Data Abundance and Asset Price Informativeness

On-Line Appendix

Not Intended for Publication

March 25, 2017

This note is the on-line appendix for "Data Abundance and Asset Price Informativeness." It contains the proofs for findings mentioned in the paper but not proven there for the sake of brevity. It is not intended for publication. The on-line appendix is organized as follows.

- Section 1 derives the equilibrium demand for the raw signal (Lemma 2 in the paper).
- Section 2 analyzes the case in which the upper bound on the mass of speculators, i.e., $\bar{\alpha}$, is less than 2.
- Section 3 analyzes the case in which speculators can make their decision to buy the processed signal contingent on the realization of the price at date 1.
- Section 4 completes Proposition 3 (case 2).
- Section 5 completes Proposition 5 by considering the case in which $C_p > C_{min}(\theta, \alpha_1^e)$.
- Section 6 shows that $C_{max}(\theta, \alpha_1)$ decreases with α_1 as claimed in the proof of Proposition 4.

1 Equilibrium demand for the raw signal (Proof of Lemma 2).

Let $\pi_1^{gross,a}(\alpha_1) = \alpha_1 \bar{\pi}_1(\alpha_1) = \alpha_1 \max\left\{\frac{\theta}{2}(1-\alpha_1), 0\right\}$ be the aggregate gross expected profit for speculators who receive the raw signal. We represent by α_1^e the equilibrium value of α_1, the demand for the raw signal at date 1. Proceeding as in the market for the processed signal, we deduce that if $\alpha_1^e > 0$ then α_1^e solves:

$$\pi_1^{gross,a}(\alpha_1^e) = \alpha_1^e \max\left\{\frac{\theta}{2}(1-\alpha_1^e), 0\right\} = C_r. \tag{1.1}$$

As, $\pi_1^{gross,a}(\alpha_1)$ reaches its maximum for $\alpha_1 = 1/2$, we deduce that the previous equation has no solution if $C_r > \frac{\theta}{8}$. In this case, for all values of α_1, $\pi_1^{gross,a}(\alpha_1) < C_r$ and therefore there is no fee at which trades between the seller of the raw signal and buyers of this signal are mutually beneficial. Thus, in this case, $\alpha_1^e = 0$.

For $C_r < \frac{\theta}{8}$, eq.(1.1) has two solutions in $(0, 1)$. As explained in the text, we select the highest as the equilibrium demand since it yields the lowest fee charged by the information seller. This solution is:

$$\alpha_1^e = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{2C_r}{\theta}}, \tag{1.2}$$

and the corresponding equilibrium fee is therefore $F_1^e = C_r/\alpha_1^e$. Last, if $C_r = \frac{\theta}{8}$, the only solution to eq.(1.1) is $\alpha_1^e = 0$.

2 The case in which $\bar{\alpha} \leq 2$.

In equilibrium speculators' demands for each type of signal (i.e., α_1^e and α_2^e) must satisfy:

$$0 \leq \alpha_1^e \leq \bar{\alpha}, \text{and} \quad 0 \leq \alpha_2^e \leq \bar{\alpha}.$$

When $\bar{\alpha} > 2$, the right hand sides of these constraints are never binding (as $\alpha_2^e \leq 2$ and $\alpha_1^e \leq 1$). This is the case on which we focused in the baseline version of the model. In contrast, when $\bar{\alpha} \leq 2$, these constraints can be binding, in which case one obtains a corner solution for equilibrium demands. We discuss each case in turn below. Note that a corner solution obtains for the demand for the raw signal iff $C_r > \frac{\theta}{8}$ (in which case $\alpha_1^e = 0$) or $C_r < \hat{C}_r$ (in which case $\alpha_1^e = 1$), where $\hat{C}_r = \frac{\theta}{2}(\frac{1}{4} - (\min\{(\bar{\alpha} - 0.5), 0\})^2)$.

1. **Case 1.** The zero profit conditions eq.(12) in the text and eq.(1.1) in this appendix have solutions α_1^e and α_2^e such that $0 < \alpha_1^e < \bar{\alpha}$ and $0 < \alpha_2^e < \bar{\alpha}$. In this case, the equilibria of the markets for the raw signal and the processed signal are interior. This case is identical to that analyzed in the paper and all results are identical to those obtained when $\bar{\alpha} > 2$.

2. **Case 2.** The equilibrium demand for (i) the raw signal is a corner solution, that is $\alpha_1^e = 0$ (if $C_r > \frac{\theta}{8}$) or $\alpha_1^e = \bar{\alpha}$ (if $C_r < \hat{C}_r$) and (ii) the equilibrium demand for the processed signal is interior ($0 < \alpha_2^e < \bar{\alpha}$). In this case, the cost of producing the raw signal C_r has (locally) no impact on α_1^e and thus has no local effect on α_2^e. Therefore, if $C_r < \hat{C}_r$ or $C_r > \frac{\theta}{8}$ then a small change in the cost of the raw signal has no effect on asset price informativeness or trade and price patterns. For other values of C_r, results are identical to that in the text.

3. **Case 3.** The equilibrium demand for the processed signal is a corner solution, that is $\alpha_2^e = 0$ or $\alpha_2^e = \bar{\alpha}$. Thus, α_2^e does not depend on α_1^e since its value is either zero or $\bar{\alpha}$. Therefore, a change in cost of the raw signal C_r has (locally) no impact on α_2^e. Thus, it does not affect price informativeness in the long run.

3 Speculators can buy the processed signal at date 2.

In this section, we consider the case in which speculators can buy the processed signal *after* observing the price of the asset at date 1, i.e., just before trading at date 2. The equilibrium of the market for the raw signal is unchanged in this case and is still given by Proposition 2. The main difference with the case considered in the text is that (i) the demand for the processed signal at date 2, α_2, depends on p_1, the price of the asset at the end of the first period and (ii) the seller the processed signal can charge a different fee for this signal depending on the price realized at date 1.

We consider two subcases. In the first case (Section 3.1), we assume that the seller of the processed signal pays the fixed cost of information processing *after* observing the realization of the price at date 1. In the second case (Section 3.2), we assume that this cost is paid *before* observing the realization of the price at date 1. The break-even conditions for the information seller differ in each case. In the second case, the information seller must set its fees for the processed signal so that it covers its fixed cost of producing the signal *on average across the possible realizations for the price at date* 1. In the first case, the information seller sets its fee so that it breaks even conditional on *each* realization of the price at date 1. The two cases yield an identical equilibrium outcome when there is a demand for the processed signal for each realization of the price at date 1. Equilibrium outcomes (e.g., the fees for the processed signal) differ when the demand for the processed signal is nil for some realization of the price at date 1. However, we show that the implications obtained in the baseline model (about price informativeness and price and trade patterns) are preserved in each case.

3.1 Investment in information processing takes place after the realization of the price at date 1

At the end of the first period, there are three possible outcomes (see Panel A of Figure 3 in the text) : (i) the asset price has not changed, $p_1 = p_0$; (ii) the asset price has increased $p_1 = \frac{1+\theta}{2}$; (iii) the asset price has decreased $p_1 = \frac{1-\theta}{2}$. Due to the symmetry of the model, the expected profit from trading on the processed signal is the same in the last two cases. Thus, the decision to buy the processed signal is identical whether the price has increased or decreased during the first period. This means that only two states are relevant for the analysis of the market for information at date 2: either (i) the price has not changed in the first period ($p_1 = p_0$) or (ii) the price has changed ($p_1 \neq p_0$). We denote by $(\alpha_2^{nc}, F_2^{nc})$ the equilibrium of the market for information in the first case and by (α_2^{c}, F_2^{c}) the equilibrium of this market in the second case (superscript "c" stands for

”change” while superscript ”nc” stands for no change).

Case 1: $p_1 \neq p_0$**.** We first derive the equilibrium of the market for the processed signal when $p_1 \neq p_0$, i.e., (α_2^c, F_2^c). As in the baseline model, (α_2^c, F_2^c) must satisfy the following zero profit conditions when $\alpha_2^c > 0$:

$$\text{Zero profit for speculators:} \quad \bar{\pi}_2^{net,c}(\alpha_2^c, F_2^c, \theta) = \pi_2^c(\alpha_2^c, \theta) - F_2^c = 0. \tag{3.1}$$

$$\text{Zero profit for the information seller:} \quad \bar{\Pi}_2^{seller,c}(\alpha_2^c, F_2^c) = \alpha_2^c \times F_2^c - C_p = 0, \tag{3.2}$$

where superscript c for expected profits indicates that these profits are computed conditional on a price *change* in the first period. If there is no solution (α_2^c, F_2^c) to this system of equations for which $\alpha_2^c > 0$ then the market for the processed signal is inactive when $p_1 \neq p_0$ and we set $\alpha_2^c = 0$.

