Annals of Finance (2006) 2: 369–395
DOI 10.1007/s10436-006-0045-z

RESEARCH ARTICLE

Martin Dierker

Endogenous information acquisition with Cournot competition

Received: 18 January 2006 / Accepted: 5 March 2006 / Published online: 18 May 2006

Abstract This paper studies information production in a model where both entry of analysts and their optimal information quality is endogenous. We show existence of the Bayesian–Nash equilibrium and solve for it in closed form. The model displays rich behavior. In particular, we find that the precision of an individual signal will always be bounded from above by the precision of the prior belief on payoff uncertainty. Furthermore, we give examples that contradict the naive intuition about information acquisition. For instance, we show how a change in the cost structure that makes information cheaper decreases price informativeness, while at the same time market liquidity and the amount of resources society spends on information acquisition can change either way. The model gives a simple, fully rational explanation on why the number of analysts following a stock can be quite large. Endogenizing the cost of information by allowing the manager to choose an optimal informational policy, we find a variety of optima that depend discontinuously on the model parameters. As a consequence, among two similar firms, one may find it optimal to attract many analysts, the other will cooperate with only a few.

Keywords Analyst coverage · Informed trading · Disclosure

JEL Classification Numbers G10 · G14 · M41

I am greatly indebted to Antonio Bernardo and Avanidhar Subrahmanyam, whose support, encouragement, insight, and many valuable comments made this paper possible. In addition, I also thank Justin Chan, Sudipto Dasgupta, Jason Hsu, Martin Nielsen, and Elena Sernova, the editor and an anonymous referee.

M. Dierker
C.T. Bauer College of Business, University of Houston,
Houston, TX 77204, USA
E-mail: mdierker@uh.edu

1 Introduction

The number of analysts following a stock varies strongly from firm to firm. This has been documented by a vast empirical literature that studies the determinants of analyst coverage, e.g., Bhushan (1989) and Brennan and Subrahmanyam (1995). Many researchers believe that the surprisingly high numbers of analysts following certain stocks can only be explained by behavioral reasons, herding, or time patterns of information arrival. But the decision to follow a stock also implies a choice of how much information to acquire. This paper presents a theory that links analyst coverage with the choice of an optimal research intensity, and thus equilibrium forecast dispersion, a variable that has been shown to explain future stock returns (Diether et al. 2002). Intuition tells us that both of these variables are closely related. We formalize this relationship and study the joint determinants of analyst activity and research intensity in the presence of interaction effects.

In our model, some agents ("analysts") can acquire information which allows them to make forecasts about a stock's fundamental value. The quality of the information can also be chosen and we make the natural assumption that the cost of acquiring information is increasing in its precision. Thus, analysts must decide whether to become informed and, if so, the amount of information to acquire. The rest of the model is in the spirit of Kyle (1985). The benefit of acquiring information comes in the form of trading gains at the expense of uninformed traders in the market. Analysts behave strategically in the Cournot sense. As analysts choose their research intensity and trading aggressiveness, they take the behavior of other analysts and the effect of their own decisions on them into account. Prices in the model are set by competitive market makers. We define the Bayesian–Nash equilibrium and proceed to show existence. The equilibrium is solved for in closed form, and it is shown to be unique among the class of equilibria in which informed agents behave symmetrically and trading strategies and the market maker pricing rule are linear.

We find the following results. In equilibrium, the optimal research intensity is a constant multiple of ex ante payoff uncertainty. This multiple is exclusively determined by the convexity of the cost function; other parameters, such as the extent of liquidity trading, play no role. With a convex cost function, analysts will never acquire very precise signals. In fact, the precision of an analyst's forecast will never exceed the precision of the prior belief on payoff uncertainty. This is a direct consequence of the interaction of entry of analysts and endogenous choice of information quality. Intuitively, high signal precision will lead to high trading profits, which will encourage the entry of new analysts. This, in turn, will reduce the optimal research intensity.

The model has rich comparative statics. In particular, it allows us to study how the cost structure of information acquisition affects price informativeness and market liquidity. Without endogenous precision choice (as in Admati and Pfleiderer 1988) cheaper information will always lead to more informative prices. In our model with endogenous information quality, however, cheaper information can lead to less informative prices when the payoff uncertainty is sufficiently small. In fact, cheaper information can result in a less informative stock price, while at the same time market liquidity deteriorates and the amount of resources society spends on information acquisition increases.

Since the quality of information is not observable, the implication that it will always be bounded by the precision of the prior belief cannot be tested directly. But the fact that information quality will be low in equilibrium leads to widely dispersed beliefs about the asset value, a large number of informed agents. We will show later that both these effects lead to high trading volume. This prediction is consistent with the empirical evidence in Bhushan (1989) showing that analyst coverage is greater for stocks with high trading volume. Furthermore, our result of high volume of informational trade may help to explain the high levels of observed trading volume, which is difficult to explain theoretically.

Finally, we endogenize the cost of information by allowing the manager to choose an informational policy towards the shareholders. The manager sets the parameters of the cost function as to maximize benefits from informational efficiency. At the same time, he will also want to minimize the cost of illiquidity.[1] Since these two effects do not necessarily go in the same direction, we find a variety of optimal policies. Some firms favor many analysts by making information cheap or inducing low research intensities. Other optimal policies result in low analyst coverage by making information expensive (to deter informed trade) or implementing high (i.e., close to the upper feasible bound) research intensities. Optimal policies are discontinuous in model parameters. This can explain the observed cross-sectional variation in analyst coverage. Among two similar firms, one may find it optimal to attract many analysts, the other will prefer few analysts.

The seminal papers of Grossman and Stiglitz (1980), Kyle (1985, 1989), and Admati and Pfleiderer (1988) all stress the importance of endogenous information acquisition. However, little theoretical work has been done subsequently. Virtually all papers model information acquisition as "agents have the option to acquire a signal of some prespecified precision about the asset value", and allow the number of informed agents to vary. Verrecchia (1982) studies information quality by taking the opposite approach: holding the number of informed agents in the economy N fixed, the precision of their signals is determined within the model. However, Verrecchia uses the noisy rational expectations approach of Hellwig (1980), thus not allowing for market making, noise traders and strategic behavior. Kyle (1989) compares endogenous information acquisition in strategic and competitive models. Endogenous information acquisition is also addressed in the literature on markets for information, e.g., Admati and Pfleiderer (1986), Caballe (1993).

In reality, the same amount of information can be produced with few analysts acquiring precise signals or many informed agents with widely dispersed beliefs. No paper makes any prediction on this matter. From a single informed agent with very precise information to a vast number of agents each knowing very little, everything is possible, depending on parametric assumptions. The key innovation of this paper is to study equilibria in which both entry and signal precision vary. From this, we can draw several new insights. Information precision will always be relatively low, since otherwise substantial profits from informed trading will entice more analysts to enter the market. Secondly, we find a separating result,

[1] Market liquidity determines the liquidity trader's expected loss to informed analysts, who use the gains to cover their research expenses. Since information acquisition has no tangible output, a social planner needs to trade off price discovery with the cost of information acquisition, which is determined by market liquidity. The manager faces the same problem, since we assume that he has to compensate liquidity traders for their expected loss to induce them to hold the stock.

telling us that optimal research intensity will only depend on two of the four model parameters. Furthermore, differences in optimal informational policies can explain strong cross-sectional variation in analyst coverage. Comparative statics allow us to better understand the effect of the cost structure of information on informational efficiency and market liquidity (which includes some counterintuitive examples). On the other hand, we find that comparative statics with respect to noise trading and payoff uncertainty remain qualitatively the same as in the simpler models like Kyle (1985, 1989) and Admati and Pfleiderer (1988).

It has often been argued that the cross-sectional variation in analyst coverage is too large to be rational. Hirshleifer et al. (1994) explain high analyst coverage of certain stocks by patterns in information arrival. Also, behavioral reasons or analyst herding (Welch 2000) are possible explanations. This paper, in contrast, gives a very simple and fully rational explanation of variation in analyst following: similar firms can have informational policies geared at either low or high analyst coverage.

The paper is organized as follows. Section 2 presents the model. Section 3 discusses the equilibrium concept and presents the closed-form solution. Furthermore, existence of the equilibrium is proved along with a uniqueness statement. We then proceed to study the properties of the model in section 4. In section 5 we endogenize the cost of information by allowing the firm's manager to select an optimal informational policy. In section 6, we explore the empirical implications of our model, while section 7 studies the robustness of our results. Finally, section 8 concludes.

2 The model

We consider a one-period economy with a single risky asset paying an uncertain amount

$$F = \bar{F} + \delta.$$

The term $\bar{F}$ is the constant mean return of the project and is known to all agents with certainty. Since it does not play a role, we set $\bar{F} = 0$. The term δ represents the part of the value about which informed agents can collect information and is assumed to be normally distributed with mean zero and variance σ_δ^2.[2]

The economy consists of $\bar{N}$ potentially-informed, risk-neutral agents ("analysts"), a mass of uninformed "noise traders", and a competitive, risk-neutral market maker. Analysts have access to a research technology. At a cost, analyst i can observe the private signal s_i where

$$s_i = \delta + \epsilon_i.$$

We assume that analyst i's forecast error, ϵ_i, is unbiased ($E(\epsilon_i) = 0$), normally distributed, and independent of all other random variables. In our framework, analysts not only choose whether or not to gather information but also the precision

[2] We could also allow for an additional unpredictable idiosyncratic shock θ to the project value that the analysts cannot learn about, e.g., letting $F = \bar{F} + \delta + \theta$ where θ is assumed to be normally distributed with mean zero, variance σ_θ^2, and independent of δ. This would not change the subsequent analysis at all, since agents are risk neutral and the θ term would drop out of the expression for expected profits.

of information. We assume that gathering more precise signals (lower standard deviation of ϵ_i) is more costly and that information cannot be credibly sold to other agents. Let σ_ϵ denote the desired standard deviation of ϵ and let $\tau = 1/\sigma_\epsilon$ denote the "research intensity" (the inverse of the standard deviation). We assume the research technology is given by an increasing and convex cost function $c(\tau)$. For the sake of tractability and interpretability, we consider cost functions of the form:

$$c(\tau) = k \cdot \tau^{\alpha}.$$

The parameter $k > 0$ represents the level of the cost function, while the parameter α determines the degree of convexity of the cost function. Economically, this relates to returns to scale. The higher α, the cheaper the same amount of information can be produced by two analysts in contrast to only one analyst. Unfortunately, empirical proxies for these parameters may be hard to obtain.

