-1} + \sigma^*\beta_u^2\tau_u^{-1} - \tau_v^{-1}(\sigma^* - \tau_v^{-1})\beta_v} \end{aligned}$$

The solution is

$$\begin{aligned} \beta_v &= \frac{\tau_v^{-1}[1 + \gamma^2\tau_u^{-1}(\sigma^* - \tau_v^{-1})]}{\tau_v^{-1} + \gamma^2\sigma^*\tau_u^{-1}(\sigma^* - \tau_v^{-1})} \\ \beta_u &= \gamma(\sigma^* - \tau_v^{-1})\beta_v \end{aligned}$$

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.

D.1 Equilibrium information-acquisition

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

(1982), first compute the ex-ante utility

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

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

$$
\begin{aligned}
\varphi(s,p) &= \frac{1}{2\pi}K_0^{1/2}\exp\Big[-\frac{1}{2}\big(k_s s^2 + 2k_{sp}sp + k_p p^2\big)\Big] \\
K_0 &:= \frac{1}{\det(\tilde{\Sigma})} = \frac{1}{(\sigma(z)-\tau_v^{-1})\beta_v^2\tau_v^{-1} + \sigma(z)\beta_u^2\tau_u^{-1}} \\
k_s &= \tilde{\Sigma}_{11}^{-1} = K_0(\beta_v^2\tau_v^{-1} + \beta_u^2\tau_u^{-1}) \\
k_{sp} &= \tilde{\Sigma}_{12}^{-1} = \tilde{\Sigma}_{21}^{-1} = -K_0\beta_v\tau_v^{-1} \\
k_p &= \tilde{\Sigma}_{22}^{-1} = K_0\sigma(z)
\end{aligned}
$$

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

$$
\begin{aligned}
a &:= -\frac{1}{2\tilde{\sigma}_v} = -\frac{1}{2}K_0^{-1}\frac{1}{(\sigma(z)-\tau_v^{-1})\beta_u^2\tau_v^{-1}\tau_u^{-1}} \\
b_s &:= K_0\beta_u^2\tau_u^{-1}\tau_v^{-1} \\
b_p &:= K_0[(\sigma(z)-\tau_v^{-1})\beta_v\tau_v^{-1} - K_0^{-1}]
\end{aligned}
$$

Using this expression and the normal density, compute

$$
\begin{aligned}
\mathbb{E}\Big[\exp\Big(-\frac{1}{2}\frac{(\tilde{\mu}_v - p)^2}{\tilde{\sigma}_v}\Big)\Big] &= \mathbb{E}[e^{a(b_s s + b_p p)^2}] \\
&= \frac{1}{2\pi}K_0^{1/2}\iint e^{a(b_s s + b_p p)^2} e^{-\frac{1}{2}(k_s s^2 + 2k_{sp}sp + k_p p^2)}\mathrm{d}s\mathrm{d}p \\
&= \frac{1}{2\pi}K_0^{1/2}\iint \exp\Big[-\frac{1}{2}\binom{s}{p}^{\top}\Omega^{-1}\binom{s}{p}\Big]\mathrm{d}s\mathrm{d}p,
\end{aligned}
$$

where

$$
\Omega^{-1} := \begin{pmatrix} k_s - 2ab_s^2 & k_{sp} - 2ab_s b_p \\ k_{sp} - 2ab_s b_p & k_p - 2ab_p^2 \end{pmatrix}
$$

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

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

Therefore,[12]

$$\frac{K_0}{\det(\Omega^{-1})} = \left(\tau_v + (\sigma(z) - \tau_v^{-1})^{-1} + \frac{\beta_v^2}{\beta_u^2 \tau_u^{-1}}\right)^{-1} \left(\beta_u^2 \tau_u^{-1} + (1-\beta_v)^2 \tau_v^{-1}\right)^{-1}.$$

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

$$\max_z -\gamma c(z) + \frac{1}{2}\log\left(\tau_v + (\sigma(z) - \tau_v^{-1})^{-1} + \frac{\beta_v^2}{\beta_u^2 \tau_u^{-1}}\right) + \frac{1}{2}\log\left(\beta_u^2 \tau_u^{-1} + (1-\beta_v)^2 \tau_v^{-1}\right)$$

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

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

[12]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)

$$\frac{K_0}{\det(\Omega^{-1})} = \left(\tau_v + (\sigma(z) - \tau_v^{-1})^{-1} + \frac{\beta_v^2}{\beta_u^2 \tau_u^{-1}}\right)^{-1} \frac{1}{\beta_u^2 \tau_u^{-1}}.$$

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.