As $p_1 \neq p_0$, the *gross* expected profit (per speculator) from trading on the processed signal is (see Step 3 in the proof of Proposition 2 in the text for a derivation):

$$\pi_2^c(\alpha_2^c) = \theta(1-\theta)(1-\alpha_2^c).$$

Thus, in *aggregate*, speculators’ net expected profit is:

$$\pi_2^{net,c,a}(\alpha_2^c, C_p) = \alpha_2^c \times (\pi_2^c(\alpha_2^c) - F_2^c) = \alpha_2^c\theta(1-\theta)(1-\alpha_2^c) - C_p,$$

where the second equality follows from eq.(3.2). In equilibrium, if $\alpha_2^c > 0$, speculators’ net expected profit must be equal to zero (see eq.(3.1)). Thus, we deduce from the previous equation that α_2^c solves:

$$\pi_2^{net,a,c}(\alpha_2^c, C_p) = \alpha_2^c\theta(1-\theta)(1-\alpha_2^c) - C_p = 0. \tag{3.3}$$

As in the baseline model, if eq.(3.3) has multiple positive solutions then we retain the highest because it yields the smallest (hence most competitive) fee for the information seller. If eq.(3.3) has no positive solution then the market for information for the processed signal is inactive (i.e., $\alpha_2^c = 0$) when $p_1 \neq 1/2$. Solving for eq.(3.3), we obtain:

$$\alpha_2^c(\theta, C_p) = \begin{cases} \frac{1}{2} + \left(\frac{1}{4} - \frac{C_p}{\theta(1-\theta)}\right)^{\frac{1}{2}} & \text{if } 0 \leq C_p \leq \frac{\theta(1-\theta)}{4}, \\ 0 & \text{if } C_p > \frac{\theta(1-\theta)}{4}, \end{cases} \tag{3.4}$$

It follows from eq.(3.2) that the equilibrium fee for the processed signal is $F_2^c = \frac{C_p}{\alpha_2^c}$ when $\alpha_2^c > 0$.

Case 2: $p_1 = p_0 = 1/2$. In this case, there is no change in the price at date 1. The net expected profit (per speculator) from trading on the processed signal is then (see Step 3 in the proof of Proposition 2 in the text for a derivation):

$$\pi_2^{nc}(\alpha_2) = \begin{cases} \frac{\theta}{2(2-\theta)}\left(2 - \theta - \alpha_2\right) & \text{if } \alpha_2 \leq 1, \\ \frac{\theta}{2}\frac{1-\theta}{2-\theta}\left(2 - \alpha_2\right) & \text{if } \alpha_2 > 1, \end{cases} \tag{3.5}$$

where superscript *nc* for the expected profit indicates that it is computed conditional on no price change in the first period. We can then solve for the equilibrium demand for the processed signal as in the case in which $p_1 \neq 1/2$. After some algebra, we obtain:

$$\alpha_2^{nc}(\theta, C_p) = \begin{cases} 1 + \left(1 - \frac{2(2-\theta)}{\theta(1-\theta)}C_p\right)^{\frac{1}{2}} & \text{if } C_p \leq \frac{\theta}{2}\frac{1-\theta}{2-\theta}, \\ \frac{2-\theta}{2} + \left(\frac{(2-\theta)^2}{4} - \frac{2(2-\theta)}{\theta}C_p\right)^{\frac{1}{2}} & \text{if } \frac{\theta}{2}\frac{1-\theta}{2-\theta} < C_p \leq \frac{\theta(2-\theta)}{8} \\ 0 & \text{if } C_p > \frac{\theta(2-\theta)}{8}, \end{cases} \tag{3.6}$$

Thus, when $p_1 = 1/2$, it follows from eq.(3.2) that the equilibrium fee for the processed signal is $F_2^{nc} = \frac{C_p}{\alpha_2^{nc}}$ when $\alpha_2^{nc} > 0$. .

Using the expressions for α_2^c and α_2^{nc} (eq.(3.4) and eq.(3.6)), it is easily shown that $\alpha_2^c < \alpha_2^{nc}$ when $C_p > 0$. Thus, the demand for processed information is smaller when prices have changed at date 1 than when they have not. In equilibrium, the expected demand for the processed signal in equilibrium is:

$$E(\alpha_2^e) = Pr(p_1 \neq p_0)\alpha_2^c + (1 - Pr(p_1 \neq p_0))\alpha_2^{nc} = \alpha_1\alpha_2^c + (1 - \alpha_1)\alpha_2^{nc},$$

where α_2^e denotes the realization of the demand for the processed signal at date 2 (i.e., α_2^c or α_2^{nc}). An increase in the demand for the raw signal increases the probability that the price changes at date 1 since $Pr(p_1 \neq p_0) = \alpha_1$. As the demand for the processed signal is smaller when the asset price changes at date 1 than when it does not ($\alpha_2^c < \alpha_2^{nc}$), we obtain the following result.

Proposition 3.1. *The expected demand for the processed signal in equilibrium decreases with the demand for the raw signal, that is,* $\frac{\partial E(\alpha_2^e)}{\partial \alpha_1} < 0$. *Thus, a decrease in the cost of the raw signal reduces the expected demand for the processed signal in equilibrium.*

This result is the analog of Proposition 3 in the baseline model. It is stronger in the sense that it holds for all parameter values. Figure 3.1 illustrates the previous proposition for specific parameter values.

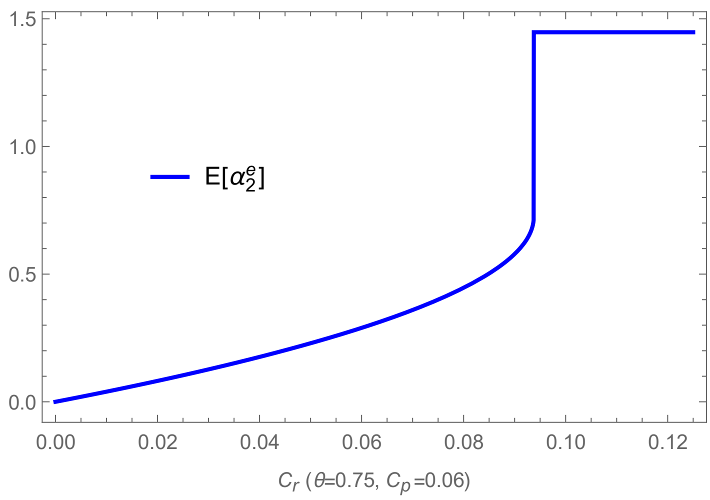

Figure 3.1: Expected demand for the processed signal (E(α_2^e))

Asset price informativeness at date t=2. We now study the effect of a reduction in the cost of producing the raw signal on the informativeness of the price at date 2. As in the baseline model, we define asset price informativeness at date 2 as:

$$\mathcal{E}_2(C_r, C_p) = \frac{1}{4} - E[(\tilde{V} - p_2^*)^2], \tag{3.7}$$

where p_2^* denotes the realization of the equilibrium price at date 2. We obtain the following result.

Proposition 3.2. *If* $C_p \leq \frac{\theta(2-\theta)}{8}$ *then*

1. *When* $\theta \leq 1/2$*, a decrease in the cost of producing the raw signal reduces asset price informativeness at date* 2*.*

2. *When* $\theta > 1/2$*, a decrease in the cost of producing the raw signal reduces asset price informativeness at date* 2 *iff* $C_p \leq \bar{C}_p$ *where* $\bar{C}_p$ *is a threshold that belongs to* $\left[\frac{\theta(1-\theta)}{2(2-\theta)}, \frac{\theta(2-\theta)}{8}\right]$

Thus, when $C_p < \frac{\theta(2-\theta)}{8}$, the informativeness of the price at date 2 can decline when C_r declines as emphasized by Proposition 5 in the paper. Figure 3.2 illustrates this claim for specific parameter values.

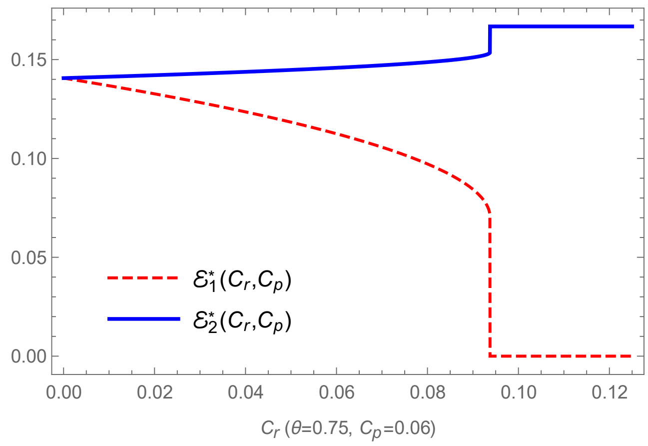

Figure 3.2: The informativeness of the price at date 2

When $C_p > \frac{\theta(2-\theta)}{8}$, the cost of producing the processed signal is so high that there is no fee at which it can be profitably produced (i.e., a which buyers and sellers of the processed signal can mutually trade), for all possible realizations of the price at date 1. Thus, the processed signal is not produced and the informativeness of the price at date 2 is equal to that at date 1. The latter increases when α_1 increases and therefore, in this case, a reduction in the cost of the raw signal raises price informativeness at date 2 as well.