Informed analysts submit orders x_i to the market maker.[3] Analysts understand the impact of their trades on equilibrium prices and explicitly account for this when choosing their optimal orders. Because they are better-informed, analysts make trading profits which allows them to recoup research expenses. It is straightforward to show that it is optimal for analysts who choose not to gather information to abstain from trading in the trading round. Noise traders collectively submit random orders z which are assumed to be normally distributed with mean zero, variance σ_z^2, and independent of all other random variables. Thus, the net total order flow (i.e., order imbalance) is given by $y = z + \sum_i x_i$. The presence of these noise traders makes it impossible for the market-maker to learn the asset value perfectly, since they only observe the net order flow. Market-makers are assumed to absorb any excess demand or supply of the asset. Moreover, competition among market-makers ensures that the equilibrium price is set equal to the expected value of the asset payoff conditional on the observed net order flow.[4]

Timing in our model is as follows. First, analysts have to simultaneously decide whether to collect information at all, how much information to acquire, and how aggressively to trade on their information. At the same time, the market maker sets a pricing scheme. After that, analysts receive costly private signals and both informed analysts and noise traders submit orders to the market maker. Finally, the market maker sets prices, trading takes place, and profits are realized.

The main goal of our paper is study the analysts' decision to collect information and if so, how much information to acquire. Obviously, the analyst's problem is quite complex since the returns to gathering information depend on how many other analysts choose to gather information, how much information they acquire, and how aggressively she and others can trade in the asset market.

In relationship with the existing literature, our model is most closely related to the one presented in Kyle (1985) and Admati and Pfleiderer (1988). The type of competition in our model closely parallels theirs. The main difference is that information acquisition in their model is a zero-one decision, whereas here they have

[3] One could imagine potential conflicts of interest between the research and trading arms of the analysts' firms. We abstract from these conflicts and assume that information flows perfectly between these two groups and that the orders submitted optimally reflect the information gathered by the analysts.

[4] This is usually justified with Bertrand competition among market makers, e.g., Holden and Subrahmanyam (1992).

to choose whether or not to acquire information as well as the optimal information quality they want to acquire. Endogenous precision has been studied by Verrecchia (1982), who uses the competitive rational expectations approach of Hellwig (1980). The difference in modeling approaches makes comparison difficult.[5]

3 Equilibrium

Analysts compete in both information acquisition and trading. We assume that competition is of the Cournot type and since agents also have private information, the formal equilibrium concept is Bayesian–Nash equilibrium. In this section, we consider the symmetric Bayesian–Nash equilibrium,[6] which is defined below:

Definition *A symmetric Bayesian–Nash equilibrium with endogenous information acquisition is a set of N informed analysts, $\bar{N}-N$ uninformed analysts, a set of research intensities $(\tau_1, \ldots, \tau_N)$, a set of informed analysts' trading strategies $(x_i(s_i), i = 1, \ldots, N)$, and a market-maker pricing rule $p(y)$ such that:*

1. *Given $(\tau_j, x_j)_{j=1,\ldots,i-1,i+1,\ldots N}$ and $p(y)$, τ_i and x_i maximize informed analyst i's expected profit*
2. *Given $(\tau_i, x_i)_{i=1,\ldots,N}$ and $p(y)$, $\bar{N} - N$ analysts find it optimal to choose not to gather information*
3. *Given the trading rules $(x_i)_{i=1,\ldots,N}$; the market maker sets $p(y) = E[F \mid y]$*
4. *$\tau_i = \tau$ and $x_i(\cdot) = x(\cdot) \forall i$; and*
5. *The expected profit for all informed and uninformed analysts is zero.*

The first condition of our definition simply states that given the research intensity and trading strategies of the other informed analysts, each informed analyst chooses their research intensity and trading strategy optimally. The second condition ensures that given the research intensity and trading strategies of the informed analysts, uninformed analysts find it optimal not to gather information. The third condition ensures that market-makers make zero expected profit. The fourth condition imposes symmetry on the research and trading strategies of the informed analysts. Finally, since all analysts are identical ex ante, the fifth condition ensures that they have the same expected profits. In reality, the number N of informed analysts clearly has to be an integer, however, we allow N to be a fraction and ignore the issue of discreteness, since it complicates the analysis without changing the qualitative nature of the results.

The following proposition provides a solution to the symmetric Bayesian–Nash equilibrium. This solution has linear trading rules and linear pricing rules and is shown to be unique among all linear, symmetric Bayesian–Nash equilibria. As in most Kyle type models, it is possible that other, non-linear equilibria exist. But note that the linearities are not ex ante imposed in the agents' strategy sets: As long as the informed analysts use a linear trading strategy, the market maker will use a linear pricing rule and vice versa.

[5] Section 10 of Kyle (1989) compares endogenous information acquisition in strategic and competitive models.

[6] We conjecture that, with symmetric agents, symmetry of the equilibrium follows from the nature of Cournot competition. However, we have not tried to prove this.

Proposition 1 *Let $\alpha \in (1,2)$. There exists a unique, symmetric Bayesian equilibrium among the class of equilibria with linear trading rules and linear pricing rules. The informed analysts' equilibrium trading strategy is given by $x_i = \beta \cdot s_i$, where*

$$\beta = \frac{\sigma_z}{\sqrt{N\left(\sigma_\delta^2 + \tau^{-2}\right)}}.$$

The market-makers pricing rule is given by $p(y) = \lambda \cdot y$, where

$$\lambda = \frac{\sigma_\delta^2}{(N+1)\sigma_\delta^2 + 2\tau^{-2}} \frac{\sqrt{N\left(\sigma_\delta^2 + \tau^{-2}\right)}}{\sigma_z}.$$

The equilibrium number of informed agents is given by the $min\{N, \bar{N}\}$, where

$$N = \frac{\left(-4k^2 + \alpha^2 k^2 + B\right)^2}{3(-2+\alpha)^2 k^2 B}$$

with

$$B = -\left((\alpha-2)^3 k^4 \left((2+\alpha)^3 k^2 - 3\,(\alpha-2)\left(\frac{\alpha}{2-\alpha}\right)^{\frac{\alpha}{2}} A \sigma_\delta^{1+\alpha}\sigma_z\right)\right)^{\frac{1}{3}}$$

and

$$A = -9\,(\alpha-2)\left(\frac{\alpha}{2-\alpha}\right)^{\frac{\alpha}{2}} \sigma_\delta^{1+\alpha}\sigma_z + \sqrt{3}\sqrt{2(2+\alpha)^3 k^2 + 27(\alpha-2)^2\left(\frac{\alpha}{2-\alpha}\right)^{\alpha} \sigma_\delta^{2+2a}\sigma_z^2}.$$

The equilibrium research intensity of the informed analysts is

$$\tau = \sqrt{\frac{2-\alpha}{\alpha}\frac{1}{\sigma_\delta}}$$

which implies an error variance of each individual signal of

$$\sigma_\epsilon^2 = \frac{\alpha}{2-\alpha} \cdot \sigma_\delta^2.$$

Proof See Appendix. □

Note that we impose the restriction $\alpha \in (1,2)$ in the above proposition. It can be shown that for $\alpha \geq 2$ the equilibrium number of informed analysts is $\bar{N}$, and if $\bar{N} \to \infty$, the equilibrium research intensity approaches zero. In this case, we approach the perfectly competitive limit as in Verrecchia (1982). On the other hand, we use the convexity of the cost function (i.e., $\alpha \geq 1$) in the proof of proposition 1. Section 7 presents the equation for equilibrium research intensity with a general cost function, which is still very simple.

The presence of noise traders enables the informed analysts to disguise their trades from market makers. Optimal disguise is critical to maximize profits from informed trading. This is achieved when the expected magnitude of total order flow of informed traders equals that of noise trades. This will in fact be the case:[7]

Corollary 1 *With or without endogenous information acquisition, informed traders as a group manage to optimally conceal their trades, i.e.,*

$$\mathrm{Var}\left(\sum_{i=1}^{N} x_i\right) = \frac{N^2\sigma_z^2}{\sigma_\delta^2 + \tau^{-2}} \cdot \left(\sigma_\delta^2 + \tau^{-2}\right) = \sigma_z^2$$

(which immideately follows from our expression of equilibrium trading aggressiveness, β). So even though the entry of more analysts leads to increased competition among them, the second moment of total informed order flow remains the same. The information content of informed trading, however, will still vary depending on N and τ.

4 Properties of the model

The unique, linear symmetric Bayesian–Nash equilibrium is somewhat complex so in this section we characterize the interesting properties of the equilibrium. While many of comparative statics results can be proven, we have to rely on numerical examples in some cases. The proposition below summarizes the results that we are able to prove analytically. Due to the considerable complexity of the expressions, proofs are not included here, but available from the author' website.