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

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

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

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

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

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

Next, we will examine equilibrium information and its pricing consequences in two limits, $\tau_u^{-1} \to 0$ 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.

D.2 Small-noise limit

As a first result, notice that as noise vanishes ($\tau_u^{-1} \to 0$), the left-hand-side of (D.2) blows up for any $z^* > 0$. Hence, a requirement is $z^* \to 0$ as $\tau_u^{-1} \to 0$, assuming an equilibrium exists for each τ_u^{-1} along this sequence.[13] 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

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

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 (D.2) shows that $\tau_u^{-1}/(z^*)^2$ must not explode as $\tau_u^{-1} \to 0$. We can also show that $\tau_u^{-1}/(z^*)^2$ cannot be vanishing, because if it did, the left-hand-side of (D.2) would necessarily exceed 1 for all small enough τ_u^{-1} (given we know $z^* > 0$ along the sequence). Hence, $\tau_u^{-1}/(z^*)^2$ has a finite, non-zero limit. Now, returning to the equilibrium price in (D.3), and using this fact, we find that β_v and $\sqrt{\tau_u^{-1}}\beta_u$ both converge to a finite constants. The limiting value for β_v is necessarily positive but less than 1. The limiting value for $\sqrt{\tau_u^{-1}}\beta_u$ 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 $z^* \sim A\tau_u^{-1\zeta/2}$ for some $A > 0$ and $\zeta > 0$. The left-hand-side of (D.2) behaves asymptotically like

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

For this expression to remain non-trivial, hence remain consistent with (D.2), we see that $\zeta < 1$ is required. In that case, $\beta_v \to 1$ and $\sqrt{\tau_u^{-1}}\beta_u \to 0$, 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, $c'(z) = \kappa z^\rho$, so

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

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 $z^*(\tau_u^{-1})$ vanishing at rate $\zeta = (1+\rho/2)^{-1}$ such that the equilibrium condition (D.2) holds with equality for every τ_u^{-1} small enough. Consequently, equilibria with information-acquisition exist for all τ_u^{-1} small enough, and these converge to a fully-revealing equilibrium.

D.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 (D.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 (D.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 (D.2) asymptotically looks like

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

If $\zeta > 1/2$, then this expression vanishes in the $\gamma \to 0$ limit, which from (D.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.

D.4 Price informativeness, volatility, and liquidity

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

$$\begin{aligned}
\mathcal{I} &:= 1 - \frac{\text{Var}[v \mid p]}{\text{Var}[p]} \\
\mathcal{V} &:= \text{Std}[p] \\
\mathcal{L} &:= [\partial p / \partial u]^{-1}
\end{aligned}$$

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 (D.3) and properties of the joint normal distribution, compute

$$\begin{aligned}
\mathcal{I} &= \frac{\beta_v^2 \tau_v^{-1}}{\beta_u^2 \tau_u^{-1} + \beta_v^2 \tau_v^{-1}} = \frac{1}{1 + \frac{\gamma^2}{4}\left(\frac{1-z^*}{z^*}\right)^2 \tau_u^{-1}} \\
\mathcal{V} &= \sqrt{\beta_u^2 \tau_u^{-1} + \beta_v^2 \tau_v^{-1}} \\
\mathcal{L} &= \beta_u^{-1}
\end{aligned}$$

Recall that in the limits $\tau_u^{-1} \to 0$ or $\gamma \to 0$, the equilibrium can become fully-revealing (for $\tau_u^{-1} \to 0$, 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 $\mathcal{V} \to \tau_v^{-1}$. For the small-noise limit, taking $\tau_u^{-1} \to 0$ with $z^* \sim \tau_{u^{\zeta}}^{-1}$ and $\zeta < 1/2$, notice that $\mathcal{I} \to +\infty$ and $\mathcal{V} \to \tau_v^{-1}$. Unlike our strategic-trading model, these limiting cases are the only ones where the fully-efficient "magical markets" outcomes arise in an NREE.

D.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

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

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 $s_i = \frac{2}{\tau_v^{-1}}\frac{z_i}{1-z_i}$. Hence, the function (D.4) represents quadratic costs over this signal error precision.

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

$$\gamma\kappa\Big[\frac{z^*}{1-z^*}+2\big(\frac{z^*}{1-z^*}\big)^2+\frac{4}{\gamma^2\tau_u^{-1}\tau_v^{-1}}\big(\frac{z^*}{1-z^*}\big)^3\Big]=1 \tag{D.5}$$

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

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

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 $s_i = \frac{2}{\tau_v^{-1}}\frac{z_i}{1-z_i}$, the function (D.6) represents linear costs over this signal error precision.