Price and trade patterns. If $C_r \leq \theta/8$, some speculators buy the raw signal (see Lemma 2 in the text of the paper) . Using Proposition 1 in the paper, we can write the equilibrium strategy of a speculator receiving the raw signal, s, as

$$x_1^*(s, C_r) = \mathbb{I}_{s=1} - \mathbb{I}_{s=0} = u \times [\mathbb{I}_{V=1} - \mathbb{I}_{V=0}] + (1-u) \times [\mathbb{I}_{\epsilon=1} - \mathbb{I}_{\epsilon=0}]. \tag{3.8}$$

If $C_r > \frac{\theta}{8}$, no speculator buys the raw signal in equilibrium and therefore $x_1^*(s, C_r) = 0$. In sum:

$$x_1^*(s, C_r) = \begin{cases} \mathbb{I}_{s=1} - \mathbb{I}_{s=0} = u \times [\mathbb{I}_{V=1} - \mathbb{I}_{V=0}] + (1-u) \times [\mathbb{I}_{\epsilon=1} - \mathbb{I}_{\epsilon=0}] & \text{if } 0 \leq C_r \leq \theta/8, \\ 0 & \text{if } C_r > \theta/8. \end{cases} \tag{3.9}$$

Similarly, using Proposition 2 in the paper, we can write the optimal trading strategy of speculators at date 2 as:

$$x_2^*(p_1, u, C_p) = \begin{cases} u \times [\mathbb{I}_{V=1} - \mathbb{I}_{V=0}] + (1-u) \times [\mathbb{I}_{p1=(1-\theta)/2} - \mathbb{I}_{p1=(1+\theta)/2}] & \text{if } 0 \leq C_p \leq \frac{\theta(1-\theta)}{4}, \\ u \times [\mathbb{I}_{V=1} - \mathbb{I}_{V=0}] \times \mathbb{I}_{p_1=1/2} & \text{if } \frac{\theta(1-\theta)}{4} < C_p \leq \frac{\theta(2-\theta)}{8}, \\ 0 & \text{if } C_p > \frac{\theta(2-\theta)}{8}, \end{cases} \tag{3.10}$$

where the last equality follows from the fact that there is no demand for the processed signal if $C_p > \frac{\theta(2-\theta)}{8}$.

Now, assume $C_r \leq \frac{\theta}{8}$ and $C_p \leq \frac{\theta(2-\theta)}{8}$. These assumptions guarantee that some speculators trade at date 1 ($x_1^* \neq 0$) and that some speculators trade at date 2 at least if $p_1 = 1/2$. Using eq.(3.9) and eq.(3.10), we obtain:

$$Cov(x_1^*, x_2^*) = \begin{cases} \theta(1 - \alpha_1^e), \text{ if } C_p > \frac{\theta(1-\theta)}{4}, \\ \theta - (1-\theta)\alpha_1^e, \text{ if } C_p \leq \frac{\theta(1-\theta)}{4}. \end{cases}$$

Thus, as implied by Corollary 4 in the baseline model, we obtain that $Cov(x_1^*, x_2^*)$ declines when α_1^e increases, i.e., when the cost of the raw signal decreases. Moreover, when $C_p \leq \frac{\theta(1-\theta)}{4}$, one obtains exactly the same expression for the covariance as that in Corollary 4 in the baseline model. Thus, as in this case, $Cov(x_1^*, x_2^*)$ becomes negative if $\theta < 1/2$ and

C_r is small enough.

Using the expression for the equilibrium price at date 1 (Proposition 1 in the paper) and eq.(3.10), we also obtain

$$Cov(r_1, x_2^*) = E\left[\left(p_1 - \frac{1}{2}\right) x_2\right] = \begin{cases} 0, \text{ if } C_p > \frac{\theta(1-\theta)}{4}, \\ \theta(2\theta - 1)\alpha_1^e, \text{ if } C_p \leq \frac{\theta(1-\theta)}{4} \text{ (as in the paper)}, \end{cases}$$

where $r_1 = p_1^* - p_0$ is the return from date 0 to date 1. Thus, when $C_p \leq \frac{\theta(1-\theta)}{4}$ (so that some speculators buy the processed signal whether the price at date 1 has changed or not), we obtain the same expression for $Cov(r_1, x_2^*)$ as that in Corollary 5 in the paper. Otherwise $Cov(r_1, x_2^*)$ is zero. Hence, Corollary 5 holds true even if speculators can make their decision to buy the processed signal contingent on the price realized at date 1.

Last consider $Cov(x_1^*, r_2)$ where $r_2 = (p_2^* - p_1^*)$ is the return from date 1 to date 2. Calculations yield:

$$Cov(x_1^*, r_2) = \begin{cases} \frac{\theta(1-\alpha_1^e)\alpha_2^{nc}}{2(2-\theta)}, \text{ when } C_p > \frac{\theta(1-\theta)}{2(2-\theta)}, \\ \frac{\theta(1-\alpha_1^e)(1+(1-\theta)(\alpha_2^{nc}-1))}{2(2-\theta)}, \text{ when } C_p \leq \frac{\theta(1-\theta)}{2(2-\theta)}. \end{cases}$$

Observe that this is the same expression as that obtained in the baseline model (see eq.(19)), with α_2^{nc} replacing α_2^e. The intuition is the following. In equilibrium, the innovation in $r_2 = (p_2^* - p_1^*)$ is orthogonal to dealers' information set set at $t = 1$ since equilibrium prices follow a martingale. If the order flow at date 1 reveals the raw signal then x_1^* is known to dealers at date 1 and therefore $r_2 = p_2^* - p_1^*$ is orthogonal to x_1^*. This means that if $p_1^* \neq 1/2$ then r_2 is orthogonal to x_1^*. Thus, the covariance between x_1^* and r_2 is only driven by the cases in which $p_1^* = 1/2$, in which case, the demand for the processed signal is α_2^{nc}. This explains why only α_2^{nc} affects $Cov(x_1^*, r_2)$.

As α_2^{nc} does not depend on α_1^e, we obtain that $Cov(x_1, r_2)$ decreases when α_1^e increases, i.e., when C_r decreases, for the same reason as in the baseline model.

Proofs for Section 3.1

Proof of Proposition Proof 3.1. When $C_p > \frac{\theta(2-\theta)}{8}$, the demand for the processed signal is nil whether the price changes at date 1 or not. Thus, $E[\alpha_2^e] = 0$ in this case. When $\frac{\theta(2-\theta)}{8} \geq C_p > \frac{\theta(1-\theta)}{4}$, the demand for the processed signal is nil if the price changes at date 1 and strictly positive otherwise. Thus, in this case, $E[\alpha_2^e] = (1-\alpha_1)\alpha_2^{nc}$, which is clearly decreasing with α_1. Finally, when $C_p \leq \frac{\theta(1-\theta)}{4}$, we have, using the expressions for the equilibrium demand for the processed signal:

$$E[\alpha_2^e] = (1-\alpha_1)\left[1 + \left(1 - \frac{2(2-\theta)}{\theta(1-\theta)}C_p\right)^{\frac{1}{2}}\right] + \alpha_1\left[\frac{1}{2} + \left(\frac{1}{4} - \frac{C_p}{\theta(1-\theta)}\right)^{\frac{1}{2}}\right],$$

which decreases with α_1 since $\left(\frac{1}{4} - \frac{C_p}{\theta(1-\theta)}\right)^{\frac{1}{2}} \leq 1/2$.

Proof of Proposition Proof 3.2. First, observe that $\mathcal{E}_2(C_r, C_p)$ is inversely and linearly related to $E[(\tilde{V} - p_2^e)^2]$. It is easily shown that $E[(\tilde{V} - p_2^*)^2] = E[p_2^*(1-p_2^*)]$. To obtain the proposition, we can therefore analyze how $E[p_2^*(1-p_2^*)]$ varies with C_r.