Proposition 2
1. *The number of informed analysts is decreasing in the level of the cost function, k; increasing in payoff uncertainty, σ_δ; and increasing in the amount of liquidity trading, σ_z.*
2. *The equilibrium research intensity is independent of the level of the cost function, k, and the variance of liquidity trading, σ_z; decreasing in the initial payoff uncertainty, σ_δ; and decreasing in the convexity parameter, α. Also, research intensity is bounded from above: $\tau \leq 1/\sigma_\delta$.*
3. *Trading aggressiveness is increasing in k.*
4. *The remaining payoff uncertainty after trading $\Sigma = \mathrm{Var}(\delta \mid \mathrm{price})$ is increasing in k and σ_z; and decreasing in α except when ex-ante payoff uncertainty, σ_δ, is small, i.e.,*

$$\sigma_\delta^2 \leq \exp\left[\frac{3 - 2\log\frac{\alpha}{2-\alpha} - \alpha\log\frac{\alpha}{2-\alpha}}{2+\alpha}\right].$$

5. *Liquidity is decreasing in k for k below a critical value k^* and increasing in k otherwise. Also, while the effect of α on liquidity is generally ambiguous, numerical simulations indicate that the effect of α on liquidity changes sign whenever the effect of α on Σ changes sign.*

[7] I want to thank Jack Hughes for pointing this out to me.

The main innovation of our research is the effect of the cost structure of information acquisition. Here, we find novel implications for price informativeness and market liquidity. Furthermore, we find that the level of the cost, k, does not play a role for the determination of the optimal research intensity, τ, which is determined by the convexity of the cost function, α.

To present complete comparative statics, we also analyse the effect of changes in liquidity trading, σ_z, and ex ante payoff uncertainty σ_δ. Since the former does not affect the choice of optimal research intensity, its behavior is the same as in a model without endogenous precision such as Kyle (1985). The latter has an impact on equilibrium signal precision, thus changing the quantitative results. However, we find these effects small, so that they do not change the qualitative nature of the results.

4.1 Number of informed analysts

Despite the complexity of the expression for the number of informed analysts, N, we can still say a lot about it. It can be shown that N is decreasing in the level of the cost function, k. Note that the error variance of each signal is $\sigma_\epsilon^2 = (\alpha/(2-\alpha)) \cdot \sigma_\delta^2$ which is independent of k; therefore, increasing k must reduce the number of informed analysts. The effect of the convexity of the cost function, summarized by the parameter α, on the number of informed analysts is ambiguous. On one hand, from our expression for the error variance of each signal it is easy to show that increasing α decreases the research intensity which, all else equal, increases the number of informed analysts N. On the other hand, increasing α impacts the cost of information. For $\tau < 1$ (which will always be the case for payoff uncertainty sufficiently large, e.g., $\sigma_\delta^2 > 2.71$), a higher value of α makes information cheaper, which will also raise the number of informed agents. However, for high research intensities τ, an increase in α makes information more expensive, and this effect can be strong enough to overcome the decrease in research intensity. In sum, an increase in α increases the number of informed analysts except possibly when there is a small level of payoff uncertainty.

One can also prove that the number of informed analysts is always increasing in the payoff uncertainty, σ_δ. From our expression for the error variance of each signal, more payoff uncertainty leads to a lower optimal research intensity which, in turn, leads to a greater number of informed analysts. Moreover, more payoff uncertainty results in greater trading profits which encourages entry of informed analysts. Finally, N is increasing in the amount of liquidity trading, σ_z, since more liquidity trading leads to higher trading profits for the informed which again encourages entry by analysts.

4.2 Research intensity

While the expression for the number of analysts N is quite involved, the equilibrium research intensity is of a very simple form, $\sigma_\epsilon^2 = (\alpha/(2-\alpha))\sigma_\delta^2$. Strikingly, it is completely determined by payoff uncertainty σ_δ and the convexity parameter, α, of the cost function. The error variance of a private signal is a constant multiple of the payoff uncertainty of the company where the constant is determined by the convexity parameter, α, of the cost function. A higher α implies a lower research intensity, since it decreases returns to scale, which makes the cost function more favorable to many analysts each producing little information. The extent of liquidity trading, σ_z^2, and the level of the cost function, k, do not have an effect on the signal quality but only influence the number of analysts following the stock.

Mathematically, this is a result of the interaction between endogenous entry and endogenous information quality. Endogenous entry corresponds to a zero profit condition: cost of information $c(\tau)$ equals the gains from informed trading $\pi(\tau)$. Endogenous information quality leads to a first order condition for the optimal research intensity $c'(\tau) = f(\tau)$. Now we can combine these two equations, which both have to hold in equilibrium. The resulting expression is very simple:

$$\frac{c'(\tau)}{c(\tau)} = \frac{2}{\sigma_\delta^2 + \tau^{-2}},$$

which explains the simple expression for the equilibrium research intensity.

Of further interest, σ_ϵ^2 can only take on certain values. For a convex cost function $c(.)$, we have $\alpha \geq 1$, implying that $\frac{\alpha}{2-\alpha} \geq 1$ thus $\sigma_\epsilon \geq \sigma_\delta$, i.e., the precision of a private signal will never be higher than the precision of the prior belief on payoff uncertainty, independent of the overall level of the cost function. This is in contrast to the models of Verrecchia (1982) and Admati and Pfleiderer (1988), where arbitrarily precise signals can be supported. The above bound on signal precision is a direct consequence of the interaction of endogenous entry and endogenous precision. The only way to obtain very precise signals is by incorporating strong non-convexities into the cost function.

In the model of Admati and Pfleiderer (1988), signal precision is an exogenous parameter for which every value is feasible. In Verrecchia (1982), signal precision is endogenous but the number of informed agents, N, is exogenous which implies that by varying the level of the cost function, k, all possible values of signal precision can be supported. In our model, however, signal precision is bounded because both signal precision and entry are endogenous. The reason for this is as follows. Suppose information is very precise. Then analysts will trade aggressively and reap large profits. However, large profit opportunities from informed trading encourage entry of new analysts, which will reduce the equilibrium trading aggressiveness of any one informed analyst, thereby also reducing the marginal benefit of collecting a very precise signal.[8] Thus, our model with endogenous entry and signal precision yields qualitatively different results than models assuming endogeneity of one or the other but not both.

4.3 Trading aggressiveness

The parameter β represents how aggressively informed analysts will trade based on their private information. Note that there is an important link between

[8] In equilibrium, research intensity and the number of informed analysts are inversely related. This is reasonable since the same amount of information can be produced with either few analysts each generating very precise signals or with many analysts each knowing very little. In fact, one can prove that, with endogenous precision only, an increase in the number of informed analysts always leads to lower optimal signal quality. This is the formal statement we have in mind when arguing that more analysts lead to lower research intensities and thus lower trading aggressiveness.

trading aggressiveness and information acquisition, which is given by the equilibrium equation for trading aggressiveness. All else equal, a high research intensity will lead to more aggressive trading behavior. Also, it will tend to decrease the number of informed analysts, N, which will also increase trading aggressiveness.

Trading aggressiveness depends on the level of the cost function, k, only through N. It can be shown that an increase in k will, in general, result in more aggressive trading strategies, since the number of informed agents will be smaller. To understand the effect of other variables on trading aggressiveness, we must rely on numerical examples. The effect of the convexity parameter, α, is again ambiguous. In general, raising α will lead to less aggressive trading, since it implies less precise private signals and, in general, the presence of more informed agents in the market. However, if payoff uncertainty is small and thus research intensity is high, increasing α will lower the number of analysts N. This can result in trading aggressiveness actually rising with α.

For fixed N, an increase in the liquidity trading, σ_z, leads to more aggressive trading behavior. However, the endogenous number of informed analysts increases in σ_z which counteracts this effect. Our numerical examples suggest that typically the former effect dominates the latter so that an increase in liquidity trading leads to more aggressive trading behavior by any one analyst. The impact of an increase in σ_δ on trading aggressiveness includes a third effect; namely, a reduction in the endogenous precision of the informed analysts' signals which tends to reduce trading aggressiveness. This effect together with the entry of new analysts is so strong that we find that typically trading aggressiveness β decreases in σ_δ.

4.4 Price informativeness

We are also interested in the informativeness of the stock price. Let us define Σ as the payoff variance remaining after trade occurs, i.e., $\Sigma = \mathrm{Var}(\delta \mid \text{price})$. Then it can be shown that:

$$\Sigma = \frac{\sigma_\delta^2\left(\sigma_\delta^2 + 2\tau^{-2}\right)}{(N+1)\sigma_\delta^2 + 2\tau^{-2}} = \frac{2+\alpha}{2(N+1) - 2N\alpha},$$

where lower Σ implies a more informative stock price. Note that our bound on information quality due to $\alpha \geq 1$ implies that with one informed analyst, $\Sigma \geq \frac{3}{4}\sigma_\delta^2$, while in the original Kyle model (the insider knows everything), $\Sigma = \frac{1}{2}\sigma_\delta^2$.

