How to distribute information

by

Espen Sirnes

Department of Economics and Management
Norwegian College of Fishery Science
University of Tromsø
Norway

How to distribute information[1]

Espen Sirnes

Abstract

In this paper a difficult question is answered with a surprisingly simple answer. A monopolist who possesses nested information can earn money from selling it at different levels of precision to investors. The problem is to maximize profits by choosing the optimal distribution of information among the investors. I show that, the optimal distribution is to give all informed investors the same level of precision. The model belying this result is a continuous version of Grossman and Stiglitz (1980).

Keywords: Distribution of information, finance, asset pricing.

JEL Classification: G12, G14

Say you are an insider and you are willing to commit the crime of selling information to some investors for a handsome sum. The sources of information for investors are only official accounts and yourself, so for really precise information you are an information monopolist. Then, how should you distribute more precise and less precise information among investors in order to maximize your profit? This paper investigates this question, and the answer turns out to be very simple; there will be no discriminatory distribution of information among informed investors at all! The profit maximizing strategy for an information monopolist is to supply the same information to every informed investor.

What this paper does is therefore to ask a very complicated question, and provide a surprisingly simple answer. The information monopolist will be allowed to distribute information according to any partially continuous density function over a continuum of investos at no extra "discrimination cost". Even then, the simplest, and probably cheapest, way of distributing information is the optimal one.

The fraction of informed investors will depend on the level of noise trading, quality of the information sold and volatility of the asset. If there is much noise in the prices (high level of noise traders) and the unconditional variance of the traded stock relatively low it will be optimal to supply a relatively large amount of information. This is because a high level of noise relative to unconditional variance describes an opaque market where information is not easily transferred to those not paying for it. Distributing more information thus has a less negative effect on monopolist profits in terms of making the prices more informative.

If the unconditional variance of the asset is high the monopolist will inform a relatively large mass of investors. This is because the monopolist has to balance the number of non-paying uninformed investors to those who receive information and pay. The higher the unconditional variance, the more investors will gain by buying information, and so the more the monopolist

earns per informed investor. This therefore leads to a higher fraction of investors being informed when the unconditional variance is high.

Since the monopolist does not sell all of his information, there is a kind of information "dead weight loss" where useful information is never disclosed. There will be information that investors would be willing to pay for, but releasing it would deplete the value of the information sold.

The example of the insider serves as an instructive illustration of the issues discussed in this paper. The results are however applicable for other issues, such as analysts' incentives to do research and the quality of information that is possible to obtain in an information market. This paper can also be useful for understanding how information is treated in general. The main point of the financial market approach is that the more who have the information the less value it has, but this may be the case for other types of information.

The extent to which this model relates to real insider trading may however also be used by regulators to deal with this particular problem.

In addition one may look at the monopolist as an abstraction used to evaluate the value of information in a market. Within this perspective, the model presented here will also give some insight into how distribution of information in general affects its value.

In this paper the information is nested. That is the information differs between agents in a way such that a well-informed trader has all the information of the less informed, and some more. The effect of this is that the covariance between two signals equals the variance of the more precise signal. The motivation for a nested information structure is that it is a likely structure for information sold by a monopolist. A non-nested structure would resemble an insider who sells different pages to different investors from unpublished account books. A nested model can be illustrated as the insider in addition provides all preceding pages to the buyer as well. In a discussion about the distribution of information, a nested model therefore seems more appropriate than a model with independent information units sold to

different investors.

It is assumed that the monopolist cannot trade himself. There might be different reasons for this, such as legal issues or constraints on liquidity. Other authors have assumed a trading monopolist such as Admati and Pfleiderer (1988) (AP). This does, however, seem to be possible only with strict assumptions on the price setting mechanism. AP use a price setting mechanism of Kyle (1985) where the price is set so that uninformed investors have no incentive to trade at all. Kyle assumes the price is set by a risk neutral market maker, equal to his expectation. This has the implication that the monopolist need not concern himself with the effect seen in Grossman and Stiglitz (1980) (GS), where information sold is partially transferred to uninformed investors through prices, reducing its value. It is therefore reasonable that without this indirect effect on the value of information seen in GS, it might actually be beneficial for a trading monopolist to sell information.

In the framework of Grossman and Stiglitz (1980) there is no market maker and all agents are risk averse. The GS model does not seem to permit an information monopolist that can trade freely. The value of information in the GS model is the same for everyone, and so there should be no incentive to sell it for a trader. In this paper, prices are set competitively as in GS and so it is just assumed that the monopolist is unable to profit directly by the information. Thus, the entire profit of the monopolist is due to resale of information to traders. The conclusion of Admati and Pfleiderer (1988) is that the monopolist may not sell information to all traders, as is also the case in this paper.

Admati and Pfleiderer (1990) assume personalized noise, and compare (in a GS competitive model) the direct sale of information and sale through a mutual fund. It is found that if the monopolist can choose an asset pricing function as well as a fixed fee (a two part pricing scheme) then he will always prefer to sell information through the fund.

Verrecchia (1982) lets investors be able to choose a certain level of pre-

cision of signals, where the signal disturbance is independently distributed among investors. The level of precision is set equal for investors with the same level of risk aversion. The usual properties for the cost function are assumed. Thus given a cost function, investors demand a certain amount of information. Investors demand different levels of information because they have different risk aversions. Investors with higher risk aversion demand less information. Furthermore, demand for information is lower when the price signal has high precision.

As a rational expectation model with an endogenous state variable one unavoidably has to deal with the problem of "infinite regress". When agents' behavior affects variables, that determine the same agents' strategies, an initially simple model can easily become so complicated that it prevents a solution. Sargent (1991) proposed a general method to deal with this. In the model of Sargent, investors are uncertain about the real stochastic process of all state variables, but observe a subset of these. As time passes agents learn from these observations and in equilibrium, their "perceived law of motion" will correspond to the actual one. The method of Sargent is widely used, for example in the business cycle model of Townsend (1983) and the financial model of Zhou (1998) who considers specialized investors.

A different approach to the problem of infinite regress is however taken here. In this model, information is continuous and nested so that investors only condition on demand generated by more informed investors. Thus, the equilibrium condition is a differential equation. The solution to this equation then determines the moments of the price and signals. Based on this, profits for the monopolist can be computed.