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

$$\gamma\chi\Big[1+2\big(\frac{z^*}{1-z^*}\big)+\frac{4}{\gamma^2\tau_u^{-1}\tau_v^{-1}}\big(\frac{z^*}{1-z^*}\big)^2\Big]=1 \tag{D.7}$$

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 equilibrium.) This unique solution is

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

Market measures in the examples. Figures D.1-D.3 display informativeness, volatility, and liquidity in the two examples. We vary τ_u^{-1}, τ_v^{-1}, 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 $\tau_v^{-1} \to 0$, the equilibrium becomes fully-revealing, with $\mathcal{I} \to +\infty$ and $\mathcal{V} \to \tau_v^{-1}$. As $\tau_u^{-1} \to 0$, the equilibrium becomes fully-revealing only if $c'(0)=0$ but not if $c'(0)>0$.

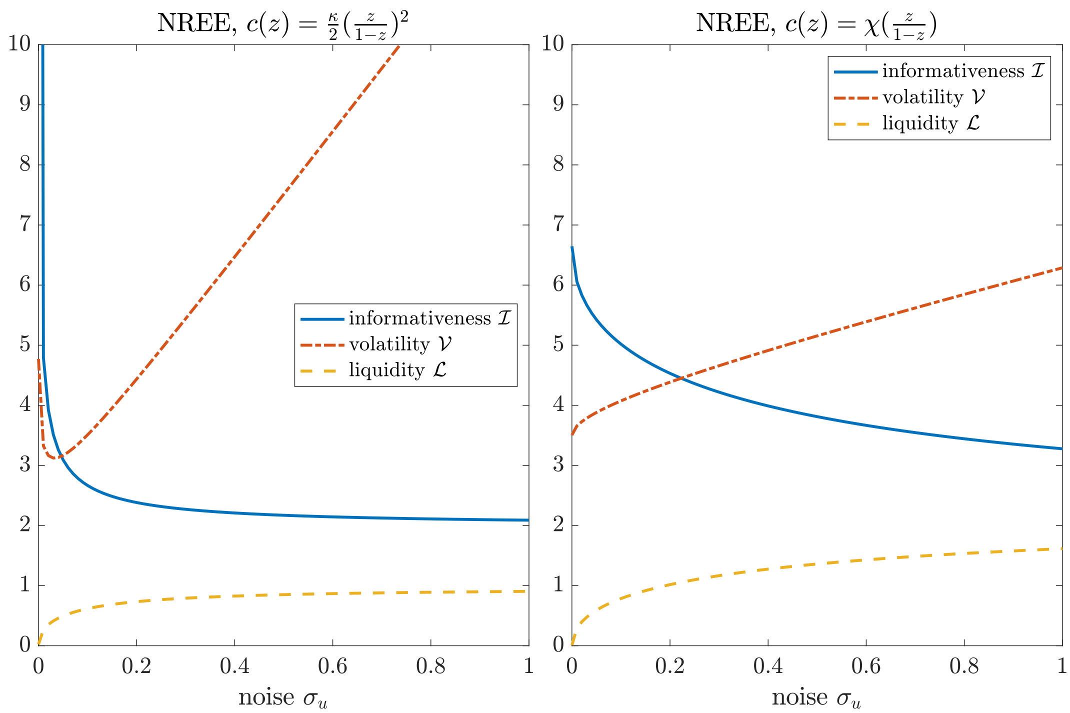

Figure D.1: Measures as a function of τ_u^{-1}. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\tau_u^{-1} = 0.5$, $\tau_v^{-1} = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.