We first compute $E[p_2^*(1-p_2^*)]$. If $\alpha_2^{nc} \leq 1$, we obtain (e.g., using Figure 3, in the paper):

$$\begin{aligned}
E[p_2^*(1-p_2^*)] &= \alpha_1 \times \left[(1-\theta)\alpha_2^c \times \frac{1}{4} + (1-\alpha_2^c) \times \frac{(1-\theta)(1+\theta)}{4}\right] \\
&\quad + (1-\alpha_1) \times \left[(1-\alpha_2^{nc}) \times \frac{1}{4} + \left(\frac{\alpha_2^{nc}}{2} + \frac{(1-\theta)\alpha_2^{nc}}{2}\right) \times \frac{1-\theta}{(2-\theta)^2}\right] \\
&= \frac{1}{4}\alpha_1[1-\theta + (1-\theta)\theta(1-\alpha_2^c)] + \frac{1}{4}(1-\alpha_1)\left[1 - \left(1 - 2(2-\theta)\frac{1-\theta}{(2-\theta)^2}\right)\alpha_2^{nc}\right] \\
&= \frac{1}{4}\alpha_1[1-\theta + (1-\theta)\theta(1-\alpha_2^c)] + \frac{1}{4}(1-\alpha_1)\left(1 - \frac{\theta}{2-\theta}\alpha_2^{nc}\right) \\
&= \frac{1}{4}\alpha_1(1-\theta) + \frac{1}{4}\alpha_1(1-\theta)\theta(1-\alpha_2^c) + \frac{1}{4}(1-\alpha_1)\theta\left(1 - \frac{1}{2-\theta}\alpha_2^{nc}\right) + \frac{1}{4}(1-\alpha_1)(1-\theta) \\
&= \frac{1}{4}(1-\theta) + \frac{1}{4}\alpha_1\pi_2^c(\alpha_2^c) + \frac{1}{2}(1-\alpha_1)\pi_2^{nc}(\alpha_2^{nc}).
\end{aligned}$$

Similarly, if $1 \leq \alpha_2^c \leq 2$, we obtain:

$$
\begin{aligned}
E[p_2^*(1-p_2^*)] &= \alpha_1 \times \left[(1-\theta)\alpha_2^c \times \frac{1}{4} + (1-\alpha_2^c) \times \frac{(1-\theta)(1+\theta)}{4} \right] \\
&+ (1-\alpha_1) \times \left[(1-\theta)(\alpha_2^{nc}-1) \times \frac{1}{4} + \left(\frac{2-\alpha_2^{nc}}{2} + \frac{(1-\theta)(2-\alpha_2^{nc})}{2} \right) \times \frac{1-\theta}{(2-\theta)^2} \right] \\
&= \frac{1}{4}\alpha_1(1-\theta) + \frac{1}{4}\alpha_1 \pi_2^c(\alpha_2^c) \\
&+ \frac{1}{4}(1-\alpha_1)(1-\theta) + (1-\alpha_1) \times \left[(1-\theta)(\alpha_2^{nc}-2) \times \frac{1}{4} + \frac{2-\alpha_2^{nc}}{2}(2-\theta) \times \frac{1-\theta}{(2-\theta)^2} \right] \\
&= \frac{1}{4}\alpha_1(1-\theta) + \frac{1}{4}\alpha_1 \pi_2^c(\alpha_2^c) \\
&+ \frac{1}{4}(1-\alpha_1)(1-\theta) + \frac{1}{2}(1-\alpha_1) \times \left[\frac{1-\theta}{2-\theta} - \frac{1-\theta}{2} \right](2-\alpha_2^{nc}) \\
&= \frac{1}{4}\alpha_1(1-\theta) + \frac{1}{4}\alpha_1 \pi_2^c(\alpha_2^c) \\
&+ \frac{1}{4}(1-\alpha_1)(1-\theta) + \frac{1}{2}(1-\alpha_1)\frac{\theta(1-\theta)}{2(2-\theta)}(2-\alpha_2^{nc}) \\
&= \frac{1}{4}(1-\theta) + \frac{1}{4}\alpha_1 \pi_2^c(\alpha_2^c) + \frac{1}{2}(1-\alpha_1)\pi_2^{nc}(\alpha_2^{nc})
\end{aligned}
$$

Thus, in all cases:

$$
\frac{\partial E[p_2^*(1-p_2^*)]}{\partial \alpha_1} = \frac{\pi_2^c(\alpha_2^c) - 2\pi_2^{nc}(\alpha_2^{nc})}{4}. \qquad (3.11)
$$

Thus, a reduction in C_r (i.e., an increase in α_1) increases the informativeness of the price at date 2 (i.e., reduces $E[p_2^*(1-p_2^*)]$) if and only if $\pi_2^c(\alpha_2^c) < 2\pi_2^{nc}(\alpha_2^{nc})$. Now we analyze when this is the case and when it is not. For this, note that: $\pi_2^c(\alpha_2^c) = C_p/\alpha_2^c$ if $\alpha_2^c > 0$ and $\pi_2^c(0) = \theta(1-\theta)$ if $\alpha_2^c = 0$. Similarly, $\pi_2^{nc}(\alpha_2^{nc}) = C_p/\alpha_2^{nc}$ if $\alpha_2^{nc} > 0$ and $\pi_2^c(0) = \theta/2$ if $\alpha_2^{nc} = 0$ (see eq.(3.5)).

- If $C_p > \frac{\theta(2-\theta)}{8}$, $\alpha_2^{nc} = \alpha_2^c = 0$. In this case, using eq.(3.11), we deduce from the previous remarks that:

  $$
  \frac{\partial E[p_2^*(1-p_2^*)]}{\partial \alpha_1} = -\frac{\theta^2}{4} < 0,
  $$

  which is decreasing with α_1.

- If $\frac{\theta(1-\theta)}{4} < C_p \leq \frac{\theta(2-\theta)}{8}$ we have $\alpha_2^c = 0$ and $\alpha_2^{nc} > 0$. In this case, In this case, using eq.(3.11), we obtain that:

$$\frac{\partial E[p_2^*(1-p_2^*)]}{\partial \alpha_1} = \theta(1-\theta) - C_p/\alpha_2^{nc}. \tag{3.12}$$

Thus, the pricing error increases with α_1 if and only if

$$\alpha_2^{nc}(\theta, C_p) \geq \frac{2C_p}{\theta(1-\theta)}. \tag{3.13}$$

Using eq.(3.6), we deduce that $\alpha_2^{nc}(\theta, C_p)$ increases when C_p decreases. And using eq.(3.11), we deduce that the derivative of the pricing error with respect to α_1 increases when C_p decreases. We also observe that when $C_p \leq \frac{\theta(1-\theta)}{2(2-\theta)}$, $\alpha_2^{nc} \geq 1$ and therefore eq.(3.13) is satisfied. Moreover, when $C_p = \frac{\theta(2-\theta)}{8}$, $\alpha_2^{nc} = \frac{2-\theta}{2}$. Thus, in this case, eq.(3.13) is satisfied if and only if $\theta < 1/2$. Combining these two observations, we deduce that when $\frac{\theta(1-\theta)}{4} < C_p \leq \frac{\theta(2-\theta)}{8}$, there is a $\bar{C}_p \in \left[\frac{\theta(1-\theta)}{2(2-\theta)}, \frac{\theta(2-\theta)}{8}\right]$ such that asset price informativeness decreases when C_r decreases if (i) $\theta < 1/2$ or if (ii) $\theta > 1/2$ and $C_p < \bar{C}_p$.

- If $C_p \leq \frac{\theta(1-\theta)}{4}$, using eq.(3.11), we obtain that:

$$\frac{\partial E[p_2^*(1-p_2^*)]}{\partial \alpha_1} = \frac{C_p}{2}\left(\frac{1}{2\alpha_2^c} - \frac{1}{\alpha_2^{nc}}\right) > 0, \tag{3.14}$$

where the inequality follows from the fact that $2\alpha_2^c \leq \alpha_2^{nc}$ when $C_p \leq \frac{\theta(1-\theta)}{4}$.

3.2 Investment in information processing takes place before the realization of the price at date 1

Now we consider the case in which the seller and the buyers of the processed signal contract *before* observing the realization of the price at date 1 but the buyers can make their decision to eventually buy the signal contingent on the realization of the price at date 1. This corresponds to the case in which (i) information processing must start early at date 1 for the processed signal to be delivered in time at date 2 and (ii) the seller offers a schedule (F_2^c,F_2^{nc}), i.e., a fee for the processed signal that is contingent on the realization of the price. A speculator can then decide to buy the processed signal only when (i) the price has changed at date 1, (ii) only when the price has not changed or (iii) in both cases.