Although there seems to be little theoretical work on nested information in financial markets, Dupont (1997) assumes correlated signals between investors with varying precision, similar to a nested structure. The approach taken here is to assume that such correlations arise specifically by a nested structure on information where the covariance between two signals is the

4

variance of the one with the best precision.

The paper is organized as follows. In the first section, the nested information structure is defined and the formation of expectations is presented.

In the section II we assume that the information monopolist can only discriminate between informed and uninformed investors. The resulting profit maximizing amount of precision, and number of informed investors consistent with a rational expectations Nash equilibrium is then found.

We then allow in section three for infinitely many investors, and derive their competitive strategies.

In section four it is shown that the best the monopolist can do is to distribute the same level of precision to all informed investors.

Results are then discussed in the final concluding section.

I   Continuous information distribution

Let there be a risk free asset which yields a zero rate of return and a risky asset which returns θ in the next period. The risky return is stochastic, but realized before the decision of investment. θ is normally distributed as

$$\theta \sim N\left(\mu_\theta, \sigma_\theta^2\right) \tag{1}$$

θ is however not directly observable by investors. Instead investors can draw continuous independent information increments du_t from a related distribution, in order to learn about θ. The amount of information an investor possesses is a fraction t of full information and is bought at cost c_t. t is proportional to the precision of the stochastic process

$$U = \left\{u_t : u_t = \theta + \tfrac{1}{t}\sigma_\varepsilon\varepsilon_t\right\}, t \in (0,1] \tag{2}$$

where the error increments are distributed independently according to a

Brownian motion

$$d\varepsilon_t \sim N\left(0, \sqrt{t}\right) \tag{3}$$

(2) is a continuous version of the process of sampling an increasing number of discrete observations from a distribution related to θ. In particular it is the case that $\operatorname{var}(u_t) = \frac{1}{t}\sigma_\varepsilon^2$ and $\operatorname{cov}(u_t, u_s) = \frac{1}{\max(t,s)}\sigma_\varepsilon^2$. This covariance property of the information u_t is what makes this a nested information model. The covariance between two different signals is the same as the variance of the most precise signal. Thus, a more informed investor has all the information of less informed investors, but then some extra, which reduces variance.

The information level t is restricted to be in the unit interval. This gives σ_ε^2 the interpretation of the smallest possible variance that the monopolist can offer investors.

A.   Expectation of the asset return

Less informed investors may observe a price signal P in the market as a result of demand from all investors. This price signal will carry information from more informed investors. However, if investors are asymmetrically informed, they will not use the price signal as observed, but rather a refined version of it, P_t, for the following reason. For an investor with positive information t part of the information contained in the price is superfluous. The price is generated partly by the process up to t, that is u_t. The investor does however know u_t, which of course unlike the price is a precise measure of u_t. Just using P as observed is therefore a very naive approach since it brings in imprecise information which could potentially be filtered out. The investor will therefore filter the price observation so that it only depends on the part of the process U which is not known to him. Denote this filtered price with respect to information t as P_t. The exact way to make such adjustments will be presented later.

Due to normality of the stochastic variables, the conditional expectation

is based on a linear relationship:

$$\theta = \mu_\theta + \beta_{u,t}\left(u_t - \mu_\theta\right) + \beta_{P,t}\left(P_t - \mu_\theta\right) + \kappa_t \qquad (4)$$

κ_t being independently normally distributed. The conditional expectation is then $\mathrm{E}\left(\theta|\,t, P_t\right) = \theta - \kappa_t$. For rational expectations the parameters $\beta_{u,t}$ and $\beta_{P,t}$ must be set so that κ_t is independent of the regressors (the signal u_t and the price measure P_t). This will give us two equations per investor type which in principle can be used to solve for the variance of P_t and its covariance with θ, and thereby determining the market equilibrium. In practice however, an explicit solution becomes close to impossible to find for just a few different levels of investors.

II   Optimal distribution of information with two levels of information

Agents have a CARA utility and the same initial wealth. The choice for an investor with information t between holding the amount X_t of the risky asset, and investing M_t in a risky asset is thus subject to the budget constraint

$$W_0 = X_t P + M_t \qquad (5)$$

As in GS, we use the well-known fact that with one risky asset the demand of a CARA investor will be the ratio of discounted expected return to the variance and the CARA-coefficient. The absolute risk aversion and risk free return are however set to unity here, as they play no role in what will be presented. The demand consistent with utility maximization can then be shown to be

$$X_t = \frac{\mathrm{E}\left(\theta|\,u_t, P_t\right) - P}{\mathrm{var}\left(\theta|\,u_t, P_t\right)} \qquad (6)$$

A. Market equilibrium

Let there be λ informed traders and $1 - \lambda$ uninformed traders. The market is in equilibrium whenever

$$(1 - \lambda) \frac{\mathrm{E}\left(\theta\mid P\right) - P}{\mathrm{var}\left(\theta\mid P\right)} + \lambda \frac{\mathrm{E}\left(\theta\mid \mu_t\right) - P}{\mathrm{var}\left(\theta\mid \mu_t\right)} + \eta = 0 \tag{7}$$

holds. η is the stochastic net demand (noise traders). The price P that solves this is the equilibrium price.

The informed investors will disregard the price information of course, because it does not contain any new information for them. The uninformed will use the observed price P, but have no signal to condition on. For informed traders, holding information t, the parameters in the regression equation (4) are

$$\beta_{u,t} = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \frac{1}{t}\sigma_\varepsilon^2}, \beta_{P,t} = 0 \tag{8}$$

$$\mathrm{var}\left(\theta\mid u_t\right) = \sigma_\theta^2 \frac{\sigma_\varepsilon^2}{\sigma_\theta^2 t + \sigma_\varepsilon^2} \tag{9}$$

For the uninformed

$$\beta_{u,0} = 0, \beta_{P,0} = \frac{\sigma_{\theta P}}{\sigma_P^2} \tag{10}$$

$$\mathrm{var}\left(\theta\mid P\right) = \frac{\sigma_\theta^2 \sigma_P^2 - \sigma_{\theta P}^2}{\sigma_P^2} \tag{11}$$

where $\sigma_{\theta P}$ is the covariance between the equilibrium price and the asset return. σ_P^2 is the price variance.