We can prove that Σ is increasing in the level of the cost function, k. Increasing k does not change signal quality, but always decreases the number of informed agents, which results in less informative prices. The intuition is simple – more expensive information results in less informative prices. Again, the effect of the convexity parameter, α, is ambiguous because the effect of α on the number of informed analysts and trading aggressiveness is ambiguous (see above). Typically, however, price informativeness is increasing in α except when payoff uncertainty is relatively small; specifically,

$$\sigma_\delta^2 \leq \exp\left[\frac{3 - 2\log\frac{\alpha}{2-\alpha} - \alpha\log\frac{\alpha}{2-\alpha}}{2+\alpha}\right].$$

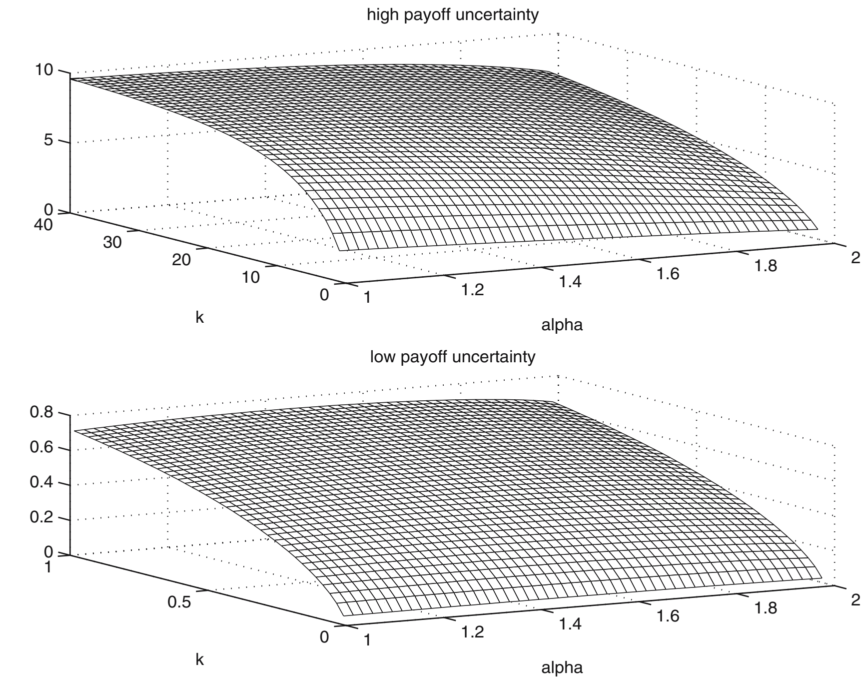

Fig. 1 Remaining uncertainty after trading Σ.This figure illustrates how the parameters α and k of the cost function influence price informativeness as measured by Σ, the remaining payoff uncertainty after trading. The higher Σ, the less informative the price. The *top* figure shows the typical case for high payoff uncertainty $\sigma_\delta^2 > 2.71$: prices are more informative the lower the k and the higher the α parameter of the cost function. Intuitively: lower cost and strongly decreasing returns-to-scale (high number of analysts N, low research intensity τ) lead to more informative prices. The parameters here are $\sigma_\delta^2 = 10$ and $\sigma_z^2 = 10$. The *bottom* figure depicts the case of $\sigma_\delta^2 < 2.71$. Lower cost of information k still leads to more informative prices, but the effect of α is now mildly hump-shaped (see top of figure). Price informativeness is highest (i.e., Σ is lowest) when α is at either one of its extremes 1 and 2. The parameters here are $\sigma_\delta^2 = 1$ and $\sigma_z^2 = 10$

We will provide intuition about this finding in a separate subsection below. Figure 1 illustrates how the parameters of the cost function affect the payoff uncertainty remaining after trade Σ. Figures 3 and 4 focus on the effect of the α parameter alone.

An increase in liquidity trading, σ_z, leads to lower Σ and thus to more informative prices. This may seem counterintuitive because an increase in liquidity trading decreases the informativeness of prices, all else equal. However, the presence of more liquidity traders yields better profit opportunities which encourages both entry and more aggressive trading by analysts. This latter effect dominates the former as in most Kyle-type models. Perhaps surprisingly, given our previous intuition, our numerical results demonstrate that an increase in the initial payoff uncertainty, σ_δ, tends to decrease the price informativeness after trade which is again consistent with most Kyle-type models. Our numerical results demonstrate that the speed at which Σ increases, however, can get arbitrarily small. Furthermore, the fraction of the initial payoff uncertainty which is not revealed by trade, Σ/σ_δ^2, is monotonically increasing and approaches zero when σ_δ gets large.

4.5 Liquidity

The parameter λ measures the price impact per unit of trading volume (inverse of market depth) and can be interpreted as a measure of illiquidity. In addition, λ also captures the cost of uninformed trading. The expected trading cost of the uninformed traders is found to be $E\left[z\left(p-(\bar{F}+\delta)\right) \mid z\right]=\lambda z^2$. Ex-ante (unconditionally), their trading cost is $\lambda\sigma_z^2$.

The comparative statics of λ display a rich behavior. Although ex post price informativeness is monotone in the level of the cost function, k, this will not necessarily be true for liquidity. In fact, it can be shown that illiquidity is always hump-shaped in the level parameter k of the cost function. We know from above that lowering k increases the number of analysts N. This has two effects. On one hand, the presence of more analysts also leads to increased competition among them, which will tend to increase liquidity. On the other hand, since individual signals are not identical, the entry of new analysts also changes the total amount of information collected by the analysts together. This leads to increased informational asymmetry, which tends to decrease liquidity. When information is comparatively cheap (k is low), the competition effect dominates, and liquidity and price informativeness move in the same direction (λ is decreasing in k). For large values of k, however, there are only few informed analysts in the market, and the second effect dominates (λ is increasing in k). Thus, less informative markets can actually be the more liquid ones. For comparative statics with respect to the other model parameters, we have to rely on numerical examples.

With respect to the convexity parameter α, different cases are possible. When payoff uncertainty, σ_δ, is large, λ will be monotonic in α, but it can be either increasing or decreasing in α depending on the parameter k. In particular, it will be increasing in α for large values of k and decreasing for small values of k. Again, this result follows from the relative strength of the competition and the information aggregation effects. When payoff uncertainty is small, however, the effect of α on Σ changes sign at

$$\sigma_\delta^2=\exp\left[\frac{3-2\log\frac{\alpha}{2-\alpha}-\alpha\log\frac{\alpha}{2-\alpha}}{2+\alpha}\right].$$

At this point, the effect of α on λ will also change sign. As a result, λ will display hump-shaped behavior for small levels of k, and bow-shape for k sufficiently large. Figure 2 illustrates how the parameters of the cost function affect market illiquidity λ. Figures 3 and 4 focus on the effect of the α parameter alone.

Numerical examples also show that higher payoff uncertainty, σ_δ, will result in less liquid markets (higher λ). An increase in the payoff uncertainty leads to more profitable trading opportunities, resulting in the presence of more informed analysts, each one of whom will collect less precise information and trade less aggressively. In sum, the confluence of these effects will always lead to less liquid markets (just as we argued above that increased payoff uncertainty tends to decrease

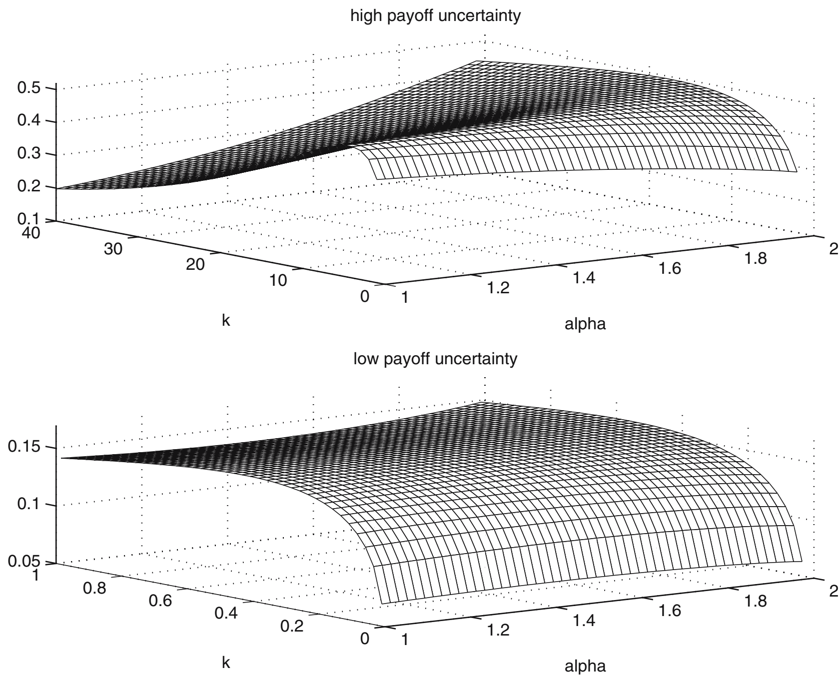

Fig. 2 Market illiquidity λ. This figure illustrates how the parameters α and k of the cost function influence market liquidity, as measured by λ, the price impact of trading an additional share. The higher λ, the less liquid is the market. The *top* figure shows the case of high payoff uncertainty $\sigma_\delta^2 > 2.71$: λ is hump-shaped in the level k of the cost function. Furthermore, λ is decreasing in α for low levels of k, while it is increasing in α for k high. The parameters are $\sigma_\delta^2 = 10$ and $\sigma_z^2 = 10$. The *bottom* figure depicts the case of low payoff uncertainty $\sigma_\delta^2 < 2.71$: again, λ is hump-shaped in the level k of the cost function. But the effect of α again gets more complex: now λ is mildly hump-shaped in α for k small and λ is bow-shaped in α for k large. The parameters are $\sigma_\delta^2 = 1$ and $\sigma_z^2 = 10$

ex post payoff uncertainty). An increase of liquidity trading, σ_z, results in the presence of more informed analysts and also to more aggressive trading behavior which will always yield more liquid markets.

4.6 The cost structure of information: examples and intuition

In a model with endogenous entry alone, such as Admati and Pfleiderer (1988), a decrease in the cost of information always increases the number of informed agents, while holding everything else constant. This will unambiguously result in higher price informativeness. Below we present examples that contradict the naive intuition about information acquisition.