Denote $R_2^{nc} = \alpha_2^{nc} F_2^{nc}$ and $R_2^c = \alpha_2^c F_2^c$. An equilibrium of the market for the processed signal is a vector $(\alpha_2^{nc}, \alpha_2^c, F_2^{nc}, F_2^{nc})$ such that:

$$\text{Zero profit for speculators if } p_1 \neq p_0: \quad \bar{\pi}_2^{net,c}(\alpha_2^c, F_2^c, \theta) = \pi_2^c(\alpha_2^c, \theta) - F_2^c = 0. \tag{3.15}$$

$$\text{Zero profit for speculators if } p_1 = p_0: \quad \bar{\pi}_2^{net,c}(\alpha_2^c, F_2^c, \theta) = \pi_2^c(\alpha_2^c, \theta) - F_2^c = 0. \tag{3.16}$$

Zero profit for the information seller:

$$\bar{\Pi}_2^{seller}(\alpha_2^{nc}, \alpha_2^c, F_2^c, F_2^{nc}) = Pr(p_1 \neq p_0) R_2^c + (1 - Pr(p_1 \neq p_0)) R_2^{nc} - C_p = \alpha_1 R_2^c + (1 - \alpha_1) R_2^{nc} - C_p = 0. \tag{3.17}$$

The only difference with the case analyzed in Section 3.1 is the zero profit condition for the seller of the processed signal. This condition accounts for the fact that the information seller must pay the fixed cost of information processing before knowing whether $p_1 \neq p_0$ (which happens with probability $Pr(p_1 \neq p_0) = \alpha_1$) or $p_1 = p_0$ (which happens with probability $(1 - \alpha_1)$).

Note that $R_2^{nc} = \alpha_2^{nc} F_2^{nc}$ and $R_2^c = \alpha_2^c F_2^c$ are the revenues of the seller of the processed signal conditional on the two possible outcomes at the end of period 1: (i) no change (nc) in the price of the asset or (i) a change (c) in the price of the asset. These revenues are bounded by the largest possible value for the aggregate expected profits from trading on the processed signal conditional on each possible outcome, that is:

$$R_2^{nc} \leq \frac{\theta(2-\theta)}{8}, \text{ and } R_2^c \leq \frac{\theta(1-\theta)}{4}$$

Thus, the largest possible expected profit for the seller of the processed signal is $\bar{\Pi}(\alpha_1) \equiv (1-\alpha_1)\frac{\theta(2-\theta)}{8} + \alpha_1 \frac{\theta(1-\theta)}{4}$ We deduce that a necessary condition for the seller of the processed signal to pay the fixed cost of producing this signal is that:

$$C_p \leq \bar{\Pi}(\alpha_1). \tag{3.18}$$

Henceforth we assume that this condition is satisfied. Otherwise the processed signal is not produced and $\alpha_2^c = \alpha_2^{nc} = 0$.

Observe that there might be multiple pairs $(R_2^c,\ R_2^{nc})$ for the zero expected profit condition for the seller of the processed signal (in eq.(3.2)), which leads to the possibility of multiple equilibria in the market for the processed signal. When this happens, we we select the equilibrium that minimize the variance of revenues for the information seller, i.e., $(R_2^{nc} - R_2^c)^2$ since the seller's revenue is binomial random variable. Indeed, this would be the strictly preferred outcome for the information seller if it is risk averse.

We can rewrite the zero profit condition (3.17):

$$R_2^{nc} - R_2^c = \frac{C_p - R_2^c}{1 - \alpha_1} \tag{3.19}$$

We then consider two cases. The first case arises when $C_p \leq \frac{\theta(1-\theta)}{4}$. In this case, the variance of revenues is minimized for $R_2^c = C_p$ and is equal to zero, so that $R_2^{nc} = C_p$ as well. This outcome is possible because $C_p < \frac{\theta(1-\theta)}{4}$, the largest possible value for R_2^c. The zero profit condition (3.2) can thus be replaced by two conditions:

$$\alpha_2^c F_2^c = C_p,$$

and

$$\alpha_2^{nc} F_2^{nc} = C_p.$$

The analysis is then identical to the case analyzed in Section 3.1. In particular, as in Section 3.1, we obtain that the equilibrium demand for the processed signal when $p_1 \neq p_0$ and when $p_1 = p_0$ are respectively:

$$\alpha_2^c(\theta, C_p) = \frac{1}{2} + \left(\frac{1}{4} - \frac{C_p}{\theta(1-\theta)}\right)^{\frac{1}{2}} \tag{3.20}$$

and

$$\alpha_2^{nc}(\theta, C_p) = 1 + \left(1 - \frac{2(2-\theta)}{\theta(1-\theta)}C_p\right)^{\frac{1}{2}} \tag{3.21}$$

The second case arises when $\frac{\theta(2-\theta)}{8} \geq C_p > \frac{\theta(1-\theta)}{4}$. In this case, the constraint that $R_2^c \leq \theta(1-\theta)/4$ is binding. Inspection of eq.(3.19) shows that the equilibrium such that the variance of revenues for the information seller is minimal is such that R_2^c is set at its maximal possible value, i.e., $\frac{\theta(1-\theta)}{4}$, which is therefore the aggregate expected trading profits obtained by speculators buying the processed signal when the price changes at date 1. This implies that in this case the demand for the processed signal must be:

$$\alpha_2^c = \frac{1}{2}, \tag{3.22}$$

and therefore $F_2^c = \theta(1-\theta)/2$. Moreover, using eq.(3.19), we deduce that R_2^{nc} must be such that

$$R_2^{nc} = \alpha_2^{nc} F_2^{nc} = \frac{\theta(1-\theta)}{4} + \frac{C_p - \frac{\theta(1-\theta)}{4}}{1-\alpha_1}.$$

Let denote the R.H.S of this equation by $R(\alpha_1, C_p)$ (i.e., $R(\alpha_1, C_p) = \frac{\theta(1-\theta)}{4} + \frac{C_p - \frac{\theta(1-\theta)}{4}}{1-\alpha_1}$). We can then derive the equilibrium demand for the processed signal when $p_1 = p_0$ as we did in Section 3.1, except that the role of C_p is now played by $R(\alpha_1, C_p)$. We deduce that the demand the processed signal conditional on no change in the price of the asset at date 1 is:

$$\alpha_2^{nc}(\alpha_1, \theta, C_p) = \begin{cases} 1 + \left(1 - \frac{2(2-\theta)}{\theta(1-\theta)} R(\alpha_1, C_p)\right)^{\frac{1}{2}} & \text{if } \frac{\theta(1-\theta)}{4} < R(\alpha_1, C_p) \leq \frac{\theta}{2}\frac{1-\theta}{2-\theta}, \\ \frac{2-\theta}{2} + \left(\frac{(2-\theta)^2}{4} - \frac{2(2-\theta)}{\theta} R(\alpha_1, C_p)\right)^{\frac{1}{2}} & \text{if } \frac{\theta}{2}\frac{1-\theta}{2-\theta} < R(\alpha_1, C_p) \leq \frac{\theta(2-\theta)}{8} \end{cases} \tag{3.23}$$

In sum, the only difference with the case studied in Section 3, is the case where $\frac{\theta(2-\theta)}{8} \geq C_p > \frac{\theta(1-\theta)}{4}$. There are two differences relative to what we obtain in 3.

First, even if the price changes at date 1, there is a demand for the processed signal when $C_p > \frac{\theta(1-\theta)}{4}$ while this demand was nil in the case considered in Section 3. The reason is that conditional on the realization of this state (a change in price), the information seller sells the processed signal at a price that is too low to recover its fixed cost of producing the signal. As this cost has been sunk, selling the signal is optimal anyway but, in this state, the fee charged for the signal is less than the cost of producing it. This implies that when $p_1 = p_0$, the seller of the processed signal sells it at a price that exceeds his fixed cost of producing the processed signal. This price cannot be profitably undercut because the decision to produce the processed signal (and pay the associated fixed cost) must be made before observing the realization of the price of the asset at date 1.

Second, α_2^{nc} is negatively related to α_1. Yet, as in Section 3.1, we still have $\alpha_2^{nc} > \alpha_2^{c}$. Moreover, the expected demand for information is equal to $E(\alpha_2^e) = \alpha_1 \frac{1}{2} + (1-\alpha_1)\alpha_2^{nc}(\alpha_1)$. As in Section 3.1, the expected demand for the processed signal decreases with the demand for the raw signal since $E(\alpha_2^e) = \alpha_1 \frac{1}{2} + (1-\alpha_1)\alpha_2^{nc}(\alpha_1)$ and $\alpha_2^{nc}(\alpha_1)$ decreases with α_1.

Asset price informativeness at date t=2. As in Section 3, we analyze the effect of a reduction of C_r (i.e., an increase in α_1) on the informativeness of prices at $t = 2$.

Proposition 3.3. *If $C_p > \bar{\Pi}(\alpha_1)$, a decrease in the cost of producing the raw signal improves asset price informativeness at date 2. If $C_p \leq \bar{\Pi}(\alpha_1)$ then there is a $\bar{C}_p(\alpha_1) \in \left[\frac{\theta(1-\theta)}{4}, \bar{\Pi}(\alpha_1)\right]$ such that a decrease in the cost of producing the raw signal reduces asset price informativeness at date 2 if and only if $C_p < \bar{C}_p(\alpha_1)$.*

Qualitatively, we find the same result as in Section 3.1: when the cost of producing processed information is "high", a decrease in the cost of producing the raw signal improves asset price informativeness at date 2, and conversely when the cost of producing processed information is "low".