From (7) we can find an expression for P. Similar to Grossman and Stiglitz (1980) we can use this to find expressions for $\sigma_{\theta P}$ and σ_P^2, giving us two

equations to solve these two unknowns. Doing this it can be found that

$$\mathrm{var}\left(\theta\middle|\,P\right) = \sigma_\theta^2 \frac{1}{t}\sigma_\varepsilon^2 \frac{\lambda^2 + \frac{1}{t}\sigma_\varepsilon^2\sigma_\eta^2}{\lambda^2\left(\sigma_\theta^2 + \frac{1}{t}\sigma_\varepsilon^2\right) + \left(\frac{1}{t}\sigma_\varepsilon^2\right)^2\sigma_\eta^2} \tag{12}$$

B. The profit function

The question now is how may the monopolist profit as much as possible by selling an amount of information t to some fraction λ of the traders. Since the monopolist earns zero from the uninformed, his problem is first to charge as much as possible for the information sold. Second he must maximize total revenue by choosing an amount t to sell, and a fraction of investors λ to sell to. When the utility from holding zero information and holding information t is the same, the maximum amount charged for the information is reached. That is, the cost of information consistent with monopolist profit maximization must be such that the utility for an informed investor is the same as that for an uninformed.

A CARA investor with absolute risk aversion equal to one who has bought information t at cost c_t and demands X_t of the risky asset, has utility after realizations of θ and u_t of

$$V_t = -\exp\left[-\left((W_0 - c_t) + X_t\left(\theta - P\right)\right)\right]$$

After being informed the expected utility is

$$\mathrm{E}\left(V_t\middle|\,u_t, P_t\right) = -\exp\left[-\left((W_0 - c_t) + \frac{(\mathrm{E}(\theta|u_t, P_t) - P)^2}{2\,\mathrm{var}(\theta|u_t, P_t)}\right)\right] \tag{13}$$

Prior to being informed u_t is however unknown and so (13) is stochastic when the choice of information level is made. $\mathrm{E}\left(\theta\middle|\,u_t, P_t\right) - P$ is normally distributed with a known expectation and variance and so $\mathrm{E}\left(V_t\right)$ can be found by taking the apropriate integral. It can be shown that (see Appendix A for

details)
$$\mathrm{var}\left(\theta - P\right) = \mathrm{var}\left(\theta | \, u_t, P_t\right) + \mathrm{var}\left(\mathrm{E}\left(\theta | \, u_t, P_t\right) - P\right) \tag{14}$$

which implies that the expectation can be found to be

$$\mathrm{E}\left(V_t\right) = -\exp\left(-\left(W_0 - c_t\right)\right) \tfrac{\mathrm{SD}(\theta | u_t, P_t)}{\mathrm{SD}(\theta - P)} \tag{15}$$

The above expression has a very heuristic interpretation. The fraction on the right hand side is the amount of standard deviation in expected profit that the model (4) cannot explain. Say the investor by use of the model (4) can fully predict returns, that is $\mathrm{SD}\left(\theta | \, u_t, P_t\right) = 0$. If it now was possible for prices (public information) not to track this prediction perfectly in this situation, we would have $\mathrm{SD}\left(\theta - P\right) > 0$ and expected utility equals zero, corresponding to an infinite level of utility.

As mentioned the best the monopolist can do is to charge an amount such that investors are indifferent between buying zero an t information, that is $\mathrm{E}\, V_t = \mathrm{E}\, V_0$. We can solve this for c_t by using (15), yielding

$$c_t = \frac{1}{2} \ln\left(\frac{\mathrm{var}\left(\theta | \, u_0, P_0\right)}{\mathrm{var}\left(\theta | \, u_t, P_t\right)}\right) \tag{16}$$

The monopolist charges this amount for a mass of λ investors. His total profit is therefore λc_t.

C. Optimal mass and precision

The maximization problem is

$$\max_{t,\lambda} \lambda c_t\left(\lambda\right) \tag{17}$$

where the dependence of c_t on λ is apparent from the conditional variance (12) and the definition of the profit function (16). As defined earlier in (2), the best precision the monopolist can supply is defined to be 1. Now define

a variable q solving

$$\frac{1}{2}\sigma_\eta\sigma_\theta \ln\left(\frac{\left(\sigma_\theta^2 t+\sigma_\varepsilon^2\right)\left(q^2\sigma_\theta^2 t+\sigma_\varepsilon^2\right)}{q^2\sigma_\theta^4 t^2+q^2\sigma_\theta^2 t\sigma_\varepsilon^2+\sigma_\varepsilon^4}\right) = \frac{t^2\sigma_\varepsilon^2\sigma_\theta^5\sigma_\eta q^2}{\left(q^2\sigma_\theta^2 t+\sigma_\varepsilon^2\right)\left(q^2\sigma_\theta^4 t^2+q^2\sigma_\theta^2 t\sigma_\varepsilon^2+\sigma_\varepsilon^4\right)} \quad (18)$$

If t has an internal solution, then[1] $q \approx 0.65146$. If the optimal information level is $t^* = 1$, then q depends σ_η, σ_θ and σ_ε as observed in (18), but no explicit solution exists. We can now state the following proposition

Proposition 1 *If the monopolist can discriminate between informed and uninformed investors only, he should optimally supply the amount of information $t^* = \min\left(\frac{1}{\lambda^*}\frac{\sigma_\eta}{\sigma_\theta}\sigma_\varepsilon^2, 1\right)$ to the mass $\lambda^* = \min\left(q\sigma_\eta\sigma_\theta, 1\right)$ of investors.*

The proposition follows from the first order conditions. The profit function is not globally concave with respect to precision and mass, but $\partial\lambda c_t/\partial t$ is positive up to the internal solution for t and decreasing thereafter and so t^* is unique. The uniqueness of λ^* is a bit more tricky. One can however establish that at some positive point the profit λc_t becomes convex with respect to λ and stays convex. Since $\lim_{\lambda\to\infty}\lambda c_t = 0$ and $\lambda c_t > 0$ for all finite λ, a λ satisfying the first order condition gives the global maximum (see appendix B for details).