As we have seen, the effect of α on price informativeness is non-monotonic. This can have the following surprising effect: an increase in α can lower the cost of information, and at the same time result in prices becoming less informative.

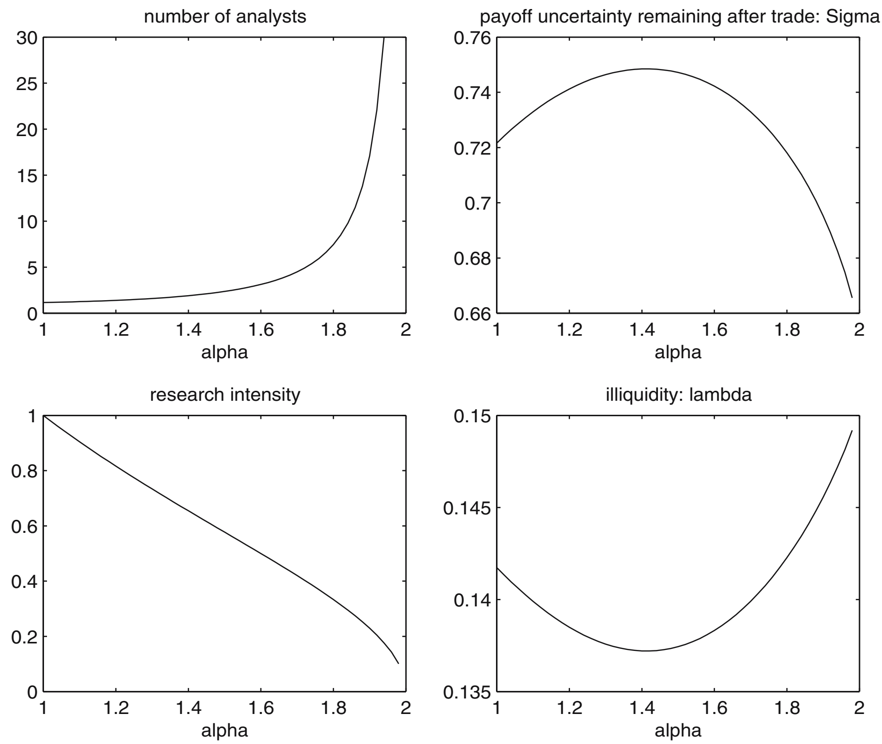

Fig. 3 The effect of α with small payoff uncertainty – k large. This figure illustrates how α affects the equilibrium. α is the exponent of the cost function, and thus determines returns to scale. The higher α, the more favorable it becomes for many agents to become analysts (*top left*), but to collect less information (*bottom left*). This results in a and hump-shaped pattern for the payoff uncertainty remaining after trade (thus a bow-shaped behavior of price informativeness, see *top right*). In this example (k large) I find a bow-shaped pattern for illiquidity λ (*bottom right*). The parameters of the model are $\sigma_\delta^2 = 1$, $\sigma_z^2 = 1$, $k = 1$

This will happen whenever

$$1 \leq \sigma_\delta^2 \leq \exp\left[\frac{3 - 2\log\frac{\alpha}{2-\alpha} - \alpha\log\frac{\alpha}{2-\alpha}}{2+\alpha}\right].$$

For example, this will always be true for $\alpha = 1$ and σ_δ^2 between 1 and 2.71.[9]

Example 1 Figure 3 illustrates the effect of λ on various model parameters. Here, we find that even though an increase in α makes information cheaper, it also changes equilibrium research intensities, which will make asset prices less informative. At the same time, however, it affects liquidity in the opposite way. Even though informational asymmetry increases, markets become more liquid, and the amount of resources spent on information acquisition decreases.

[9]Figures 1 and 2 show how price informativeness and market liquidity depend on both parameters of the cost function, α and k.

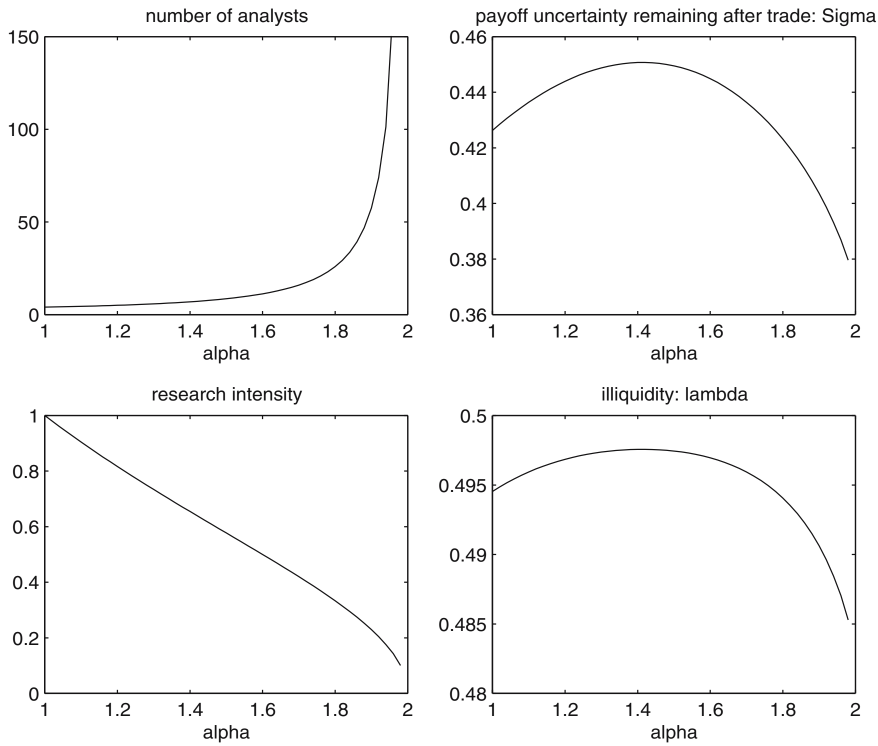

Fig. 4 The effect of α with small payoff uncertainty – k small. This figure illustrates how α affects the equilibrium. α is the exponent of the cost function, and thus determines returns to scale. The higher α, the more favorable it becomes for many agents to become analysts (*top left*), but to collect less information (*bottom left*). This results in a and hump-shaped pattern for the payoff uncertainty remaining after trade (thus a bow-shaped behavior of price informativeness, see *top right*. In contrast to Fig. 3, we have a smaller k, which results in a hump-shaped pattern for illiquidity λ (*bottom right*). So for small values of α, even though an increase in α makes information cheaper, it has only negative consequences: it results in the stock price being less informative and markets being less liquid at the same time. The parameters of the model are $\sigma_\delta^2 = 1, \sigma_z^2 = 1, k = 0.1$

Example 2 Figure 4 illustrates a "worst-case" scenario. Even though raising α results in lower cost of information, it also changes equilibrium research intensities, which leads to more analysts collecting less precise information. In this example, this has unambiguously negative consequences: prices become less informative, while markets become less liquid at the same time, and the amount of resources spent on information acquisition increases!

The non-monotonic effect of α on price informativeness is again the result of the interplay of competition and information acquisition. Imperfect competition implies that, all else equal, the same amount of information is incorporated into prices more efficiently if many analysts each possess a little piece of the information rather than one insider knowing everything. But in the above parameter region, an increase in α will always lead to the presence of more analysts. Consequently, the non-monotonicity of price informativeness can only be explained by information acquisition. In fact, for low payoff uncertainty, σ_δ, we find that total information acquisition is non-monotonic in α. In particular, an increase in α can lead to a decrease in total information acquisition that is strong enough to offset the increase in competition.

The reason for the non-monotonicity in information acquisition is as follows. A change in α impacts both total cost $c(.)$ and marginal cost $c'(.)$, which both affect information acquisition. While the profitability of informed trading and thus the number of analysts depend on total cost $c(.)$, marginal cost $c'(.)$ determines the optimal signal precision. Interestingly, a change in α can have opposite effects. For small values of payoff uncertainty, an increase in α increases the marginal cost of information, while at the same time decreasing total cost $c(.)$. The increase in marginal cost induces informed analysts to acquire less precise information, and this will lead to a decrease in total information production.

Consequently, there exist two ways of optimally incorporating information into prices. Typically, a large α is efficient. In this case, many analysts with widely dispersed beliefs will lead to the highest price informativeness. If payoff uncertainty is small, however, a low α will lead to more informative prices, meaning that it is optimal to have relatively few analysts who each generate relatively precise information. Which one will be the case is exclusively determined by the comparison of payoff uncertainty σ_δ^2 with a function of the convexity parameter, α.

5 Optimal disclosure

This section explores in more depth the relationship between the cost of information and the properties of the resulting asset market, i.e., informational efficiency and market liquidity. For this purpose, we will endogenize the cost function by allowing the firm's manager to choose an optimal informational policy toward the shareholders.[10] In general, different informational policies will require different amounts of the firm's resources for collecting, processing and presenting the information. For the sake of simplicity we will abstract from this issue and simply assume that the firm can freely select parameters from a certain range, i.e., choose a $k^* \in \left[\underline{k}, \bar{k}\right]$ and an $\alpha^* \in \left[\underline{\alpha}, \bar{\alpha}\right]$.

