Online Appendix to "Incentives for Information Production in Markets where Prices Affect Real Investment"

James Dow* Itay Goldstein† Alexander Guembel‡

June 21, 2016

1 Welfare

In this Appendix, we provide a welfare analysis and characterize the socially optimal amount of information production. This Appendix completes the discussion of Section 5 of our paper. In the model we assume that noise trading is distributed according to $n \sim N\left(0, \sigma^2\right)$. In order to carry out a welfare analysis, we now rationalize this distribution of noise trading. Specifically, we assume that some agents in the economy derive a private benefit from either selling or buying one unit of the stock. This could be due to a private value that they derive from holding a long or short position in the stock. The simplest way to think about this benefit is that it stems from the need to hedge an exogenous risky position. That is, agents may have a positive or a negative exposure correlated with the fundamental of the stock, and so wish to sell it or buy it, respectively. Other possible reasons for private valuations include liquidity needs, differential tax treatment and financing costs (for a similar approach see, for example, Duffie, Garleanu and Pedersen, 2005).

We assume that the measure of these agents in the economy is $|n|$, which is distributed according to a half-normal distribution. We also assume that for any realization of n, with probability 0.5 all these agents have a positive exposure leading them to sell a unit, and with probability 0.5 all these agents have a negative exposure leading them to buy a unit.[1] More precisely, trading a unit in the opposite direction to their exposure leads the agents to obtain a fixed benefit b. They also make an endogenously determined trading profit of $V - P$ when they buy and $P - V$ when they sell (of course, in equilibrium this will be negative on average). Note that trading profit/loss is bounded, and so by assuming a high enough b, we can ensure that

*London Business School, Sussex Place, Regent´s Park, London, NW1 4SA, UK. Phone: +44 207 000 7000, email: jdow@london.edu.

†Department of Finance, Wharton School, University of Pennsylvania, Steinberg Hall - Dietrich Hall, Suite 2300, Philadelphia, PA 19104-6367, USA. Phone: +1 215 746 0499, e-mail: itayg@wharton.upenn.edu.

‡Toulouse School of Economics, Manufacture des Tabacs, 21, allée de Brienne, 31000 Toulouse, France. Phone: +33 561 128574, e-mail: alexander.guembel@univ-tlse1.fr.

[1] This treatment is reminiscient of Zhu (2014).

these agents will always trade to offset their exposure. This leads to the noise trading distribution previously assumed in the paper: $n \sim N\left(0, \sigma^2\right)$.

Welfare is the sum of the utilities of players in the financial market (speculators, noise traders, and market maker) and the value of the firm. Note that the trading profits in the financial market sum up to zero: the market maker breaks even and the expected gain of the speculators is the expected loss of the noise traders. Hence, giving equal weights to the different agents, overall welfare is given by $V(\alpha) - \alpha c + E[b|n|]$, that is, the value of the firm minus the cost of information production incurred by the speculators plus the expected private benefit from trading for the noise traders.

Suppose that the social planner can choose the amount of information production α in the economy to maximize this expression for social welfare. The third term in the welfare expression above – expected private benefits for the noise traders – does not depend on α. Hence, the social planner has to solve:

$$\max_{\alpha} V(\alpha) - \alpha c. \tag{1}$$

I.e., maximize the value of the firm (which is higher when there is more information) minus the cost of information production (which is also higher when there is more information). The way we have rationalized noise trading is important for deriving this simple specification for welfare. One could consider alternative specifications, in which noise traders' behaviour depends on the amount of information in the market, and welfare would therefore depend on information in a more complex way. We view such effects as being unrelated to the main issues highlighted here, and so we leave investigation of these alternative settings to future research.