The intuition behind this result is that if the monopolist only supplies a small amount of information, the difference in precision levels will be very small and thus so will the willingness to pay and the profit. Increasing the amount of information will increase income at first, but it also makes the market more transparent. In the extreme case where almost perfect precision is supplied the price will almost perfectly reveal all information as the informed investors will take very large positions. In this case there will be no willingness to pay for information, and so there is a limit to the amount of information the monopolist will sell.

The same argument can be used to explain how the number of informed investors, λ, affect profits. If only a small number of investors are informed, the volume sold is small and so is the income. If however a large fraction of

11

investors are informed, it may reduce the value of the information because more of it is revealed in the price. The optimal number of investors supplied with information may therefore lie somewhere between zero and unity.

It is thus optimal to distribute as much information as possible up to $t^* = \frac{\sigma_\varepsilon^2}{q\sigma_\theta^2}$. If the monopolist has more information than this, that is $\frac{\sigma_\varepsilon^2}{q\sigma_\theta^2} < 1$, he is better off keeping it to himself. σ_ε^2 is the residual variance of the best information that the monopolist can supply. It is therefore the case that if the information quality is poor, there is a high probability that the monopolist should optimally supply all his information.

Interestingly the optimal amount of information is actually independent of the level of noise trading when λ^* is the internal solution. This is because as the level of noise trading increases, the increase in number of informed traders will exactly offset the positive effect of noise trading on the profit.

If we look at the optimal signal variance we find that inducing $\frac{\sigma_\varepsilon^2}{t} = \sigma_\theta^2 q$ maximizes the profit of the monopolist. Thus if there is a lot of uncertainty in the variance of the asset's return, that is a high σ_θ, then the investor would profit by selling relatively imprecise information to the market. If there is a lot of disturbance to the market price through noise traders, a high σ_η, then more precise information should be sold because the information value is less degraded by the price signal.

III   Optimal distribution of information with infinitely many levels of information

We now turn to the more general case where the monopolist is able to discriminate between an infinite number of investors. That is, the monopolist is able to assign precision t to a density f_t of investors at any level of precision. In this setting, with a continuum of investors, the demand is an integral over these investors. Thus the market equilibrium condition corresponding to (7) in the previous section is:

$$\int_0^1 f_y \frac{\mathrm{E}\left(\theta\mid u_y, P_y\right) - P}{\mathrm{var}\left(\theta\mid u_y, P_y\right)} dy + \eta = 0 \qquad (19)$$

η being random supply/demand from noise traders with expectation 0 for simplicity. Since investors know exactly the demand of less informed investors (due to the nested information structure assumed), they will only consider the demand generated by better-informed investors. This means that for an investor with information t only, the part of the integral (19) above t is of interest (it can be shown that this indeed yields a more efficient estimate).

A. The price process

Solving for P in (19) and censoring the part of the integral with precision less than t gives

$$(P_t - \mu_\theta) = \frac{1}{v} \int_t^1 f_y \frac{\mathrm{E}\left(\theta\mid u_y, P_y\right) - \mu_\theta}{\mathrm{var}\left(\theta\mid u_y, P_y\right)} dy + \frac{1}{v}\eta \qquad (20)$$

Where $v = \int_0^1 f_y / \mathrm{var}\left(\theta\mid u_y, P_y\right) dy$. (20) is a stochastic differential equation. Due to linearity however, we can treat it as an ordinary differential equation. The general solution can be found by the standard formula for a linear first order differential equation, and is

$$(P_t - \mu_\theta) = m_t^{-1}\left(\int_t^1 g_y\left(u_y - \mu_\theta\right) dy + \eta\right) \qquad (21)$$

where

$$g_t = f_t \frac{1}{v} m_t \frac{\beta_{u,t}}{\mathrm{var}\left(\theta\mid u_t, P_t\right)} \qquad (22)$$

$$m_t = v e^{-\int_t^T \frac{1}{v} f_z \frac{\beta_{P,z}}{\mathrm{var}\left(\theta\mid u_z, P_z\right)} dz} \qquad (23)$$

13

B.   Joint moments of the price process and signals

We want to find the joint moments of the signals u_t and the price measure P_t. To do this we need the functions defined below

$$
\begin{aligned}
b_t &= \int_t^1 g_x dx \\
\gamma_t &= \int_t^1 g_x \frac{1}{x} dx \\
a_t &= \int_t^1 g_x \left(\gamma_x - \tfrac{1}{x} b_x \right) dx
\end{aligned}
\tag{24}
$$

We can now write the joint moments in terms of the above functions (for details see appendix C):

$$
\begin{aligned}
\sigma_{P_t, u_t} &= m_t^{-1} \left(b_t \sigma_\theta^2 + \gamma_t \sigma_\varepsilon^2 \right) && (25) \\
\sigma_{P_t}^2 &= m_t^{-2} \left(\sigma_\theta^2 b_t^2 + \sigma_\varepsilon^2 \left(a_t + b_t \gamma_t \right) + \sigma_\eta^2 \right) && (26) \\
\sigma_{P_t, \theta} &= m_t^{-1} b_t \sigma_\theta^2 && (27) \\
\sigma_{u_t}^2 &= \sigma_\theta^2 + \tfrac{1}{t} \sigma_\varepsilon^2 && (28) \\
\sigma_{u_t, \theta} &= \sigma_\theta^2 && (29)
\end{aligned}
$$

As previously mentioned the parameters $\beta_{u,t}$ and $\beta_{P,t}$ must be set so that the residual in (4) is independent of the regressors. This is obtained when

$$
\begin{aligned}
\beta_{u,t} &= \frac{\sigma_\theta^2}{|\Sigma| m_t^2} \left(\sigma_\varepsilon^2 a_t + \sigma_\eta^2 \right) && (30) \\
\beta_{P,t} &= \frac{\sigma_\theta^2 \sigma_\varepsilon^2}{|\Sigma| m_t} \left(\frac{1}{t} b_t - \gamma_t \right) && (31)
\end{aligned}
$$

where the corresponding covariance matrix is

$$
\Sigma = \begin{pmatrix} \sigma_{u_t}^2 & \sigma_{P_t, u_t} \\ \sigma_{P_t, u_t} & \sigma_{P_t}^2 \end{pmatrix}
\tag{32}
$$

14

This implies the conditional variance (which in turn determines the willingness to pay for information)

$$\text{var}\left(\theta\mid u_t, P_t\right) = \frac{\sigma_\theta^2 \sigma_\varepsilon^2}{\left|\Sigma\right| m_t^2}\left(\frac{1}{t}\sigma_\varepsilon^2\left(a_t + b_t\gamma_t\right) - \gamma_t^2\sigma_\varepsilon^2 + \frac{1}{t}\sigma_\eta^2\right) \tag{33}$$

It is not possible to find an explicit expression of this variance, since the function g_t can only be found through numerical procedures with limited accuracy.

