Chapman University

Chapman University Digital Commons

Accounting Faculty Articles and Research | Accounting

3-15-2018

Clarity Trumps Content: An Experiment on Information Acquisition in Beauty Contests

Sanjay Banerjee
University of Alberta

Hong Qu
Penn State University

Ran Zhao
Chapman University, rzhao@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/accounting_articles

Part of the Economic Theory Commons, and the Other Economics Commons

Clarity Trumps Content: An Experiment on Information Acquisition in Beauty Contests

Comments

Working paper.

Clarity Trumps Content: An Experiment on Information Acquisition in Beauty Contests

Sanjay Banerjee

Hong Qu

Ran Zhao

ALBERTA SCHOOL OF BUSINESS ESTABLISHED 1916

UNIVERSITY OF ALBERTA
ALBERTA SCHOOL OF BUSINESS

Clarity trumps content: An experiment on information acquisition in beauty contests[*]

Sanjay Banerjee [†] Hong Qu [‡] Ran Zhao [§]

March 15, 2018

Abstract

We provide experimental evidence that under strong beauty contest incentives, players ignore signals from an information source with high content if the source has low clarity. Instead, they acquire equally costly signals from a source with higher clarity despite its lower content. Content measures how precisely an information source identifies an economic situation, whereas clarity measures how precisely the source content is commonly interpreted. Low clarity impairs players' ability to coordinate. When signals are provided exogenously, our experimental results are less severe than theoretical predictions, but consistent with level-2 reasoning in a cognitive behavioral model. When players acquire signals endogenously, ignoring a high-content source is more severe than theoretical predictions. Our results imply that when beauty contest incentives are strong (e.g., short-horizon trading), investors can completely ignore a firm's disclosure, despite its high content, if the disclosure is not sufficiently clear.

Keywords: investor attention, coordination, disclosure, recognition

1   Introduction

When making decisions, investors often have access to multiple sources of information. They can access detailed financial statements, analyst reports, and brief company news in the media. There is evidence that sometimes investors ignore information sources with high content such as financial reports, and instead pay a lot of attention to sources with low content such as news briefs by the media. For example, Niantic, a U.S. software company, released a mobile game, "Pokemon Go," on July 6, 2016. The share price of Nintendo, a Japanese company that has little relationship with the game, rose more than 120% after the release of Pokemon Go. Investors mistakenly believed that Nintendo owned and thus benefited from the popularity of the game because the cartoon character Pokemon first appeared in a video game developed by Nintendo. Although Nintendo's financial statements indicated that the company did not make or own Pokemon Go, the market reaction suggests that investors ignored that information. On the other hand, *Mad Money*, the popular CNBC show, receives a lot of attention from retail investors and generates significant market reactions despite the fact that stock analyses by Jim Cramer, the show's host, are very brief and offer no additional information beyond that which is publicly available.[1] Why do investors ignore information with high content (financial statements), but pay attention to information with low content (Cramer's show)?

In this paper, we propose that beauty contest incentives can lead agents to ignore an information source with high content and instead pay attention to a source with low content.[2] As John Keyenes said almost a century ago, "stock picking is like a beauty contest." An investor picks a company's stock not only because the fundamentals look promising but also because the investor wants to choose what other investors find attractive. As such, when beauty contest incentives are present, investors want to acquire information not only about the economic fundamentals but

---

[1]Engelberg et al. (2012) report that stocks recommended by Cramer experience very high trading activities, and generate above 2.4% abnormal overnight returns. These recommendations do not appear to be informative in the long run because the overnight returns subsequently reverse after a few months.

[2]Agents' have beauty contest incentives when they want to "do what others do." The (Keynesian) beauty contest is characterized by strategic complementarities in agents' decision rules; i.e., the marginal payoffs for an individual agent's action increases with the average action of all other agents.

also about other investors' likely actions. Apart from the content of an information source, the clarity of that source plays a critical role in investors' information acquisition decisions. Content of a source determines how precisely the source identifies the economic fundamentals, whereas clarity determines how precisely the source content is commonly interpreted. An information source with higher clarity is similarly interpreted, and therefore more useful in predicting others' likely actions. If beauty contest incentives are sufficiently strong, investors' incentive to acquire commonly interpreted information dominates their incentive to learn about the fundamentals— i.e., clarity trumps content. As such, investors pay more attention to a high-clarity information source despite its low content, and ignore a low-clarity source despite its high content. Myatt and Wallace (2012) derive this insight in a beauty contest model. We construct a discrete information choice version of Myatt and Wallace (2012) and test this key insight in laboratory experiments.