Price and trade patterns. As speculators are either active in both cases ("c" or "nc") or not at all, a speculator's trade at $t = 2$ is as follows,

$$x_2^*(p_1, u, C_p) = \begin{cases} 0 & \text{if } C_p > \bar{\Pi}(\alpha_1), \\ u \times [\mathbb{I}_{V=1} - \mathbb{I}_{V=0}] + (1-u) \times [\mathbb{I}_{p_1=(1-\theta)/2} - \mathbb{I}_{p_1=(1+\theta)/2}] & \text{if } 0 \leq C_p \leq \bar{\Pi}(\alpha_1). \end{cases} \tag{3.24}$$

Therefore, proceeding as in Section 3.1, we obtain:

$$Cov(x_1^*, x_2^*) = \begin{cases} 0, \text{ if } C_p > \bar{\Pi}(\alpha_1), \\ \theta - (1-\theta)\alpha_1^e, \text{ if } C_p \leq \bar{\Pi}(\alpha_1). \end{cases}$$

$$Cov(r_1, x_2^*) = E\left[\left(p_1 - \frac{1}{2}\right) x_2\right] = \begin{cases} 0, \text{ if } C_p > \bar{\Pi}(\alpha_1), \\ \theta(2\theta - 1)\alpha_1^e, \text{ if } C_p \leq \bar{\Pi}(\alpha_1). \end{cases}$$

$$Cov(x_1^*, r_2) = \begin{cases} 0, \text{ if } C_p > \bar{\Pi}(\alpha_1), \\ \frac{\theta(1-\alpha_1^e)\alpha_2^{nc}(\alpha_1)}{2(2-\theta)}, \text{ if } \frac{\theta(1-\theta)}{4} < C_p \leq \bar{\Pi}(\alpha_1) \text{ and } R(\alpha_1, C_p) > \frac{\theta(1-\theta)}{2(2-\theta)}, \\ \frac{\theta(1-\alpha_1^e)(1+(1-\theta)(\alpha_2^{nc}(\alpha_1)-1))}{2(2-\theta)}, \text{ if } \frac{\theta(1-\theta)}{4} < C_p \leq \bar{\Pi}(\alpha_1) \text{ and } R(\alpha_1, C_p) \leq \frac{\theta(1-\theta)}{2(2-\theta)}, \\ \frac{\theta(1-\alpha_1^e)(1+(1-\theta)(\alpha_2^{nc}-1))}{2(2-\theta)}, \text{ if } C_p \leq \frac{\theta(1-\theta)}{4}. \end{cases}$$

The expression of covariances $Cov(x_1^*, x_2^*)$ and $Cov(r_1, x_2^*)$ are as in the paper. The only difference is the threshold $\bar{\Pi}(\alpha_1)$ that replaces C_{max}. Compared to Section 3.1, we lose the "intermediate" cases which correspond to the situation where $\alpha_2^c = 0$ and $\alpha_2^{nc} > 0$, when $\frac{\theta(2-\theta)}{8} \geq C_p > \frac{\theta(1-\theta)}{4}$.

$Cov(x_1^*, r_2)$ admits more cases than in the paper or in Section 3.1. Notice that in the second and third cases, α_1 also enters the expression via $\alpha_2^{nc}(\alpha_1)$, which is not the case in the fourth case, and in Section 3. Yet, in all cases, $Cov(x_1^*, r_2)$ decreases with α_1, as in Section 3.1.

Proofs for Section 3.2

Proof or Proposition 3.3. When $C_p > \bar{\Pi}(\alpha_1)$, we have $\alpha_2^{nc} = \alpha_2^c = 0$ then informativeness increases with α_1 as in Proposition 3.2, when $C_p > \frac{\theta(2-\theta)}{8}$.

When $C_p > \frac{\theta(1-\theta)}{4}$ Using equation (3.11), we can calculate the derivative of the pricing error with respect to α_1 as follows

$$\frac{\partial E[p_2^*(1-p_2^*)]}{\partial \alpha_1} = \frac{\pi_2^c(1/2) - 2\pi_2^{nc}(\alpha_2^{nc})}{4} = \frac{1}{4}\left[\frac{\theta(1-\theta)}{2} - \frac{2}{\alpha_2^{nc}}\left(\frac{\theta(1-\theta)}{4} + \frac{C_p - \frac{\theta(1-\theta)}{4}}{1-\alpha_1}\right)\right],$$

We can first notice that this derivative is decreasing function of C_p since α_2^{nc} declines with C_p and $\pi_2^{nc}(\alpha_2^{nc})$ declines with α_2^{nc}.

Second, the derivative can be rewritten as

$$\frac{1}{4}\left[\frac{\theta(1-\theta)}{2}\left(1 - \frac{1}{\alpha_2^{nc}}\right) - \frac{2}{\alpha_2^{nc}}\frac{C_p - \frac{\theta(1-\theta)}{4}}{1-\alpha_1}\right].$$

If $\frac{\theta}{2}\frac{1-\theta}{2-\theta} \leq R(\alpha_1, C_p) \leq \frac{\theta(2-\theta)}{8}$, we have $\alpha_2^{nc} \leq 1$, then the derivative of the pricing error is negative. It follows that price informativeness increases with α_1.

When $R(\alpha_1, C_p) < \frac{\theta}{2}\frac{1-\theta}{2-\theta}$, $\alpha_2^{nc} > 1$. Consider the case where $R(\alpha_1, C_p)$ converges to $\frac{\theta(1-\theta)}{4}$, or equivalently C_p converges to $\frac{\theta(1-\theta)}{4}$ from above. The derivative converges towards

$$\frac{\theta(1-\theta)}{8}\left(1 - \frac{1}{\alpha_2^{nc}}\right) > 0.$$

Since $\partial E[p_2^*(1-p_2^*)]/\partial\alpha_1$ is a decreasing function of C_p, it follows that there is a threshold $\bar{C}_p(\alpha_1)$ such that price informativeness decreases with α_1, i.e., $\partial E[p_2^*(1-p_2^*)]/\partial\alpha_1 > 0$, if and only if $C_p < \bar{C}_p(\alpha_1)$.

When $C_p \leq \frac{\theta(1-\theta)}{4}$, we are in the same case as in Proposition 3.2. Asset price informativeness decreases with α_1.

4 Complement to the proof of Proposition 3.

Let $\bar{\alpha}_1(\theta) = \frac{(1-\theta)^2+\theta^2}{(1-2\theta)[2(1-\theta)(2-\theta)-1]}$ and $\bar{C}_p(\theta) = \frac{(\theta(1-\theta)^2)(1-2\theta)}{(2-\theta)\left(2(1-\theta)-\frac{1}{2-\theta}\right)^2}$. We first prove the following result.

Lemma 4.1. *For $\theta < 1/2$ and $C_{min}(\theta, \alpha_1) \leq C_p < C_{max}(\theta, \alpha_1)$, $\frac{\partial \alpha_2^e}{\partial \alpha_1} > 0$ if and only if (i)*

$\theta < \frac{\sqrt{2}-1}{\sqrt{2}}$, *(ii)* $\alpha_1 > \bar{\alpha}_1(\theta)$, *and (iii)* $C_p > \bar{C}_p(\theta)$.

Proof. When $\theta < 1/2$, we know from Corollary 2 that $\frac{\partial \alpha_2^e}{\partial \alpha_1} > 0$ if and only if

$$\alpha_2^e(\alpha_1) < \hat{\alpha}_2(\theta) = \frac{(2-\theta)(1-2\theta)}{2(2-\theta)(1-\theta)-1}$$

Using the expression of α_2^e in Lemma 1 in the text, we obtain that $\alpha_2^e(\alpha_1) < \hat{\alpha}_2(\theta)$ if and only if

$$\alpha_2^{max}(\theta, \alpha_1)\left(1+\sqrt{1-\frac{C_p}{C_{max}(\theta, \alpha_1)}}\right) < \hat{\alpha}_2(\theta). \tag{4.1}$$

That is, if and only if

$$\sqrt{1-\frac{C_p}{C_{max}(\theta, \alpha_1)}} < \frac{\hat{\alpha}_2(\theta)}{\alpha_2^{max}(\theta, \alpha_1)} - 1. \tag{4.2}$$

For this inequality to be verified, a necessary condition is that the right hand side is positive. That is, we must have:

$$\hat{\alpha}_2(\theta) > \alpha_2^{max}(\theta, \alpha_1) = \frac{1}{2}\frac{1-(2\theta-1)\alpha_1}{\frac{1}{2-\theta}+\left(2(1-\theta)-\frac{1}{2-\theta}\right)\alpha_1}\left(= \frac{(2-\theta)(1-(2\theta-1)\alpha_1)}{2(1+(2(2-\theta)(1-\theta)-1)\,\alpha_1)}\right). \tag{4.3}$$