C. The profit function

As in the previous section, investors with information t are willing to pay an amount c_t such that they are indifferent between any level of precision. Although we are now considering an infinite number of investors, the function c_t largely remain the same as defined in (16), except that it will not depend on λ but indirectly on the density function f_t through the change in the variables a_y, b_y and γ_y. At a given level of information the monopolist earns $f_t c_t dt$. We can therefore write the total income to the monopolist as

$$\Pi = \int_0^1 f_t c_t\left(a_y, b_y, \gamma_y, t\right) dt \tag{34}$$

where

$$c_t\left(a_t, b_t, \gamma_t\right) = \frac{1}{2}\ln\left(\frac{\text{var}_f\left(\theta\mid u_0, P_t\right)}{\text{var}_f\left(\theta\mid u_t, P_t\right)}\right) \tag{35}$$

which we note is independent of m_t.

IV Solving for the optimal distribution of information

When solving for the optimal distribution of information, optimal control theory will be used. Since we do not have an explicit expression for g_t the

solution will not rely on an explicit specification of the profit function.

The monopolist wants to maximize (34). His problem is to

$$\max_{f_t} \int_0^1 f_y c_y \left(a_y, b_y, \gamma_y\right) dy \qquad (36)$$

using the density f_t as the only control variable. The state at precision t is described by the state variables a_t, b_t and γ_t. We now state the following proposition

Proposition 2 *The distribution of information that maximizes overall information value* Π *is to have all informed investors concentrated at a single level of precision* t^*

Proof: By the definition of the state variables (24) as well as (22), the dynamic constraints follow

$$\dot{b}_t = -f_t \frac{t\left(a_t \sigma_\varepsilon^2 + \sigma_\eta^2\right)^2}{\sigma_\varepsilon^2 \sigma_\eta^2 \left(\left(a_t + \gamma_t \left(b_t - t\gamma_t\right)\right) \sigma_\varepsilon^2 + \sigma_\eta^2\right)} \qquad (37)$$

$$\dot{\gamma}_t = \tfrac{1}{t}\dot{b}_t \qquad (38)$$

$$\dot{a}_t = -\left(\frac{1}{t}b_t - \gamma_t\right)\dot{b}_t \qquad (39)$$

Denote the corresponding Hamilton function as H_t. Note here that the derivative of all state variables are proportional to the control f_t and that this is the case for the profit function, $f_y c_y$, at a given precision level too. This implies that

$$\frac{\partial H_t}{\partial f_t} = h_t \left(a_t, b_t, \gamma_t\right) \qquad (40)$$

for some function h_t dependent on a_t, b_t, γ_t and precision level only. Thus $\partial H_t / \partial f_t$ is independent of the control f_t, which is an important point when proving Proposition 2. The monopolist will choose the density of investors f_t at any precision level t that maximizes H_t. This can be stated by the

optimality condition

$$h_t^* (f_t^* - f_t) \geq 0 \qquad (41)$$

for all possible t and f_t. Asterisks indicate values calculated with the optimal control f_t^*. We see that (39) is satisfied if $f_t^* = 0$ whenever $h_t^* < 0$ and f_t^* set to its maximum value whenever $h_t^* \geq 0$.

Now denote by t^* the smallest t where $h_t^* \geq 0$ is true. Let $\lambda\Delta$ be the length of an interval $I = [t^*, t^* + \lambda\Delta]$, where $\lambda \in [0, 1]$ is set such that $h_t^* \geq 0 \forall t \in I$.

A closed solution requires a maximum value for f_t. We therefore assume the density is bounded as

$$f_t \in \left[0, \tfrac{1}{\Delta}\right] \qquad (42)$$

and so in the interval I the density of investors is set to $f_t = \frac{1}{\Delta}$ by the monopolist. This implies that the mass of informed investors in this interval is λ.

Now two possibilities arise:

Conjecture 1 $h_{t^*}^* \geq 0$ *at* $t^* + \Delta$, *that is* $\lambda = 1$, *meaning the total mass of investors have been assigned information, and consequently this is the only interval with a strictly positive density of investors.*

Conjecture 2 $h_t^* < 0$ *for some* $t > t^* + \lambda\Delta$ *and for* $\lambda \in [0, 1)$

We now need to investigate if in case of Conjecture 2 there are any subsequent points at which $h_t^* \geq 0$ and $f_t^* > 0$. We do this by noting the following facts:

a) We see from the dynamic constraints (37) - (39) that a_t, b_t and γ_t are in fact constants when $f_t^* = 0$.

b) The associated adjoint functions are constant too when $f_t^* = 0$.

c) h_t^* is continuous (due to continuity of the state variables, the definition of c_t and its independence of f_t)

d) Statement c) and Conjecture 2 together implies that there must be some $t > t^* + \lambda\Delta$ for which $h_t^* < 0$ and $dh_t^*/dt \leq 0$ holds simultaneously arbitrarily close to $t^* + \lambda\Delta$.

17

f) With constant state variables and adjoint functions one can show that if $dh_t^*/dt \leq 0$ and $h_t^* < 0$ for some $\bar{t}$, then $h_t^* < 0$ also hold for any $t \geq \bar{t}$. This is true because with constant variables as stated in a) and b), it is true that $dh_t^*/dt = \partial h_t^*/\partial t$. By evaluating this derivative one will find that the sign of dh_t^*/dt only changes to negative once.