In the extant literature, investors' information acquisition (attention) behavior is explained by individual investor's information processing costs. An investor tends to ignore financial statements but pays attention to Cramer's brief stock analyses because the former are harder to process than the latter. However, this intuition is incomplete because an investor has no reason to ignore financial statements with high information content if the benefit from higher content is greater than the processing costs. The strategic perspective raised by Myatt and Wallace (2012) can explain why investors rationally ignore information with high content and instead pay attention to low-content information. This perspective emphasizes that when players have strong incentive to coordinate with others, source clarity plays a more important role than source content in players' information acquisition decisions. Although companies' financial statements have more content because of their details, they often lack clarity because investors may focus on different parts of the statements and draw different interpretations.[3] In contrast, Jim Cramer's stock recommendations may have lower content, but they have greater clarity because they are concise and simple, and thus more likely to be commonly interpreted.

---

[3]Investors, firms, practitioners, and regulators are concerned that corporate disclosures have become longer and increasingly complex to navigate (KPMG, 2011; Monga and Chasan, 2015). To improve the clarity of financial statements, the SEC has taken several initiatives including the "plain English" rule (1998), and the more recent FAST ACT (2015).

We illustrate investors' information acquisition behavior by constructing a beauty contest model (Morris and Shin, 2002) similar to Myatt and Wallace (2012) but with discrete information choice and finite number of agents. Prior literature (Allen et al., 2006; Gao, 2008) has shown that a trading model with short-horizon traders has the features of a beauty contest model in which the incentive to match the action of others (strategic complementarity) emerges endogenously. Allowing players to acquire information from multiple sources before the beauty contest stage, Myatt and Wallace (2012) derive the theoretical insight that information sources which receive attention are the clearest possible even though they have low content. With continuous information acquisition choice in Myatt and Wallace (2012), the overall clarity of a signal is endogenous and the information acquisition equilibrium is unique. With discrete information choice in our model, signal clarity is exogenous and multiple acquisition equilibria emerge under certain parameter sets. We extend Myatt and Wallace by including players' limited order of beliefs in a cognitive behavioral model (e.g., Stahl and Wilson, 1994; Nagel, 1995) in both beauty contest and information acquisition stages.

In our model, two players simultaneously decide whether to acquire signals from two information sources at a cost before the beauty contest game. Each information source generates signals with two noise components: a common noise shared among all players, and an idiosyncratic noise unique to each player. The common noise may arise because a source generates information about the economic fundamentals with error. An information source with higher content has common noise with a smaller variance. The idiosyncratic noise may arise because players have different interpretations of the signal with common noise. An information source with higher clarity results in idiosyncratic noises with a smaller variance.[4] A unique feature of our setup is that we keep the total variance of (a signal from) each source the same, but vary the level of content and clarity. One information source has high content and low clarity, and the other has low content and high clarity. Although both information sources generate equally costly signals

---

[4]Liang and Zhang (2014) show that this information structure endogenously arises in a measurement system with managerial discretion and aggregation. They refer to the precision of common error as "accuracy" and the precision of idiosyncratic error as "objectivity." Similar information (exogenous) structures have been used in Indjejikian (1991), Holthausen and Verrecchia (1990) and Kim and Verrecchia (1994).

and are equally useful in reducing the fundamental uncertainty, the source with higher clarity is more useful in predicting others' actions and as such, reducing the strategic uncertainty.

When information is provided exogenously, our model predicts that in the absence of beauty contest incentives, players assign equal weight to signals from both information sources. However, when there are incentives to coordinate, players assign more weight to a signal from the high-clarity and low-content source. When players acquire information endogenously, our three main predictions are: (i) the likelihood of a player acquiring a signal from a low-clarity and high-content source increases with her opponent acquiring the same signal (complementarity in information acquisition); (ii) when information acquisition costs are high, strong beauty contest incentives lead players to acquire information only from the high-clarity source, and ignore the source with lower clarity despite its higher content; and (iii) when information acquisition costs are low, players acquire information from both information sources, but they assign more weight to information with higher clarity and lower content.