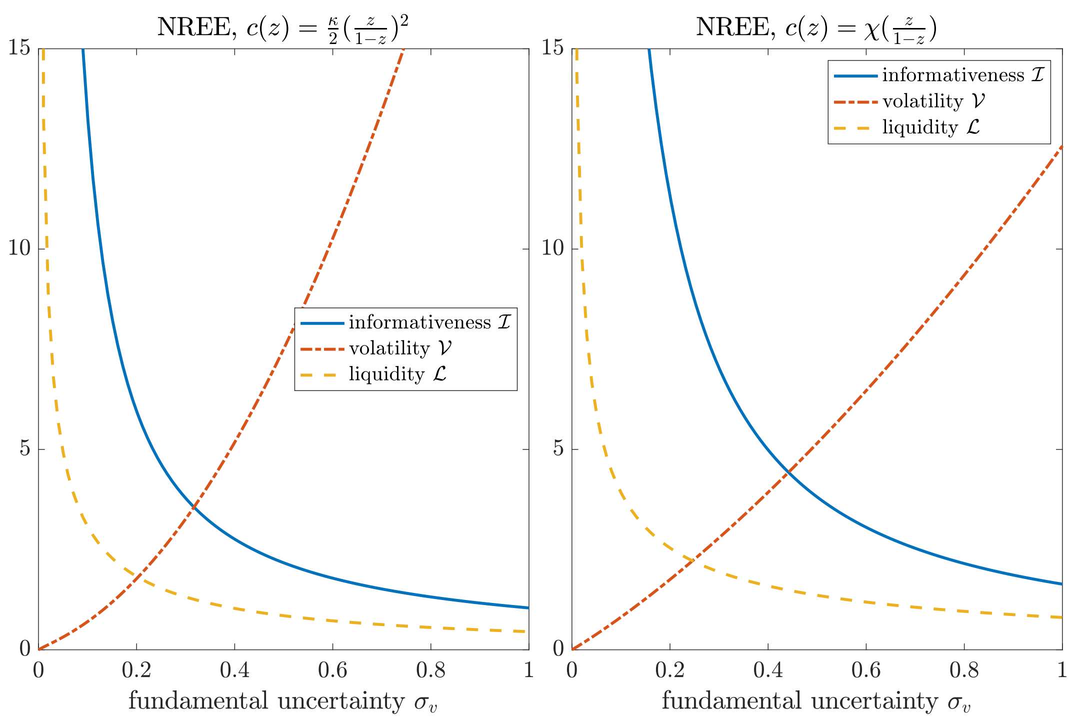

Figure D.2: Measures as a function of τ_v^{-1}. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\tau_u^{-1} = 0.5$, $\tau_v^{-1} = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.

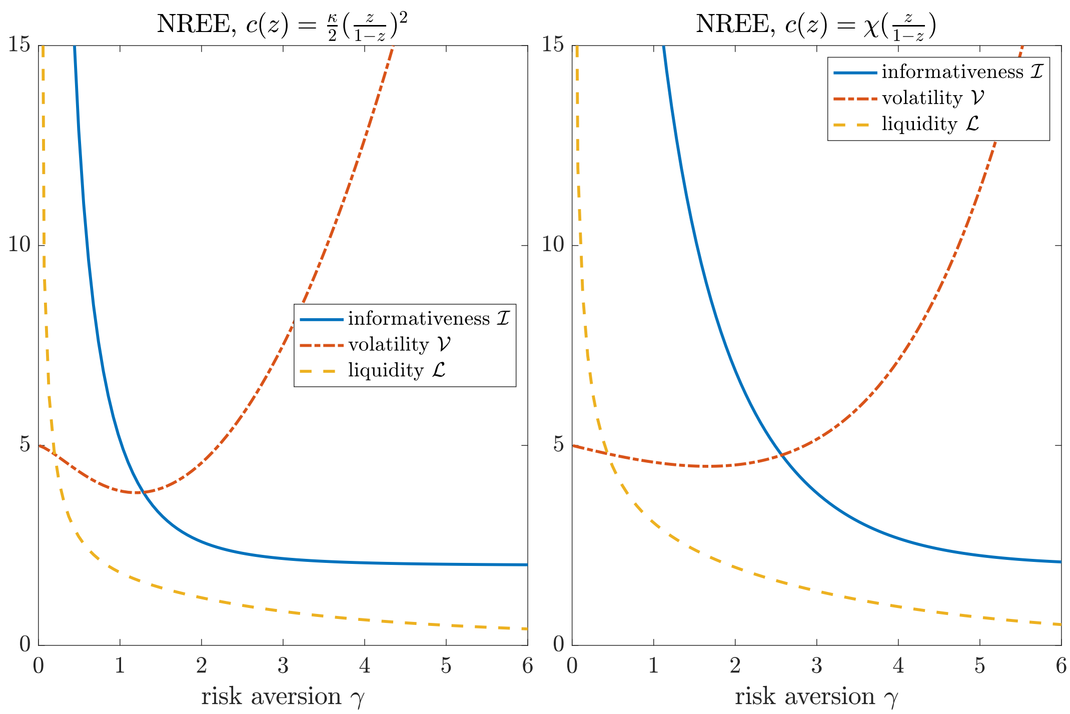

Figure D.3: Measures as a function of γ. The volatility measure is scaled for aesthetic purposes. Baseline parameters: $\tau_u^{-1} = 0.5$, $\tau_v^{-1} = 0.5$, and $\gamma = 3$. For the information cost functions, we use $\kappa = 1$ for the quadratic cost and $\chi = 0.1$ for the linear cost.