Thus, we may conclude that I is the only interval in which $f_t^* > 0$. Allowing f_t^* to approach infinity by letting $\Delta \to 0$ implies that the interval for which $f_t^* > 0$ approaches a single point of mass λ ∎

Since it has been shown that it is optimal to concentrate the information sale to a single level of precision, we can restate Proposition 1 in more general terms

Proposition 3 *A monopolist that can discriminate between an infinite number investors, will optimally supply the same amount of information $t^* = \min\left(\frac{1}{\lambda^*}\frac{\sigma_\eta}{\sigma_\theta}\sigma_\varepsilon^2, 1\right)$ to a mass $\lambda^* = \min\left(q^*\sigma_\eta\sigma_\theta, 1\right)$ of investors.*

V   Concluding remarks

This model shows that informing investors asymmetrically is never optimal. When considering only homogenously informed investors it is found that a well-informed information monopolist should sell less than all his information. If there is a lot of noise, additional information has a less degrading effect on its value, and so the monopolist can sell more information without loss. The opposite is the case if the asset itself is very risky. In that case, less information should be sold, because a smaller amount of information has substantial value.

The findings in this paper show that an insider might consider holding back information. Thus, there may be a natural limit to how much information will be supplied by insiders in an asset market. Also, the insider will actually prefer to supply market participants with the same information even when distributing different information can be done at no extra cost

and unlimited discrimination in terms of a continuous density function is possible.

In practice, of course this may be difficult since the expected loss in being caught is not taken into account in this model. It is however an interesting idea that if a legal market was created for insider information, then it should according to this model be distributed evenly and with low precision. For example one might envision the creation of a market where firms supply inside information to investors for a payment equal to the negative effect that supplying this information may have on their profit due to revelation of business strategies. This would increase transparency and thereby reduce market risk, and also reduce incentives for criminal activity.

Since there is an upper limit to how precise the information sold in a market will be, the model also suggests that analysts may limit the extent of their research because very precise information may not be advantageous to sell. This of course only holds as long as the analyst is actually able to obtain monopoly on the information he gathers. The effect that not all information will be sold also shows that there may be a "dead weight loss" in information markets as well. The unsold information would of course make investors better off by reducing variance, and they would be willing to pay for this. However, since the monopolist controls the information flow, he can take into account the negative effect that increased information in the market has on the information value of the price.

Possible extensions of this model will for example be to let the monopolist do some limited trading. Also, it would be of interest to look at other distributions of information not necessarily optimal for the monopolist. In this way, one might obtain an understanding of how the value of information is depleted by a heterogeneous distribution of it.

References

Admati, Anat R, and Paul Pfleiderer, 1988, Selling and Trading on Information in Financial Markets, *American Economic Review* 78, 96–103.

Admati, Anat R., and Paul Pfleiderer, 1990, Direct and INdirect Sale of Information, *Econometrica* 58, 901–928.

Dupont, Dominique, 1997, Trading Volume and Information Distribution in a Market-Clearing Framework, *Working paper, Board of Governors of the Federal Reserve System.*

Grossman, S. J., and J. E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *The American Economic Review* 70, No. 3, 393–408.

Kyle, Albert S., 1985, Continous Auctions and Insider Trading, *Econometrica* 53, 1315–1336.

Sargent, Thomas J., 1991, Equilibrium with Signal Extraction from Endogenous Variables, *Journal of Economic Dynamics and Control* 15, 245–273.

Townsend, Robert M, 1983, Forecasting the Forecasts of Others, *Journal of Political Economy* 91, 546–88.

Verrecchia, Robert E., 1982, Information Aquisition in a Noisy Rational Expectations Economy, *Econometrica* 50, 1415–1430.

Zhou, Chunsheng, 1998, Dynamic Portfolio Choice and Asset Pricing with Differential Information, *Journal of Economic Dynamics and Control* 22, 1027–1051.

Notes

$q = 0.651\,460\,502\,417\,619\,437$ in double precision.

VI   Appendix

A.   Unconditional variance of the profit

It will be proven that

$$\mathrm{var}\,(\theta - P) = \mathrm{var}\,(\theta|\,u_t, P_t) + \mathrm{var}\,(\mathrm{E}\,(\theta|\,u_t, P_t) - P) \tag{A.1}$$

First it must be shown that the residual of the regression (4) κ_t is uncorrelated with the price P. Recall (21):

$$(P_t - \mu_\theta)\,m_t = \int_t^1 g_y\,(u_y - \mu_\theta)\,dy + \eta \tag{A.2}$$

we can write P as

$$P = \frac{m_t}{m_0}P_t + \left(P - \frac{m_t}{m_0}P_t\right) \tag{A.3}$$

Define

$$
\begin{aligned}
\tilde{P}_t &= \left(P - \frac{m_t}{m_0}P_t\right) \\
&= \frac{1}{m_0}\int_0^t g_y\,(u_y - \mu_\theta)\,dy
\end{aligned} \tag{A.4}
$$

Using (4) we can write

$$\mathrm{cov}\left(\tilde{P}_t, \kappa_t\right) = \mathrm{E}\,(\theta - \mu_\theta)\,\tilde{P}_t \quad -\beta_{u,t}\,\mathrm{E}\,(u_t - \mu_\theta)\,\tilde{P}_t \tag{A.5}$$
$$-\beta_{P,t}\,\mathrm{E}\,(P_t - \mu_\theta)\,\tilde{P}_t$$

21

By the definition of u_t in , and the definition of $\tilde{P}_t$ (A.4) it follows that

$$\mathrm{E}\left(\theta - \mu_\theta\right)\tilde{P}_t = \frac{1}{m_0}\int_0^t g_y dy \sigma_\theta^2 \qquad (A.6)$$

$$\mathrm{E}\left(u_t - \mu_\theta\right)\tilde{P}_t = \frac{1}{m_0}\int_0^t g_y dy \sigma_{u_t}^2 \qquad (A.7)$$

$$\mathrm{E}\left(P_t - \mu_\theta\right)\tilde{P}_t = \frac{1}{m_0}\int_0^t g_y dy \sigma_{u_t,P_t} \qquad (A.8)$$

and we can therefore write the covariance between $\tilde{P}_t$ and κ_t as

$$\mathrm{cov}\left(\tilde{P}_t, \kappa_t\right) = \left(\sigma_\theta^2 - \beta_{u,t}\sigma_{u_t}^2 - \beta_{P,t}\sigma_{u_t,P_t}\right)\frac{1}{m_0}\int_0^t g_y dy \qquad (A.9)$$

$$= \mathrm{cov}\left(u_t, \kappa_t\right)\frac{1}{m_0}\int_0^t g_y dy = 0$$

because $\mathrm{cov}\left(u_t, \kappa_t\right) = 0$ (which is true because u_t is a regressor). We can thus conclude that $\mathrm{cov}\left(P, \kappa_t\right) = 0 \forall t$.