Denote by α^s the constrained efficient level of information production, which maximizes the expression in (1). Then:

Proposition 1.1 *(i) If $\gamma < \gamma_0$ or $\gamma > 1 - \gamma_0$, then the constrained efficient level of information production is $\alpha^s = 0$, where*

$$\gamma_0 = c\frac{\sqrt{2\pi\sigma^2}}{(V_h - V_l)(2\lambda - 1)}. \tag{2}$$

Otherwise, $\alpha^s > 0$.

(ii) If $\gamma_0 < \frac{1}{2}$, then α^s is increasing in γ for $\gamma_0 < \gamma < \frac{1}{2}$ and decreasing in γ for $1 - \gamma_0 > \gamma > \frac{1}{2}$.

(iii) If $\gamma_0 < \frac{1}{2}$, then for $\gamma = \frac{1}{2}$, the socially optimal level of information production is given by

$$\alpha^s\left(\gamma = \frac{1}{2}\right) = \frac{\sigma}{2\lambda - 1}\sqrt{-2\ln(2\gamma_0)}.$$

Proof of Proposition 1.1: From the proof of Proposition 4 we know that $V(\alpha)$ is monotonically increasing in α and can be written as

$$\frac{\partial V(\alpha)}{\partial \alpha} = \frac{1}{2}(V_h - V_l)\left[\gamma\varphi\left(\overline{X}(\widehat{\alpha};\gamma) - \widehat{\alpha}(2\lambda - 1)\right) + (1-\gamma)\varphi\left(\overline{X}(\widehat{\alpha};\gamma) + \widehat{\alpha}(2\lambda - 1)\right)\right]$$

Moreover, when $\gamma < \frac{1}{2}$ then $V(\alpha)$ is convex for $[\alpha(2\lambda - 1)]^2 < \frac{\sigma^2}{2}\ln\frac{1-\gamma}{\gamma}$ and concave otherwise. When $\gamma > \frac{1}{2}$ then $V(\alpha)$ is convex for $[\alpha(2\lambda - 1)]^2 < -\frac{\sigma^2}{2}\ln\frac{1-\gamma}{\gamma}$ and concave otherwise. The maximum slope of

$V(\alpha)$ is attained in its saddle point $\widetilde{\alpha}$, defined by the previous regions. In the saddle point, the exact value of $\frac{\partial V}{\partial \alpha}$ can be calculated as

$$\left.\frac{\partial V(\alpha)}{\partial \alpha}\right|_{\widetilde{\alpha}} = \begin{cases} \gamma \frac{(V_h - V_l)(2\lambda - 1)}{\sqrt{2\pi\sigma^2}} & \text{if} \quad \gamma < \frac{1}{2} \\ (1-\gamma) \frac{(V_h - V_l)(2\lambda - 1)}{\sqrt{2\pi\sigma^2}} & \text{if} \quad \gamma > \frac{1}{2} \end{cases}.$$

Clearly, if $\left.\frac{\partial V(\alpha)}{\partial \alpha}\right|_{\widetilde{\alpha}} < c$ then the first-order condition cannot be satisfied and $\alpha^s = 0$.

We can calculate $\frac{\partial \alpha^s}{\partial \gamma}$ from the implicit function theorem:

$$\frac{\partial \alpha^s}{\partial \gamma} = -\frac{\frac{\partial V(\alpha,\gamma)}{\partial \gamma}}{\frac{\partial V(\alpha,\gamma)}{\partial \alpha}}.$$

Doing the calculations yields

$$\frac{\partial \alpha^s}{\partial \gamma} = \sigma^2 \frac{\varphi\left[\overline{X}(\alpha) - \alpha(2\lambda-1)\right] - \varphi\left[\overline{X}(\alpha) + \alpha(2\lambda-1)\right] - \frac{\partial \overline{X}(\alpha;\gamma)}{\partial \gamma} \frac{\overline{X}(\alpha;\gamma)}{\sigma^2} K}{\left\{-\frac{\left(\overline{X}(\alpha)\right)^2}{\alpha} + \alpha(2\lambda-1)^2\right\} K}, \tag{3}$$

where

$$K = \gamma\varphi\left[\overline{X}(\alpha) - \alpha(2\lambda-1)\right] + (1-\gamma)\varphi\left[\overline{X}(\alpha) + \alpha(2\lambda-1)\right]$$