Observe that $\alpha_2^{max}(\theta, \alpha_1)$ decreases with α_1. Moreover, $\alpha_2^{max}(\theta, 0) = (2-\theta)/2$ and $\alpha_2^{max}(\theta, 1) = 1/2$. We have:

$$\begin{aligned}\alpha_2^{max}(\theta, 0) - \hat{\alpha}_2(\theta) &= \frac{2-\theta}{2} - \frac{1-2\theta}{2(1-\theta)-\frac{1}{2-\theta}} = \frac{2(1-\theta)(2-\theta)-1-2(1-2\theta)}{2\left(2(1-\theta)-\frac{1}{2-\theta}\right)} \\ &= \frac{2(1-\theta)(2-\theta)-4(1-\theta)+1}{2\left(2(1-\theta)-\frac{1}{2-\theta}\right)} = \frac{-2\theta(1-\theta)+1}{2\left(2(1-\theta)-\frac{1}{2-\theta}\right)} > 0,\end{aligned}$$

for $\theta \leq 1/2$. Moreover,

$$\begin{aligned}\alpha_2^{max}(\theta, 1) - \hat{\alpha}_2(\theta) &= \frac{1}{2} - \frac{1-2\theta}{2(1-\theta)-\frac{1}{2-\theta}} = \frac{1-\theta-\frac{1}{2(2-\theta)}-1+2\theta}{2(1-\theta)-\frac{1}{2-\theta}} \\ &= \frac{\theta-\frac{1}{2(2-\theta)}}{2(1-\theta)-\frac{1}{2-\theta}} = \frac{\theta(2-\theta)-\frac{1}{2}}{2(1-\theta)(2-\theta)-1} \\ &= \frac{-(1-\theta)^2+\frac{1}{2}}{2(1-\theta)(2-\theta)-1}.\end{aligned}$$

Therefore $\alpha_2^{max}(\theta,1) - \hat{\alpha}_2(\theta) > 0$ iff $\theta > \frac{\sqrt{2}-1}{\sqrt{2}}$. Consequently if $\theta > \frac{\sqrt{2}-1}{\sqrt{2}}$ then $\alpha_2^{max}(\theta,\alpha_1) - \hat{\alpha}_2(\theta) > 0$ for all α_1 and Condition (4.3) cannot hold true. Hence, $\frac{\partial \alpha_2^e}{\partial \alpha_1} > 0$ cannot hold true if $\theta > \frac{\sqrt{2}-1}{\sqrt{2}}$.

If $\theta < \frac{\sqrt{2}-1}{\sqrt{2}}$, there exists $\bar{\alpha}_1(\theta) \in [0,1]$ such that $\alpha_2^{max}(\theta,\bar{\alpha}_1(\theta)) = \hat{\alpha}_2(\theta)$ and for all $\alpha_1 > \bar{\alpha}_1(\theta)$, $\alpha_2^{max}(\theta,\alpha_1) < \hat{\alpha}_2(\theta)$, since $\alpha_2^{max}(\theta,\alpha_1)$ is decreasing with α_1. Solving the equation $\alpha_2^{max}(\theta,\bar{\alpha}_1(\theta)) = \hat{\alpha}_2(\theta)$ for $\bar{\alpha}_1(\theta)$, we obtain:

$$\bar{\alpha}_1(\theta) = \frac{1}{\left(1 - \frac{\theta}{1-\theta}\right)\left[2(2-\theta) - \frac{1}{1-\theta}\right]} + \frac{\theta^2}{(1-2\theta)[2(1-\theta)(2-\theta)-1]}. \tag{4.4}$$

We deduce that for $\theta < \frac{\sqrt{2}-1}{\sqrt{2}}$, $\bar{\alpha}_1(\theta)$ increases with θ. Moreover,

$$\bar{\alpha}_1(0) = \frac{1}{3}$$

and

$$\bar{\alpha}_1(\frac{\sqrt{2}-1}{\sqrt{2}}) = \frac{\frac{1}{2} + \left(1 - \frac{1}{\sqrt{2}}\right)^2}{(\sqrt{2}-1)\left[2\frac{1}{\sqrt{2}}\left(1+\frac{1}{\sqrt{2}}\right) - 1\right]} = 1.$$

Thus, if $\theta < \frac{\sqrt{2}-1}{\sqrt{2}}$ and $\alpha_1 > \bar{\alpha}_1(\theta)$, Condition (4.2) can be satisfied. This condition is equivalent to:

$$C_p > \bar{C}_p(\theta), \tag{4.5}$$

where

$$\bar{C}_p(\theta) \equiv \frac{C_{max}(\theta,\alpha_1)}{\alpha_2^{max}(\theta,\alpha_1)^2}(2\alpha_2^{max}(\theta,\alpha_1) - \hat{\alpha}_2(\theta))\hat{\alpha}_2(\theta)$$

We have

$$\frac{C_{max}(\theta,\alpha_1)}{\alpha_2^{max}(\theta,\alpha_1)^2} = \frac{\theta}{2}\left[\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1\right].$$

Moreover:

$$\begin{aligned} 2\alpha_2^{max}(\theta,\alpha_1) - \hat{\alpha}_2(\theta) &= \frac{1-(2\theta-1)\alpha_1}{\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1} - \frac{1-2\theta}{2(1-\theta) - \frac{1}{2-\theta}} \\ &= \frac{\frac{2(1-\theta)^2}{2-\theta}}{\left(2(1-\theta) - \frac{1}{2-\theta}\right)\left[\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1\right]} \end{aligned}$$

Thus:

$$\bar{C}_p(\theta) = \frac{\theta(1-\theta)^2}{2-\theta}\frac{\hat{\alpha}_2(\theta)}{2(1-\theta)-\frac{1}{2-\theta}} = \frac{\theta(1-\theta)^2}{2-\theta}\frac{1-2\theta}{\left(2(1-\theta)-\frac{1}{2-\theta}\right)^2} = \frac{\theta(1-\theta)^2(2-\theta)(1-2\theta)}{(2(1-\theta)(2-\theta)-1)^2}.$$

This achieves the proof of Lemma 4.1.**Q.E.D**

Finally, we observe in equilibrium, the condition $\alpha_1^e(C_r) > \bar{\alpha}_1(\theta)$ is equivalent to:

$$C_r < \frac{\theta}{2}\left(\frac{1}{4} - \max\left(\bar{\alpha}_1(\theta) - \frac{1}{2}, 0\right)^2\right) = \bar{C}_r(\theta).$$

5 Complement to Proposition 5.

In Proposition 5, we have shown that if $C_p < C_{min}(\theta, \alpha_1^e)$ then a decrease in the cost of producing the raw signal, C_r, reduce the informativeness of the price at date 2. We now show that this result can hold even when $C_{min}(\theta, \alpha_1^e) < C_p < C_{max}(\theta, \alpha_1^e)$. We also analyze the case $C_p > C_{max}(\theta, \alpha_1^e)$ at the end of this section (Case B).

Case A. Consider the case in which $C_{min}(\theta, \alpha_1^e) < C_p < C_{max}(\theta, \alpha_1^e)$. In this case, $0 < \alpha_2^e \leq 1$ (Proposition 1). Using eq.(6) and eq.(31), we have:

$$\mathcal{E}_2(C_r, C_p) = \frac{\theta}{4} - \frac{1}{2}\left(\bar{\pi}_2(\alpha_1^e, \alpha_2^e) - \frac{\theta(1-\theta)}{2}\alpha_1^e(1-\alpha_2^e)\right), \quad (5.1)$$

where we omit the arguments of functions α_1^e and α_2^e to simplify notations. As $\alpha_2^e > 0$, in equilibrium, $\alpha_2^e \bar{\pi}_2^e = C_p$ (see eq.(11)). Thus, we can rewrite eq.(5.1) as:

$$\mathcal{E}_2(C_r, C_p) = \frac{\theta}{4} - \frac{1}{2}\left(\frac{C_p}{\alpha_2^e} - \frac{\theta(1-\theta)}{2}\alpha_1^e(1-\alpha_2^e)\right), \quad (5.2)$$

Using the fact that C_r affects α_2^e only through its effect on α_1^e, we deduce from eq.(5.1) that:

$$\frac{\partial \mathcal{E}_2(C_r, C_p)}{\partial C_r} = \left(\frac{1}{2}\frac{\partial \alpha_1^e}{\partial C_r}\right)\left(\frac{\partial \alpha_2^e}{\partial \alpha_1^e}\left(\frac{C_p}{(\alpha_2^e)^2} - \frac{\theta(1-\theta)}{2}\alpha_1^e\right) + \frac{\theta(1-\theta)}{2}(1-\alpha_2^e)\right), \quad (5.3)$$