Following Fishman and Hagerty (1989), Holmstrom and Tirole (1993), and Dow and Gorton (1997), we assume that price discovery results in economic benefits B that are increasing in informational efficiency. These benefits are the payoff from informed analysts' research expenditures, $\lambda\sigma_z^2$. This is the expected loss of liquidity traders to the informed analyst. To induce the liquidity traders to hold the stock, the firm has to compensate them for their expected loss. Thus, the manager's problem is to choose parameters of the cost function as to maximize

$$B - \lambda\sigma_z^2.$$

Non-monotonicities in the comparative statics for informational efficiency and liquidity result in a multitude of optimal policies. Furthermore, we find that small

[10] Our approach is in the spirit of Fishman and Hagerty (1989). This is in contrast to a large part of the literature on information disclosure, going back to Diamond (1985), which assumes that managers can disclose information about the firm's value credibly and costlessly, and that all market participants interpret this information correctly.

changes in the firm's characteristic can change the optimal informational policy completely.

First, consider the parameter k that determines the overall level of the cost function. To obtain maximum informational efficiency, the manager will find it optimal to set k as low as possible, i.e., $k^* = \underline{k}$. This will maximize economic benefits B. But on the other hand, social cost $\lambda\sigma_z^2$ is hump-shaped in k. One way to ensure liquid markets is to make information as cheap as possible to attract competition among analysts, and also maximize the economic benefits from informational efficiency. But in other cases, it will be optimal to make information prohibitively expensive, i.e., set k^* close to $\bar{k}$. This allows the manager to protect liquidity traders by deterring informational trade, which may well be worth foregoing the benefits from informational efficiency. The possibility of different equilibria leads to discontinuities: a small change in parameters may result in a significant change in the informational policy, which can shift from one end of the parameter range to the other.

We know that α can have a non-monotonic effect on Σ. Consequently, benefits B from informational efficiency are maximized by either setting $\alpha^* = \bar{\alpha}$ or $\alpha^* = \underline{\alpha}$. A high α will result in many analysts each doing relatively little, resulting in stiff competition among analysts which translates into highly informative prices. On the other hand, it can be optimal to set α low, resulting in few analysts each doing careful research. The optimal research intensity will be so high as to overcome the disadvantage of low competition. Again, a small change in the firm's characteristics can be enough to switch from $\underline{\alpha}$ to $\bar{\alpha}$. Furthermore, interior equilibria are possible in situations where a change in α has opposite effects on informational efficiency and liquidity, allowing even more possible optimal policies.

6 Empirical implications

6.1 Determinants of analyst coverage

The model yields a number of predictions about the determinants of analyst coverage in the cross-section. For example, high payoff uncertainty and high noise trading will all lead to greater analyst coverage in equilibrium. Furthermore, small differences in form characteristics may result in widely different informational policies. In many cases, the optimal policy results in many analysts all acquiring widely dispersed signals. Thus, the model gives a very simple explanation on why some stocks are followed by many analysts, without resorting to arguments such as herding (Welch 2000), time patterns in information arrival (Hirshleifer et al. 1994), or behavioral reasons.

High values of liquidity trading, σ_z, and payoff uncertainty, σ_δ, will both result in a high number of analysts N. Consistent with our model, Brennan and Subrahmanyam (1995) find a significant positive effect of return variance on analyst following. We also predict that a low level of the cost function (k low) and strongly decreasing returns to scale (α high) will both lead to increased analyst coverage. However, since the cost function is unobservable, we cannot test these implications directly.

The multitude of optimal informational policies can explain strong cross-sectional variation in analyst coverage. Among similar firms, one may find it optimal to make information cheap to attract analysts and thus achieve high informational efficiency, the other may find it optimal to make information prohibitively expensive to deter informed trade. Moreover, among otherwise similar firms, one may find it favorable if few analysts each collect precise information, while the other may find it favorable if many analysts each know only a little bit.

Additional insights could potentially be obtained by studying of a multi asset setting along the lines of Caballe and Krishnan (1994). This would provide liquidity traders to choose which stocks to trade in. To attract liquidity traders, the manager will have an additional incentive to keep liquidity high when setting the informational policy of the firm. Also, correlations among asset payoffs may induce spillovers from one firm to the other. However, due to the added complexity, we will have to leave this direction to future research.

6.2 Trading volume

The expression for trading volume is not completely straightforward in Kyle-type models. Imagine that there are two traders, submitting orders of 10 and −16, respectively. Then total trading volume (number of shares trading hands) will be 16: ten shares are purchased by trader 1 from trader 2 and six shares are purchased by the market maker from trader 2. So trading volume equals the sum of all buys or the sum of all sales, whichever is larger which is the same as the sum of the absolute values of all orders plus the absolute value of the order imbalance (divided by two to avoid double counting). Thus, expected trading volume EV is found to be

$$\mathrm{EV} = \frac{1}{2}\sum_{i=1}^{N} E|x_i| + \frac{1}{2}E|z| + \frac{1}{2}E\left|\sum_{i=1}^{N} x_i + z\right|$$

The expected value of the absolute of a normal random variable equals its standard deviation times $\sqrt{2/\pi}$, so expected volume can be calculated as:

$$\mathrm{EV} = \frac{1}{\sqrt{2\pi}}\left(N\sigma(x_i) + \sigma_z + \sigma\left(\sum_{i=1}^{N} x_i + z\right)\right)$$

which yields

$$\mathrm{EV} = \frac{1}{\sqrt{2\pi}}\left(\frac{\sqrt{N}\sigma_\delta \tau^{-1}}{\sqrt{\sigma_\delta^2 + \tau^{-2}}} + \sigma_z + \sqrt{\frac{\sigma_\delta^2 \tau^{-2}}{\sigma_\delta^2 + \tau^{-2}} + \sigma_z^2}\right).$$

By inspection, trading volume is higher when the research intensity, τ, is lower and when the number of analysts, N, is large. The latter finding is consistent with Bhushan (1989), who reports a positive and significant effect of trading volume on N. Our result that signal precision will always be less than the payoff uncertainty is not directly testable, however, it does suggest high levels of trading volume because low signal precision results in greater analyst entry and widely dispersed beliefs, which both increase the volume of informational trade. This is an interesting finding, since the huge trading volume observed in reality is still a puzzle to explain theoretically.

7 Robustness

In this section, we want to explore the sensitivity of our results to the model assumptions, in particular, with respect to the cost structure for information.

So far, we have restricted ourselves to cost functions of the form $c(\tau) = k\tau^{\alpha}$. Holding the number of analysts, N, fixed, we can obtain the optimality condition for the research intensity τ for a general cost function $c(.)$, which is found to be:

$$c'(\tau) = \frac{\sigma_\delta^2 \sigma_z}{\sqrt{N}\sqrt{\sigma_\delta^2 + \tau^{-2}}\left((N+1)\sigma_\delta^2 + 2\tau^{-2}\right)} \cdot \frac{2}{\tau^3}.$$

When we proceed to endogenize entry, we get another equilibrium condition, namely that gains from informed trading equal the cost of information:

$$c(\tau) = k\tau^{\alpha} = \frac{\sigma_\delta^2 \sigma_z \sqrt{\sigma_\delta^2 + \tau^{-2}}}{\tau^3 \sqrt{N}\left((N+1)\sigma_\delta^2 + 2\tau^{-2}\right)}.$$

In equilibrium, both these conditions have to hold. Note that due to the similarity of the two expressions, combining them yields a very simple expression for the equilibrium research intensity:

$$\frac{c'(\tau)}{c(\tau)} = \frac{2}{\sigma_\delta^2 + \tau^{-2}}.$$

Note that the resulting equation is independent of the extent of liquidity trading, σ_z. Also, any general cost function $c(.)$ will lead to the same optimal research intensity as a rescaled cost function $kc(.)$ for any $k > 0$. Only for issues of existence and uniqueness and for the sake of a closed form solution do we need to restrict ourselves to cost functions of the power type.

We could also include a fixed cost l in the cost of information. The optimal research intensity would then be given by

$$\frac{c'(\tau)}{c(\tau) + l} = \frac{2}{\sigma_\delta^2 + \tau^{-2}}.$$

Including a small fixed cost would not change our results much. However, as the fixed cost gets larger, it counteracts the importance of endogenous precision, while putting more emphasis on endogenous entry. In the limit, we would obtain predictions similar to Admati and Pfleiderer (1988). Note that in the absence of a fixed cost, analysts could acquire just a tiny bit of information. In equilibrium, however, this will be a suboptimal strategy, since analysts with very little information are themselves exposed to the superior information of better-informed analysts.

Also, it is possible to generalize the model to allow for asymmetric equilibria. In this case, it will be possible that some analysts will acquire just very little information in equilibrium. The presence of many interacting economic effects in our model may lead to even more interesting new implications, which we leave for future research.

8 Conclusion

This paper studies information acquisition in a model where both entry of analysts and their optimal information quality is endogenous. We show existence of the Bayesian–Nash equilibrium and solve for it in closed form. The model displays a very rich behavior. In particular, we find that the precision of a individual signal will always be bounded from above by the precision of the prior belief on payoff uncertainty. Furthermore, we give examples that contradict the naive intuition about information acquisition. For instance, we show how a change in the cost structure that makes information cheaper decreases price informativeness, while at the same time market liquidity and the amount of resources society spends on information acquisition can change either way (depending on the relative strength of changes in information acquisition and competition). The model gives a simple, fully rational explanation on why the number of analysts following a stock can be quite large. The model allows for interesting extensions in future research, for instance an analysis of asymmetric equilibria.

Appendix: proof of proposition 1

The proof proceeds in three steps. First, taken the number of informed analysts N as given, and assuming that an equilibrium exists, we calculate the market maker pricing rule $p(y)$, informed analysts' trading strategies x_i, and an optimality condition for the informed analysts' research intensity τ_i. Note that we do not need any specific assumptions about the cost function $c(.)$ here.