Consider first the case $\gamma < \frac{1}{2}$. Hence, $\overline{X}(\alpha) > 0$ and $\varphi\left[\overline{X}(\alpha) - \alpha(2\lambda-1)\right] > \varphi\left[\overline{X}(\alpha) + \alpha(2\lambda-1)\right]$. Since $\frac{\partial \overline{X}(\alpha;\gamma)}{\partial \gamma} < 0$, it follows that the numerator of (3) is positive and therefore $\frac{\partial \alpha^s}{\partial \gamma} > 0$ if $-\frac{\left(\overline{X}(\alpha)\right)^2}{\alpha} + \alpha(2\lambda-1)^2 > 0$. The latter can be rewritten as

$$\alpha(2\lambda-1) > \overline{X}(\alpha). \tag{4}$$

Since $V(\alpha)$ must be concave at α^s and since (4) coincides with the region in which $V\left(\alpha; \gamma < \frac{1}{2}\right)$ is concave, it follows that $\frac{\partial \alpha^s}{\partial \gamma} > 0$.

Consider now $\gamma > \frac{1}{2}$ and therefore $\overline{X}(\alpha) < 0$ and $\varphi\left[\overline{X}(\alpha) - \alpha(2\lambda-1)\right] < \varphi\left[\overline{X}(\alpha) + \alpha(2\lambda-1)\right]$. The numerator is now negative and $\frac{\partial \alpha^s}{\partial \gamma} < 0$ if $-\frac{\left(\overline{X}(\alpha)\right)^2}{\alpha} + \alpha(2\lambda-1)^2 > 0$. The latter can be rewritten as

$$-\alpha(2\lambda-1) < \overline{X}(\alpha),$$

which again coincides with the concavity region for $V(\alpha)$.

For the special case $\gamma = \frac{1}{2}$ we know that $\overline{X}(\alpha;\gamma) = 0$ for any $\alpha > 0$. This allows us to calculate the social optimum explicitly by solving

$$\frac{1}{2}(V_h - V_l)(2\lambda-1)\varphi(\alpha^s(2\lambda-1)) = c.$$

QED

The results are intuitive. When profitability is either very high ($\gamma > 1 - \gamma_0$) or very low ($\gamma < \gamma_0$), it is optimal not to produce any information. Here, following the ex-ante optimal investment decision provides a better outcome than spending resources on information production and improving the investment decision,