Since $\mathrm{E}\left(\theta|\, u_t, P_t\right) = \theta - \kappa_t$ we have

$$\mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right) - P\right) = \mathrm{var}\left(\theta - \kappa_t - P\right)$$

which we can write out as

$$\mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right) - P\right) = \mathrm{var}\left(\theta - \kappa_t\right) + 2\,\mathrm{cov}\left(\kappa_t, P\right) - 2\sigma_{\theta,P} + \sigma_P^2$$

Using $\mathrm{cov}\left(\kappa_t, P\right) = 0$ and again that $\mathrm{E}\left(\theta|\, u_t, P_t\right) = \theta - \kappa_t$ we can write

$$\mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right) - P\right) = \mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right)\right) - 2\sigma_{\theta,P} + \sigma_P^2$$

Rearranging we get

$$\mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right) - P\right) = \left(\sigma_\theta^2 - 2\sigma_{\theta,P} + \sigma_P^2\right) - \left(\sigma_\theta^2 - \mathrm{var}\left(\mathrm{E}\left(\theta|\, u_t, P_t\right)\right)\right)$$

The first term on the right and side equals var $(\theta - P)$ and the second term is var $(\theta | u_t, P_t)$ ∎

B. Uniqueness of λ^*

The maximum of the profit function λc_t is given in Proposition 1, and we want to be sure that this is a global maximum. The only positive real solution for $\partial^2 \lambda c_t / \partial \lambda^2 = 0$ is

$$\hat{\lambda} = \frac{1}{\sqrt{2t}} \sqrt{\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2}\right) \left(1 + \sqrt{12 \frac{(\sigma_\varepsilon^2 + t\sigma_\theta^2)\sigma_\varepsilon^2}{(2\sigma_\varepsilon^2 + t\sigma_\theta^2)^2}}\right)} \sigma_\eta \sigma_\varepsilon \qquad \text{(A.10)}$$

It can be shown that $\partial^3 c_t / \partial \lambda^3 < 0$, so the function starts concave and then turns convex at $\lambda = \hat{\lambda}$. Furthermore it can be shown that $\lim_{\lambda \to \infty} \lambda c_t = 0$ and also it is a fact that $\lambda c_t > 0$ for finite λ. This implies that in the convex region $\lambda > \hat{\lambda}$ of the profit function, the profit is always strictly decreasing, since the profit is never negative. Of course λ is always less than unity, but the point is that the first order condition can only hold in the concave region of the profit function. Therefore if there exist a λ satisfying the first order condition, this must in fact be the optimal λ maximizing the profit function λc_t ∎

C. The moments of the price process

It will be shown how to derive the moments given in (25) - (29). From (21) we have that

$$(P_t - \mu_\theta) = m_t^{-1} \left(\int_t^1 g_y (u_y - \mu_\theta) \, dy + \eta \right) \qquad \text{(A.11)}$$

The derivation of the moments builds on the fact that given the process

of u_t in (2) the variance of the signal process is

$$\sigma_{u_t}^2 = \mathrm{E}\left(u_y - \mu_\theta\right)^2 = \sigma_\theta^2 + \tfrac{1}{t}\sigma_\varepsilon^2 \tag{A.12}$$

and the covariance between two signals of different precisions

$$\sigma_{u_s,u_t} = \mathrm{E}\left(u_s - \mu_\theta\right)\left(u_t - \mu_\theta\right) = \sigma_\theta^2 + \tfrac{1}{\max(s,t)}\sigma_\varepsilon^2 \tag{A.13}$$

because the innovation from s to t (or vice versa) is independent of the least precise signal. The most precise signal is θ itself, and so

$$\sigma_{u_t,\theta} = \mathrm{E}\left(u_t - \mu_\theta\right)\left(\theta - \mu_\theta\right) = \sigma_\theta^2 \tag{A.14}$$

From this, and the definitions of a_t, b_t and γ_t in (24) it follows that the price-signal covariance is

$$
\begin{aligned}
\sigma_{P_t,u_t} &= \mathrm{E}\, m_t^{-1}\left(\int_t^1 g_y\left(u_y - \mu_\theta\right) dy + \eta\right)\left(u_t - \mu_\theta\right)\\
&= m_t^{-1}\left(\int_t^1 g_y\,\mathrm{E}\left(\left(u_y - \mu_\theta\right)\left(u_t - \mu_\theta\right)\right) dy\right)\\
&= m_t^{-1}\left(\int_t^1 g_y\left(\sigma_\theta^2 + \frac{1}{y}\sigma_\varepsilon^2\right) dy\right)\\
&= m_t^{-1}\left(b_t\sigma_\theta^2 + \gamma_t\sigma_\varepsilon^2\right)
\end{aligned}
\tag{A.15}
$$

It also follows from definitions of a_t, b_t and γ_t that $m_t = \sigma_\eta^2 + a_t\sigma_\varepsilon^2$. The price

variance is

$$
\begin{aligned}
\sigma_{P_t}^2 &= m_t^{-2}\,\mathrm{E}\left(\int_t^1 g_y\left(u_y-\mu_\theta\right)dy+\eta\right)^2 \\
&= m_t^{-2}\left(\int_t^1\int_t^1 g_x g_y\,\mathrm{E}\left(u_y-\mu_\theta\right)\left(u_x-\mu_\theta\right)dydx\right)+\sigma_\eta^2 \\
&= m_t^{-2}\left(\int_t^1\int_t^x g_y g_x\left(\sigma_\theta^2+\frac{1}{x}\sigma_\varepsilon^2\right)dydx\right) \qquad\qquad \text{(A.16)} \\
&\quad +m_t^{-2}\left(\int_t^1\int_x^1 g_y g_x\left(\sigma_\theta^2+\frac{1}{y}\sigma_\varepsilon^2\right)dydx\right)+\sigma_\eta^2 \\
&= m_t^{-2}\left(\sigma_\theta^2 b_t^2+\sigma_\varepsilon^2\left(a_t+b_t\gamma_t\right)+\sigma_\eta^2\right)
\end{aligned}
$$

and the price-state covariance is

$$
\begin{aligned}
\sigma_{P_t,\theta} &= m_t^{-1}\,\mathrm{E}\left(\int_t^1 g_y\left(u_y-\mu_\theta\right)dy+\eta\right)\left(\theta-\mu_\theta\right) \\
&= m_t^{-1}\left(\int_t^1 g_y dy\right)\sigma_\theta^2 \qquad\qquad\qquad\qquad\quad \text{(A.17)} \\
&= m_t^{-1} b_t\sigma_\theta^2
\end{aligned}
$$

D. h_t^* has a single maximum if $f_t=0$

It will be proven that in intervals of t where the density f_t is zero, there is a single global maximum point of h_t^* and more importantly, if $h_t^*<0$ for some $t=\bar{t}$, then this will be true also for $t>\bar{t}$. This is important for the proof of Proposition 2.