In the second step, still taking the number of informed analysts N as given, we show that for a cost function of the type $c(\tau) = k \cdot \tau^{\alpha}$, an equilibrium exists. This is shown to be the unique equilibrium with symmetric behavior of informed analysts, linear trading strategies x_i and a linear market maker pricing scheme $p(y)$.

The third lemma finally endogenizes the entry decision, i.e., shows existence of a symmetric Bayesian–Nash equilibrium as defined in section 3. This equilibrium is solved for in closed for by solving a third-order polynomial for the number of informed analysts N.

Step 1 (Equilibrium with fixed N**)** *Assume that for a cost function* $c(.)$*, a symmetric linear equilibrium exists. Then it is described by a market maker pricing scheme* $p(y) = \lambda \cdot y$ *given by*

$$\lambda = \frac{\sigma_\delta^2}{(N+1)\sigma_\delta^2 + 2\tau^{-2}} \frac{\sqrt{N(\sigma_\delta^2 + \tau^{-2})}}{\sigma_z}$$

informed analysts' trading strategies $x_i = \beta \cdot s_i$ *given by*

$$\beta = \frac{\sigma_\delta^2}{\lambda\left((N+1)\sigma_\delta^2 + 2\tau^{-2}\right)} = \frac{\sigma_z}{\sqrt{N(\sigma_\delta^2 + \tau^{-2})}}$$

and informed analysts' research intensities $\tau_i = \tau = 1/\sigma_\epsilon$, where τ is given by the optimality condition

$$c'(\tau) = \frac{\sigma_\delta^2 \sigma_z}{\sqrt{N}\sqrt{\sigma_\delta^2 + \tau^{-2}}\left((N+1)\sigma_\delta^2 + 2\tau^{-2}\right)} \cdot \frac{2}{\tau^3}.$$

The proof proceeds as follows. First we conjecture the existence of a symmetric linear equilibrium. Then, given that an informed analyst i has a research intensity τ_i, we can calculate his optimal trading strategy, which is found to be linear in the individual signal. Next we can determine the first order condition for the optimal choice of his research intensity. We can use the symmetry of the equilibrium, to compute the informed trading strategy. Finally, we can calculate the market maker's pricing scheme from his zero-profit condition.

Proof of Step 1 Assume that all informed analysts besides i use a linear trading strategy of the form $x_j = \beta_j \cdot s_j$ (and has a research intensity τ_j), and that the market maker uses a linear pricing scheme $p(y) = \lambda \cdot y$. Informed analyst i jointly chooses market orders x_i and research intensity τ_i to maximize his expected trading profits. First, we will find the optimal trading strategy $x_i(\tau_i)$ as a function of the research intensity. Given τ_i, agent i's optimization problem is:

$$\max_{x_i} E\left[x_i\left(\delta + \theta - \lambda\left(x_i + \sum_{j \neq i} \beta_j s_j + z\right)\right) \mid s_i, \tau_i\right].$$

Plug in $s_j = \delta + \epsilon_j$:

$$\max_{x_i} E\left[x_i\left(\delta + \theta - \lambda\left(x_i + \sum_{j \neq i} \beta_j \delta + \sum_{j \neq i} \beta_j \epsilon_j + z\right)\right) \mid \delta + \epsilon_i, \tau_i\right].$$

Now we can use that the ϵ_j's, θ and z are independent of ϵ_i and have mean zero, so that they drop out when calculating the expectations. The problem simplifies to

$$\max_{x_i}\left\{-\lambda x_i^2 + x_i \cdot E\left[\delta \mid \delta + \epsilon_i, \tau_i\right] \cdot \left[1 - \lambda \sum_{j \neq i} \beta_j\right]\right\}.$$

Differentiating with respect to x_i gives us the first-order condition:

$$2\lambda x_i = E\left[\delta \mid \delta + \epsilon_i\right] \cdot \left[1 - \lambda \sum_{j \neq i} \beta_j\right].$$

We calculate $E\left[\delta \mid \delta + \epsilon_i, \tau_i\right] = \frac{\sigma_\delta^2}{\sigma_\delta^2 + \tau_i^{-2}}(\delta + \epsilon_i)$ and find that x_i is linear in s_i, so we can write it in the form $x_i = \beta_i(\delta + \epsilon_i)$:

$$x_i = \beta_i\left(\delta + \epsilon_i\right) = \frac{\sigma_\delta^2}{\sigma_\delta^2 + \tau_i^{-2}} \frac{1 - \lambda \sum_{j \neq i} \beta_j}{2\lambda}\left(\delta + \epsilon_i\right).$$

Using the above optimal trading strategy, insider i then chooses the optimal research intensity τ_i taking λ and $\bar{\beta}$ as given. He maximizes:

$$\max_{\tau_i} E\left[x_i\left(\delta-\lambda\left(x_i+\sum_{j\neq i}\beta_j\delta-\sum_{j\neq i}\beta_j\epsilon_j-z\right)\right)\right]-c(\tau_i). \tag{1}$$

The terms involving ϵ_j and z all drop out, since variables are uncorrelated and have zero expectation. Plugging in the expression for x_i, we obtain i's problem as:

$$\max_{\tau_i} E\left[-\frac{\sigma_\delta^4}{(\sigma_\delta^2+\tau_i^{-2})^2}\frac{\left(1-\lambda\sum_{j\neq i}\beta_j\right)^2}{4\lambda}(\delta+\epsilon_i)^2\right.$$
$$\left.+\frac{\sigma_\delta^2}{\sigma_\delta^2+\tau_i^{-2}}\frac{\left(1-\lambda\sum_{j\neq i}\beta_j\right)^2}{2\lambda}\delta(\delta+\epsilon_i)\right]-c(\tau_i).$$

Now we can calculate expectations using $E[\epsilon_i]=\tau_i^{-2}$ and simplify. We obtain:

$$\max_{\tau_i}\frac{\sigma_\delta^4}{\sigma_\delta^2+\tau_i^{-2}}\cdot\frac{\left(1-\lambda\sum_{j\neq i}\beta_j\right)^2}{4\lambda}-c(\tau_i).$$

Next, we can differentiate with respect to τ_i to obtain the final optimality condition:

$$c'(\tau_i)=\frac{2\sigma_\delta^4}{\tau_i^3(\sigma_\delta^2+\tau_i^{-2})^2}\cdot\frac{\left(1-\lambda\sum_{j\neq i}\beta_j\right)^2}{4\lambda}.$$

We are looking for a symmetric equilibrium, so we plug $\beta=\bar{\beta}$ and $\tau_i=\tau\ \forall\, i$ into the expression for trading aggressiveness β to obtain:

$$\beta=\frac{\sigma_\delta^2}{\lambda\left((N+1)\sigma_\delta^2+2\tau^{-2}\right)}.$$

Next, we can compute the value of λ that sets market maker profits to zero. As usual, that implies that we can obtain λ from regressing δ on the total order flow $y=\sum_{i=1}^N x_i+z$, which yields

$$\delta=\frac{\operatorname{Cov}\left(\sum_{i=1}^N x_i+z,\delta\right)}{\operatorname{Var}\left(\sum_{i=1}^N x_i+z\right)}$$

substituting the expression for x_i and simplifying yields

$$\lambda=\frac{\sigma_\delta^2}{(N+1)\sigma_\delta^2+2\tau^{-2}}\frac{\sqrt{N\left(\sigma_\delta^2+\tau^{-2}\right)}}{\sigma_z^2},$$

which finally yields the first order condition:

$$c'(\tau) = \frac{2\sigma_\delta^4}{\tau^3\left(\sigma_\delta^2+\tau^{-2}\right)^2}\cdot\frac{\left(1-\frac{\sigma_\delta^2(N-1)}{(N+1)\sigma_\delta^2+2\tau^{-2}}\right)^2}{\frac{4\sigma_\delta^2}{(N+1)\sigma_\delta^2+2\tau^{-2}}\frac{\sqrt{N(\sigma_\delta^2+\tau^{-2})}}{\sigma_z}}$$
$$= \frac{2\sigma_\delta^4}{\tau^3(\sigma_\delta^2+\tau^{-2})^2}\cdot\frac{\left(\frac{2\left(\sigma_\delta^2+\tau^{-2}\right)}{(N+1)\sigma_\delta^2+2\tau^{-2}}\right)^2}{\frac{4\sigma_\delta^2}{(N+1)\sigma_\delta^2+2\tau^{-2}}\frac{\sqrt{N\left(\sigma_\delta^2+\tau^{-2}\right)}}{\sigma_z}}.$$

Now we can simplify and obtain:

$$c'(\tau) = \frac{2\sigma_\delta^2\sigma_z}{\tau^3\sqrt{N}\sqrt{\sigma_\delta^2+\tau^{-2}}\left((N+1)\sigma_\delta^2+2\tau^{-2}\right)}.$$

□

Existence and uniqueness of an equilibrium for general types of cost functions is not completely straightforward. For example, we will later see that for $\alpha = 1$ the first order condition does not have a solution unless

$$k \leq \frac{\sigma_\delta^2\sigma_z}{N}.$$

In addition, we have to check that informed agents trading profits are big enough to cover their cost of information acquisition. So we cannot simply appeal to convex analysis arguments in general. However, if we restrict ourselves to cost functions of the type $c(\tau) = k\cdot\tau^\alpha$, we get the following result:

Step 2 (Existence and uniqueness with N **fixed)** *Depending on α, we have two different cases:*

- *For* $\alpha > 2$*, there exists a unique symmetric linear Nash equilibrium with endogenous information acquisition for every possible* N*.*
- *For* $\alpha \in [1, 2)$*, there exists a number* N^{max} *such that for all* $N \leq N^{max}$ *there exists a unique symmetric linear Nash equilibrium with endogenous information acquisition.*