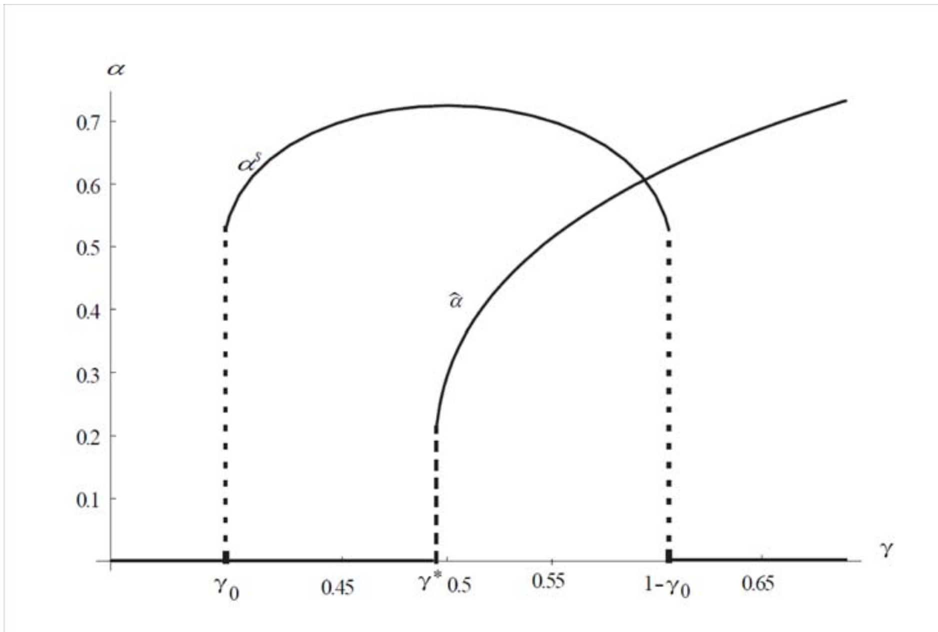

Figure 1: Illustrates the socially optimal level of information production α^s and the equilibrium $\widehat{\alpha}$ as a function of γ. Parameter values are $c = 0.3, \lambda = 0.9, \sigma^2 = 0.7, V_h - V_l = 2$.

since the prior provides very strong indication about which investment should be undertaken. From the expression for γ_0, we see that the size of the range at which no information production is optimal increases in the cost of information production c and in the amount of noise trading σ^2. The latter is true because when there is more noise trading, information is less useful in guiding investment. Moreover, the size of the range at which no information production is optimal decreases when the information is more precise (λ is higher) and when there is more uncertainty about the risky project ($V_h - V_l$ is greater).

Note that if parameters are such that $\gamma_0 > \frac{1}{2}$, then the social planner would not produce any information so that $\alpha^s = 0$ for any value of γ. Otherwise, positive information production is desirable. The maximum amount of information should be produced when $\gamma = \frac{1}{2}$, which is the point where the prior information provides the weakest indication about the optimal decision, and so information acquisition is very desirable. The amount of information produced in the constrained efficient solution will decrease as we move away from $\gamma = \frac{1}{2}$ and information becomes less important in guiding the investment decision.

The constrained efficient amount of information is illustrated in Figure 1. The figure also shows the equilibrium amount of information derived before: It is zero below γ^* and increases in γ above γ^*. It is not difficult to see that the constrained efficient solution does not coincide with the equilibrium outcome. In equilibrium, more information is produced as γ increases, whereas the constrained efficient solution is to produce more information when γ is close to $\frac{1}{2}$. The following corollary describes two observations about the deviation of the equilibrium outcome from constrained efficiency.

Corollary 1.1 *When $\gamma > 1-\gamma_0$ then there may be too much information production in equilibrium. When $\gamma < \gamma^*$ there may be too little information production in equilibrium.*

The first statement follows directly from the fact that $\alpha^s = 0$ for $\gamma > 1-\gamma_0$, while a strictly positive amount of information may be produced in equilibrium. The second statement follows from the observation that for some $\gamma^* < \frac{1}{2}$ equilibrium information production drops to zero, but if $\gamma^* > \gamma_0$ then the social planner would like some information to be produced. The numerical example of Figure 1 shows that there are parameter values for which this occurs.

1.1 Welfare ranking of multiple equilibria

When $\gamma \in \left[\gamma^*, \frac{1}{2}\right)$ there may be multiple equilibria, as shown in Proposition 2. This raises the question whether the equilibria can be welfare ranked. The answer is that they cannot be ranked in general. To see this, it is sufficient to consider numerical examples. The example provided in the figure above shows that the socially optimal level of information production exceeds the highest equilibrium amount of information in the relevant region. This suggests that for this example the equilibrium with more information is better than the one with no information. But, the opposite can also happen. In particular, for sufficiently high values of σ^2, the social value of information drops, while trading profits are significant leading to excessive information production in the most informative equilibrium.[2]

1.2 Policy implications

Consider the parameter region when there is too little information production, specifically, when $\gamma < \frac{1}{2}$ and $\widehat{\alpha} = 0$. There are several policies that could encourage information production and therefore improve welfare, for example a subsidy on share trading or a subsidy on investment.