The derivative of the Hamilton function with respect to the control is

$$
\frac{\partial H_t}{\partial f_t} = \frac{H_t}{f_t}=h_t \qquad\qquad\qquad\qquad\qquad \text{(A.18)}
$$

$$
h_t = c_t\left(a_t,b_t,\gamma_t\right)+\sum_{i=1}^3 p_{i,t}\psi_i\left(a_t,b_t,\gamma_t,t\right) \qquad\qquad \text{(A.19)}
$$

25

where the functions ψ_i are the right hand sides of the dynamic constraints (37) - (39), divided by f_t:

$$\psi_1\left(a_t, b_t, \gamma_t, t\right) = -\frac{t\left(a_t\sigma_\varepsilon^2 + \sigma_\eta^2\right)^2}{\sigma_\varepsilon^2\sigma_\eta^2\left(\left(a_t + \gamma_t\left(b_t - t\gamma_t\right)\right)\sigma_\varepsilon^2 + \sigma_\eta^2\right)} \quad \text{(A.20)}$$

$$\psi_2\left(a_t, b_t, \gamma_t, t\right) = -\frac{\left(a_t\sigma_\varepsilon^2 + \sigma_\eta^2\right)^2}{\sigma_\varepsilon^2\sigma_\eta^2\left(\left(a_t + \gamma_t\left(b_t - t\gamma_t\right)\right)\sigma_\varepsilon^2 + \sigma_\eta^2\right)} \quad \text{(A.21)}$$

$$\psi_3\left(a_t, b_t, \gamma_t, t\right) = -\frac{t\left(\frac{1}{t}b_t - \gamma_t\right)\left(a_t\sigma_\varepsilon^2 + \sigma_\eta^2\right)^2}{\sigma_\varepsilon^2\sigma_\eta^2\left(\left(a_t + \gamma_t\left(b_t - t\gamma_t\right)\right)\sigma_\varepsilon^2 + \sigma_\eta^2\right)} \quad \text{(A.22)}$$

and $c_t\left(a_t, b_t, \gamma_t\right)$ is found by inserting

$$\text{var}_f\left(\theta\middle| u_t, P_t\right) = \ln\left(\frac{\sigma_\theta^2\sigma_\varepsilon^2\left(a_t + \gamma_t(b_t - t\gamma_t)\sigma_\varepsilon^2 + \sigma_\eta^2\right)}{(a_t + \gamma_t(b_t - t\gamma_t))\sigma_\varepsilon^4 + t\sigma_\eta^2\sigma_\theta^2 + \sigma_\varepsilon^2\left(\gamma_t(b_t(b_t - t\gamma_t) + a_t t)\sigma_\theta^2 + \sigma_\eta^2\right)}\right) \quad \text{(A.23)}$$

into the cost function $c_t = \frac{1}{2}\ln\frac{\text{var}_f(\theta|u_0, P_0)}{\text{var}_f(\theta|u_t, P_t)}$

From the dynamic constraints (37) - (39) it follows that a_t, b_t and γ_t are constants if $f_t = 0$. Furthermore, since $H_t = 0$ if $f_t = 0$ it is also true that $\partial H_t/\partial a_t = \partial H_t/\partial b_t = \partial H_t/\partial \gamma_t = 0$. This in turn implies that if $f_t = 0$ then $p_{i,t} = 0, i \in \{1, 2, 3\}$. Using this, it can be found that

$$\left.\frac{dh_t}{dt}\right|_{f_t=0} = \frac{\partial h_t}{\partial t}$$

$$= \frac{a_t\sigma_\varepsilon^2 + \sigma_\eta^2}{2\left(a_t + \gamma_t(b_t - t\gamma_t)\sigma_\varepsilon^2 + \sigma_\eta^2\right)}\left(-K + \sigma_\theta^2\,\text{var}_f\left(\theta\middle| u_t, P_t\right)\right) \quad \text{(A.24)}$$

where

$$K = \frac{2}{\sigma_\eta^2\sigma_\varepsilon^2}\left(\left(\psi_1 + \psi_3\gamma_t\right)\left(a_t\sigma_\varepsilon^2 + \sigma_\eta^2\right) + \gamma_t\left(b\psi_1 + \psi_2\gamma_t\right)\sigma_\varepsilon^2\right) \quad \text{(A.25)}$$

h_t is not globally concave in t but the sign of $\partial h_t/\partial t$ depends on $\left(-K + \sigma_\theta^2\,\text{var}_f\left(\theta\middle| u_t, P_t\right)\right)$.

It can be found that

$$
\frac{\partial}{\partial t}\left(-K + \sigma_\theta^2 \operatorname{var}_f\left(\theta \mid u_t, P_t\right)\right)
$$

$$
= -\frac{\sigma_\theta^2\left(a_t \sigma_\varepsilon^2 + \sigma_\eta^2\right)^2}{\left(\left(a_t + \gamma_t(b_t - t\gamma_t)\right)\sigma_\varepsilon^4 + t\sigma_\eta^2\sigma_\theta^2 + \sigma_\varepsilon^2\left(\gamma_t(b_t(b_t - t\gamma_t) + a_t t)\sigma_\theta^2 + \sigma_\eta^2\right)\right)^2} \qquad (A.26)
$$

$$
< \quad 0
$$

Which means that if $dh_t/dt < 0$ and $f_t = 0$ at $t = \bar{t}$, then this is true $\forall t > \bar{t}$ ∎

27