Proof of Step 2 Let us first consider the case of $\alpha > 1$. We know from Step 1 that the first order condition for the analysts is:

$$c'(\tau) = \alpha k\tau^{\alpha-1} = \frac{2\sigma_\delta^2\sigma_z}{\tau^3\sqrt{N}\sqrt{\sigma_\delta^2+\tau^{-2}}\left((N+1)\sigma_\delta^2+2\tau^{-2}\right)}.$$

Let us rewrite it as

$$c'(\tau) = \alpha k\tau^{\alpha-1} = \frac{2\sigma_\delta^2\sigma_z}{\sqrt{N}\sqrt{\tau^2\sigma_\delta^2+1}\left((N+1)\tau^2\sigma_\delta^2+2\right)}.$$

The term on the left-hand side is of the order of $\tau^{\alpha-1}$, and goes to infinity for large τ. The term on the right-hand side, however, will approach zero as τ goes to infinity. So

$$g(\tau) = -\alpha k \tau^{\alpha-1} + \frac{2\sigma_\delta^2\sigma_z}{\sqrt{N}\sqrt{\tau^2\sigma_\delta^2+1}\left((N+1)\tau^2\sigma_\delta^2+2\right)}$$

will be *negative* for τ sufficiently large. For small τ, the left hand term goes to zero, while the right-hand term converges to a positive constant:

$$\lim_{\tau\to 0} \frac{2\sigma_\delta^2\sigma_z}{\sqrt{N}\sqrt{\tau^2\sigma_\delta^2+1}\left((N+1)\tau^2\sigma_\delta^2+2\right)} = \frac{\sigma_\delta^2\sigma_z}{\sqrt{N}}.$$

This guarantees the existence of (at least one) tangency ($g(\tau^*) = 0$) for $\alpha > 1$.

For $\alpha = 1$, $g(\tau)$ is a third order polynomial in τ^{-2}:

$$\begin{aligned}&\left(4Nk^2 - 4\sigma_\delta^2\sigma_z^2\right)\cdot\tau^{-6} + 4Nk^2(N+2)\sigma_\delta^2\cdot\tau^{-4}\\ &\quad + Nk^2(N+1)(N+5)\sigma_\delta 42\cdot\tau^{-2} + k^2N(N+1)^2\sigma_\delta^6.\end{aligned}$$

By Descartes' rule, $g(.)$ has a unique positive solution if and only if $N < \frac{\sigma_\delta^4\sigma_z^2}{k^2}$. So there is a unique tangency for $N < N^{\max} = \frac{\sigma_\delta^4\sigma_z^2}{k^2}$.

Also, in equilibrium, the informed analysts must make enough trading profits to recover their research expenses. We calculate the expected revenue $\pi_{\inf}$ of informed analyst trading, which is found to be

$$\begin{aligned}\pi_{\inf} &= E\left[x\left(\delta - \lambda N\beta\delta - \lambda\beta\sum_{i=1}^{N}\epsilon_i - \lambda z\right)\right]\\ &= \sigma_\delta^2\beta(1-N\lambda\beta) - \lambda\beta^2\tau^{-2}\\ &= \frac{\sigma_\delta^2\sigma_z\sqrt{N(\sigma_\delta^2+\tau^{-2})}}{N\left[(N+1)\sigma_\delta^2+2\tau^{-2}\right]}\end{aligned}$$

for the optimal research intensity τ, we equate the cost for and revenue of becoming informed:

$$k\tau^\alpha = \frac{\sigma_\delta^2\sigma_z\sqrt{\sigma_\delta^2+\tau^{-2}}}{\tau^3\sqrt{N}\left((N+1)\sigma_\delta^2+2\tau^{-2}\right)}.$$

Now we can take advantage of the similarity of the first order condition with individual profits. Substituting in the first order condition yields that equilibrium profits are positive if and only if

$$\tau \geq \sqrt{\frac{2-\alpha}{\alpha}\frac{1}{\sigma_\delta}}.$$

This always holds for $\alpha \geq 2$, but will always become binding for large N when $\alpha < 2$. As a consequence, there exist a maximum number $N^{\max}$ that we can support for $1 \leq \alpha < 2$, while we can support any N for $\alpha \geq 2$.

The positive profit condition is stronger than the second order condition, which can be shown to be:

$$\tau \geq \sqrt{\frac{2-\alpha}{2+\alpha}}\frac{1}{\sigma_\delta}.$$

This implies that the above equilibrium is unique. If there were in fact two equilibria (assume $g(\tau_1) = g(\tau_2) = 0$). Then g has strict local maxima in both τ_1 and τ_2. Consequently, there has to be an τ^* such that g has a local minimum at τ^*. This contradicts the second order condition. □

Proof of main proposition We know from the proof of step 2 that for $\alpha \in [1, 2)$ there is a maximum error variance that can be supported in equilibrium. So in the absence of fixed costs, agents will become informed until either there are no more agents that have access to the information technology or until profits equal costs, i.e., the maximum error variance is reached:

$$\sigma_\epsilon^2 = \tau^{-2} = \frac{\alpha}{2-\alpha}\sigma_\delta^2.$$

Then we can use the above expression for τ together with the fact that profits equal costs:

$$k\tau^\alpha = \frac{\sigma_\delta^2 \sigma_z \sqrt{\sigma_\delta^2 + \tau^{-2}}}{\sqrt{N}\left((N+1)\sigma_\delta^2 + 2\tau^{-2}\right)}.$$

Multiplying this out gives us a third order polynomial in N, in which the coefficient of the square term is zero. This guarantees a tractable solution to the equation.

$$\sigma_\epsilon^{-\alpha} k\sqrt{N}\left((N+1)\sigma_\delta^2 + 2\sigma_\epsilon^2\right) = \sigma_\delta^2\sigma_z\sqrt{\left(\sigma_\delta^2 + \sigma_\epsilon^2\right)}$$

substituting in $\sigma_\epsilon^2 = \frac{\alpha}{2-\alpha}$:

$$\sigma_\delta^{-\alpha}\left(\frac{\alpha}{2-\alpha}\right)^{\frac{-\alpha}{2}} kN^{\frac{3}{2}}\sigma_\delta^2 + \sigma_\delta^{-\alpha}\left(\frac{\alpha}{2-\alpha}\right)^{\frac{-\alpha}{2}}$$
$$kN^{\frac{1}{2}}\left(1+\frac{2\alpha}{2-\alpha}\right)\sigma_\delta^2 = \sigma_\delta^3\sigma_z\sqrt{\left(1+\frac{\alpha}{2-\alpha}\right)}.$$

Squaring and simplifying yields:

$$\sigma_\delta^{-2\alpha}\left(\frac{\alpha}{2-\alpha}\right)^{-\alpha} k^2N^3 + \sigma_\delta^{-2\alpha}\left(\frac{\alpha}{2-\alpha}\right)^{-\alpha}$$
$$\left(\frac{2+\alpha}{2-\alpha}\right)^2 k^2 N - \sigma_z^2\sigma_\delta^3\left(\frac{2}{2-\alpha}\right) = 0.$$

The solution to the equation is stated in the proposition. □

References

Admati, A., Pfleiderer, P.: A monopolistic market for information. J Econ Theory **39**, 400–438 (1986)
Admati, A., Pfleiderer, P.: A theory of intraday patterns: volume and price variability. Rev Financ Stud **1**, 3–40 (1988)
Bhushan, R.: Firm characteristics and analyst following. J Account Econ **11**, 255–274 (1989)
Brennan, M.J., Subrahmanyam, A.: Investment analysis and price formation in securities markets. J Financ Econ **38**, 361–381 (1995)
Caballe, J.: Transmission and production of information in imperfectly competitive financial markets. J Econ Bus **45**, 409–419 (1993)
Caballe, J., Krishnan, M.: Imperfect competition in a security market with risk neutrality. Econometrica **62**, 695–704 (1994)
Diamond, D. W.: Optimal disclosure of information by firms. J Finance **40**, 1071–1094 (1985)
Diether, K., Malloy, C., Scherbina, A.: Differences of opinion and stock returns. J Finance **57**, 2113–2141 (2002)
Dow, J., Gorton, G.: Stock market efficiency and economic efficiency: is there a connection. J Finance **52**, 1087–1129 (1997)
Fishman, M.J., Hagerty, K.: Disclosure decisions by firms and the competition for price efficiency. J Finance **44**, 633–646 (1989)
Grossman, S., Stiglitz, J.: On the impossibility of informationally efficient markets. Am Econ Rev **70**, 393–408 (1980)
Hellwig, M.: On the aggregation of information in competitive markets. J Econ Theory **22**, 477–498 (1980)
Hirshleifer, D., Subrahmanyam, A., Titman, S.: Security analysis and trading patterns when some investors receive information before others. J Finance **49**, 1665–1698 (1994)
Holden, C., Subrahmanyam, A.: Long-lived private information and imperfect competition. J Finance **47**, 247–270 (1992)
Holmstrom, B., Tirole, J.: Market liquidity and performance monitoring. J Polit Econ **101**, 678–709 (1993)
Kyle, A.S.: Continuous auctions and insider trading. Econometrica **53**, 1315–1335 (1985)
Kyle, A.S.: Informed speculation with imperfect competition. Rev Econ Stud **56**, 317–356 (1989)
Verrecchia, R.E.: Information acquisition in a noisy rational expectations economy. Econometrica **50**, 1415–1430 (1982)
Welch, I.: Herding among security analysts. J Finan Econ **58**, 369–396 (2000)