Since information production cannot be observed, it cannot be subsidized directly. Therefore, the subsidy would have to be contingent on trades or trading profits. For example, the government could directly subsidize trading activity. This, however, raises the possibility of spurious trading so as to increase the payments under the subsidy. If trading profits, but not losses, are subsidized this raises the same problem of encouraging spurious trading activity. A subsidy for informed trading activity may therefore be difficult to implement in practical terms.

We therefore now study in more detail a policy that provides a subsidy τ to the firm if it takes the risky investment project.[3] From the perspective of the firm an investment subsidy raises γ to

$$\gamma_\tau \equiv \gamma + \frac{\tau}{V_h - V_l}.$$

[2] For example, for the following parameters: $\gamma = 0.49, c = 0.3, V_h - V_l = 2, \sigma^2 = 1.1$ we get $\alpha^s = 0$ and $\widehat{\alpha}_3 > 0$.

[3] We assume that the government can distinguish between the firm's actions $A = 0$ and $A = 1$. In practice this may correspond to subsidizing new investments, but not investment that goes into the replacement of depreciated capital.

In order to develop the argument formally, denote the new cut-off level of investment by

$$\overline{X}_{\tau}(\alpha) = \frac{\sigma^2}{2\alpha(2\lambda-1)} \ln \frac{1-\gamma_{\tau}}{\gamma_{\tau}}.$$

A tax reduces this cut-off level, i.e., $\overline{X}_{\tau}(\alpha) < \overline{X}(\alpha)$. Moreover, denote by $\widehat{\alpha}_{\tau}$ the equilibrium level of information production when the cut-off employed by the firm corresponds to $\overline{X}_{\tau}(\alpha)$. From the previous analysis it is clear by now that $\widehat{\alpha}_{\tau} > \widehat{\alpha}$ unless both are at a corner.

Since a tax is purely redistributive, the social planner cares about the expected firm value before any subsidies are paid. This can be written as

$$\begin{aligned} V_{\tau} &= V_0 + \frac{1}{2}(V_h - V_l) \int_{\overline{X}_{\tau}(\widehat{\alpha}_{\tau})}^{\overline{X}(\widehat{\alpha}_{\tau})} \left[\gamma\varphi(x - \widehat{\alpha}_{\tau}(2\lambda-1)) - (1-\gamma)\varphi(x + \widehat{\alpha}_{\tau}(2\lambda-1))\right] dx \qquad (5) \\ &+ \frac{1}{2}(V_h - V_l) \int_{\overline{X}(\widehat{\alpha}_{\tau})}^{\overline{X}(\widehat{\alpha})} \left[\gamma\varphi(x - \widehat{\alpha}_{\tau}(2\lambda-1)) - (1-\gamma)\varphi(x + \widehat{\alpha}_{\tau}(2\lambda-1))\right] dx \\ &+ \frac{1}{2}(V_h - V_l) \int_{\overline{X}(\widehat{\alpha})}^{\infty} \left[\gamma\varphi(x - \widehat{\alpha}_{\tau}(2\lambda-1)) - (1-\gamma)\varphi(x + \widehat{\alpha}_{\tau}(2\lambda-1))\right]. \end{aligned}$$

Note that by definition of $\overline{X}(\widehat{\alpha}_{\tau})$ we have

$$\gamma\varphi\left(\overline{X}(\widehat{\alpha}_{\tau}) - \widehat{\alpha}_{\tau}(2\lambda-1)\right) = (1-\gamma)\varphi\left(\overline{X}(\widehat{\alpha}_{\tau}) + \widehat{\alpha}_{\tau}(2\lambda-1)\right).$$

It follows that the first integral in (5) is negative. This captures the distortion introduced by an investment subsidy: the firm now invests in cases where it would not have invested in the absence of a subsidy. In that region, the expected payoff before payment of the subsidy from investing in the risky project is actually lower than from investing in the riskless project. Investing is therefore socially undesirable. The second line captures the increase in firm value that results from the additional information that is available due to the subsidy. Since $\widehat{\alpha}_{\tau} > \widehat{\alpha}$ this is always positive. The final line captures the expected firm value that would be obtained in the absence of an investment subsidy. We can now show the following.