As $\frac{\partial \alpha_1^e}{\partial C_r} \leq 0$, we deduce that the sign of $\frac{\partial \mathcal{E}_2(C_r, C_p)}{\partial C_r}$ is opposite to the sign of the following function:

$$G(\alpha_1^e, \alpha_2^e) \equiv \frac{\partial \alpha_2^e}{\partial \alpha_1^e}\left(\frac{C_p}{(\alpha_2^e)^2} - \frac{\theta(1-\theta)}{2}\alpha_1^e\right) + \frac{\theta(1-\theta)}{2}(1-\alpha_2^e). \tag{5.4}$$

To determine the sign of $G(\alpha_1^e, \alpha_2^e)$, we first compute $\frac{\partial \alpha_2^e}{\partial \alpha_1^e}$. Using eq.(28), we obtain:

$$-\frac{\partial \alpha_2^e}{\partial \alpha_1^e} = \frac{\frac{\partial[\alpha_2^e \bar{\pi}_2(\alpha_1^e,\alpha_2^e)]}{\partial \alpha_1^e}}{\frac{\partial[\alpha_2^e \bar{\pi}_2(\alpha_1^e,\alpha_2^e)]}{\partial \alpha_2^e}} = \frac{\alpha_2^e \frac{\partial \bar{\pi}_2(\alpha_1^e,\alpha_2^e)}{\partial \alpha_1^e}}{\alpha_2^e \frac{\partial \bar{\pi}_2(\alpha_1^e,\alpha_2^e)}{\partial \alpha_2^e} + \bar{\pi}_2(\alpha_1^e,\alpha_2^e)}.$$

Moreover, as $0 < \alpha_2^e \leq 1$, we deduce from Proposition 2 that:

$$\bar{\pi}_2(\alpha_1^e, \alpha_2^e) = \frac{\theta}{2}\left\{1 - (2\theta-1)\alpha_1^e - \left[\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1^e\right]\alpha_2^e\right\}.$$

This implies that

$$\frac{\partial \bar{\pi}_2(\alpha_1^e, \alpha_2^e)}{\partial \alpha_1^e} = -\frac{\theta}{2}\left[2\theta - 1 + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_2^e\right],$$

$$\frac{\partial \bar{\pi}_2(\alpha_1^e, \alpha_2^e)}{\partial \alpha_1^e} = -\frac{\theta}{2}\left[\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1^e\right].$$

Therefore,

$$-\frac{\partial \alpha_2^e}{\partial \alpha_1^e} = \frac{\alpha_2^e\left[2\theta - 1 + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_2^e\right]}{\alpha_2^e\left[\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1^e\right] - \frac{2C_p}{\theta}\frac{1}{\alpha_2^e}}.$$

The denominator of this expression is equal to $-\frac{\partial[\alpha_2^e \bar{\pi}_2(\alpha_1^e,\alpha_2^e)]}{\partial \alpha_2^e}$, which is strictly positive (see the discussion that precedes Lemma 1 in the paper or Figure 5). Hence, we deduce that $G(\alpha_1^e, \alpha_2^e) < 0$ iff:

$$\begin{aligned}
&\alpha_2^e\left[2\theta - 1 + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_2^e\right]\left(\frac{2C_p}{\theta}\frac{1}{(\alpha_2^e)^2} - (1-\theta)\alpha_1^e\right), \\
&- (1-\theta)(1-\alpha_2^e)\left[\alpha_2^e\left(\frac{1}{2-\theta} + \left(2(1-\theta) - \frac{1}{2-\theta}\right)\alpha_1^e\right) - \frac{2C_p}{\theta}\frac{1}{\alpha_2^e}\right] > 0
\end{aligned}$$

After some algebra, one can show that this condition is equivalent to:

$$\Upsilon(\alpha_1^e, \alpha_2^e(\alpha_1^e), \theta, C_p) > 0,$$

where

$$\Upsilon(\alpha_1^e, \alpha_2^e(\alpha_1^e), \theta, C_p) \equiv \frac{1-\theta}{2-\theta}\left(\frac{2C_p}{\theta} - \alpha_2^e(1-\alpha_2^e)\right) + 2C_p\left(\frac{1}{\alpha_2^e} - 1\right) - \alpha_1^e\alpha_2^e\frac{(1-\theta)^2}{2-\theta}. \quad (5.5)$$

In sum, $G(\alpha_1^e, \alpha_2^e) < 0$ iff $\Upsilon(\alpha_1^e, \alpha_2^e(\alpha_1^e), \theta, C_p) > 0$. Thus, when $C_{min}(\theta, \alpha_1^e) < C_p < C_{max}(\theta, \alpha_1^e)$, $\frac{\partial \mathcal{E}_2(C_r, C_p)}{\partial C_r} > 0$ iff $\Upsilon(\alpha_1^e, \alpha_2^e(\alpha_1^e), \theta, C_p) > 0$. In other words, when $C_{min}(\theta, \alpha_1^e) < C_p < C_{max}(\theta, \alpha_1^e)$, a decrease in the cost of the raw signal lowers the informativeness of the price at date 2 if and only if $\Upsilon(\alpha_1^e, \alpha_2^e(\alpha_1^e), \theta, C_p) > 0$.

Case B. Last consider the case in which $C_{max}(\theta, \alpha_1^e) < C_p$. In this case, no speculator buys the processed signal ($\alpha_2^e = 0$). Thus, $\mathcal{E}_2(C_r, C_p) = \mathcal{E}_1(C_r, C_p)$ (see eq.(31) in the proof of Corollary 3). That is, long run price informativeness is equal to short run price informativeness because there is no information production after date 1. In this case, a decrease in C_r raises price informativeness simply because it raises the demand for the raw signal and thereby price informativeness at date 1.

6 Proof that C_{max} decreases with α_1 when $\theta > \frac{\sqrt{2}-1}{\sqrt{2}}$.

Case 1. For $\theta > 1/2$, we can write C_{max} as

$$C_{max}(\theta, \alpha_1) = \frac{\theta}{4}\bar{\alpha}_2(\theta, \alpha_1)(1 - (2\theta - 1)\alpha_1),$$

which is the product of two decreasing and positive functions of α_1. Thus, in this case, C_{max} decreases with α_1.

Case 2. For $\theta < 1/2$, we use the fact that for a given α_2, the aggregate speculators' profit at $t = 2$, that is $\alpha_2\bar{\pi}_2(\alpha_1, \alpha_2)$ can be written as

$$\begin{aligned}\alpha_2\bar{\pi}_2(\alpha_1, \alpha_2) &= \frac{C_{max}(\theta, \alpha_1)}{\alpha_2^{max}(\theta, \alpha_1)^2}(2\alpha_2^{max}(\theta, \alpha_1) - \alpha_2)\alpha_2 \\ &= C_{max}(\theta, \alpha_1)\left(2 - \frac{\alpha_2}{\alpha_2^{max}(\theta, \alpha_1)}\right)\frac{\alpha_2}{\alpha_2^{max}(\theta, \alpha_1)}\end{aligned}$$

We know that for $\alpha_2 = \hat{\alpha}_2(\theta)$, the expected profit $\bar{\pi}_2(\alpha_1, \hat{\alpha}_2(\theta))$ does not depend on α_1.

Therefore the aggregate profit taken in $\hat{\alpha}_2(\theta)$ is also independent of α_1 and happen to be equal to $\bar{C}_p(\theta)$ calculated previously. Indeed,

$$\hat{\alpha}_2(\theta)\bar{\pi}_2(\alpha_1, \hat{\alpha}_2(\theta)) = C_{max}(\theta, \alpha_1)\left(2 - \frac{\hat{\alpha}_2(\theta)}{\alpha_2^{max}(\theta, \alpha_1)}\right)\frac{\hat{\alpha}_2(\theta)}{\alpha_2^{max}(\theta, \alpha_1)} = \bar{C}_p(\theta) > 0. \quad (6.1)$$

Notice that the the function $h(x) = x(2-x)$ reaches a maximum for $x = 1$, increases for $x < 1$ and decreases for $x > 1$. In previous section 4, we have shown that when $\theta < 1/2$, $\bar{C}_p(\theta) > 0$, $\alpha_2^{max}(\theta, \alpha_1)$ decreases with α_1, and $\alpha_2^{max}(\theta, \alpha_1) > \hat{\alpha}_2(\theta)$ if $\theta > \frac{\sqrt{2}-1}{\sqrt{2}}$. Therefore, the ratio $r(\theta, \alpha_1) = \frac{\hat{\alpha}_2(\theta)}{\alpha_2^{max}(\theta,\alpha_1)}$ increases with α_1, and is always less than 1. Hence $r(\theta, \alpha_1)(2-r(\theta, \alpha_1))$ increases with α_1. We deduce that $C_{max}(\theta, \alpha_1)$ must decrease with α_1 since $C_{max}(\theta, \alpha_1)r(\theta, \alpha_1)(2-r(\theta, \alpha_1))$ is strictly positive and does not depend on α_1 (see eq.(6.1)).