Proposition 1.2 *There are parameters $\gamma < \gamma^*$ for which it is socially desirable to pay an investment subsidy.*

Proof of Proposition 1.2: We first prove that V_{τ} may be increasing in τ. Suppose $\gamma = \gamma^* - \varepsilon$, i.e., $\widehat{\alpha} = 0$. Let τ be just high enough that $\gamma_{\tau} = \gamma^*$ and therefore $\widehat{\alpha}_{\tau} > 0$. There always exists a ε small enough such that $\overline{X}_{\tau}(\widehat{\alpha}_{\tau})$ is arbitrarily close to $\overline{X}(\widehat{\alpha}_{\tau})$. Note also that $\overline{X}(\widehat{\alpha}) = \infty$ and therefore the second integral of (5) dominates the first. Next we need to show that the increase in V_{τ} exceeds the information production cost. That this can be the case follows from the numerical example contained in Figure 1.QED

2 Endogenous uncertainty and belief dispersion

For any given realization of order flow X define[4]

$$D(X;\widehat{\alpha}) = \Pr(\omega = h|X, s = h) - \Pr(\omega = h|X, s = l).$$

The following can then be shown.

Lemma 2.1 *Belief dispersion is decreasing in γ, i.e.,*

$$\frac{dD(X;\widehat{\alpha})}{d\gamma} < 0 \quad \forall X.$$

Proof of Lemma 2.1: Calculating the Bayesian updates, we can write D as a function of order flow X and the equilibrium fraction of informed traders $\widehat{\alpha}$:

$$\begin{aligned} D(X;\widehat{\alpha}) &= \frac{\lambda\varphi(X-\widehat{\alpha}(2\lambda-1))}{\lambda\varphi(X-\widehat{\alpha}(2\lambda-1))+(1-\lambda)\varphi(X+\widehat{\alpha}(2\lambda-1))} \\ &\quad -\frac{(1-\lambda)\varphi(X-\widehat{\alpha}(2\lambda-1))}{(1-\lambda)\varphi(X-\widehat{\alpha}(2\lambda-1))+\lambda\varphi(X+\widehat{\alpha}(2\lambda-1))}. \end{aligned} \tag{6}$$

Since $\frac{\partial\widehat{\alpha}}{\partial\gamma} > 0$ and $\frac{\partial D(X;\widehat{\alpha})}{\partial\gamma} = 0$ it is sufficient to show that $\frac{\partial D(X;\widehat{\alpha})}{\partial\widehat{\alpha}} < 0$. Using (6) and simplifying we can write

$$\frac{\partial D(X;\widehat{\alpha})}{\partial\widehat{\alpha}} = \frac{2\lambda(1-\lambda)(2\lambda-1)^2\varphi_l\varphi_h}{\sigma^2(\lambda\varphi_h+(1-\lambda)\varphi_l)^2((1-\lambda)\varphi_h+\lambda\varphi_l)^2}X\left(\varphi_l^2-\varphi_h^2\right),$$

where $\varphi_l = \varphi(X+\widehat{\alpha}(2\lambda-1))$ and $\varphi_h = \varphi(X-\widehat{\alpha}(2\lambda-1))$. The sign of the expression depends on the sign of $X\left(\varphi_l^2-\varphi_h^2\right)$. Since $\varphi_l^2 > \varphi_h^2 \Leftrightarrow X < 0$ the result follows immediately. QED

[4] Since X is a sufficient statistic for P we drop P in the conditioning