While these predictions are stark, one wonders whether they are supported by empirical evidence, as the predictions require strong assumptions on players' ability to process an infinite level of high-order beliefs to coordinate. It is unclear whether our prediction that information with high content is completely ignored if it has low clarity holds in reality. In archival data, players' private information acquisition decisions are unobservable, which makes it difficult to identify the effect of coordination on information acquisition. Thus, we test our theoretical predictions in controlled laboratory experiments, in which we can observe players' private information acquisition decisions, and manipulate the degree of strategic complementarity and the cost of information.

Our experiments employ a 2X2 between-subjects design by manipulating the strength of complementarity and the cost of information. We find experimental evidence of complementarity in information acquisition. Under endogenous information acquisition, our experimental results confirm that information with high content and low clarity is ignored at high information costs. However, contrary to our theoretical predictions, information with high content is ignored also when information costs are low. Additionally, we find that with exogenous information, the underweighting of information with high content and low clarity is less severe than theoretical

4

predictions based on an infinite level of high-order beliefs but consistent with level-2 reasoning in a cognitive behavioral model. This finding suggests that compared to equilibrium predictions, subjects overuse information with high content and low clarity. The overuse of such information at the beauty contest stage creates an endogenous cost at the information acquisition stage by driving players' actions further away from each other, and thereby reducing their payoffs. This *endogenous* cost of acquiring information with high content makes it rational for players to ignore such information even when the *exogenous* cost of information acquisition is low.

Our paper contributes to the emerging accounting literature that explores how strategic complementarity affects disclosure choices. Assuming exogenous information, prior literature primarily focuses on the insight from Morris and Shin (2002) that beauty contest incentives lead agents to put greater weight on information that is more public in nature. This effect underlies the findings in Gao (2008) and Chen et al. (2014) in the context of short-horizon trading, Arya and Mittendorf (2016) and Chen et al. (2017) in the context of investment complementarities, and Anctil et al. (2004), Banerjee and Maier (2016), Qu (2013), and Liang and Zhang (2014) in the context of bank runs. We are the first to introduce the insight of Myatt and Wallace (2012) with endogenous information acquisition. Most importantly, we are the first to provide experimental evidence on the key implication of the model.

Our study is among the first to experimentally test complementarity in information acquisition. In a two-period model with short-horizon traders, Chamley (2007) establishes the existence of strategic complementarity in the acquisition of traders' private information. In a generic model of beauty contest, Hellwig and Veldkamp (2009) show that agents' information choices exhibit complementarity in the presence of strategic complementarity: if an agent wants to do what others do, she wants to know what others know. Our study contributes to this literature by providing experimental evidence of complementarity in information acquisition.[5]

Our findings offer very different policy implications from prior literature on the role of the

---

[5]Szkup and Trevino (2012) study endogenous information acquisition in a global game of speculative attack. They find that although most subjects play the predicted equilibrium in the information acquisition stage, subjects' action choices in the coordination stage deviate from the theoretical predictions. The authors focus their analysis on the observed deviation behavior in the coordination stage. We find subjects' deviation behavior in both the coordination (beauty contest) and information acquisition stages.

public nature of information in beauty contests. With exogenous information, prior research finds that over-weighting of public information in laboratory experiments is less severe than theoretical predictions (e.g., Cornand and Heinemann, 2014; Baeriswyl and Cornand, 2016). While theory imposes the assumption of infinite level of reasoning, experimental evidence indicates that subjects' behavior is more consistent with level-2 reasoning. Using a theoretical model, Cornand and Heinemann (2015) show that the counterintuitive result that increased transparency reduces social welfare in Morris and Shin (2002) no longer holds if we assume agents' limited levels of reasoning. Thus, prior research implies that the distortion in the use of public information under beauty contest incentives may not be a serious concern. Our paper shows that this policy implication is no longer valid when information is endogenous. Our experimental evidence suggests that the publicity of information matters even more than theoretical predictions. Players completely ignore information with high content and low clarity even when theory predicts that they do not. Such ignoring can result in significant distortions in capital market efficiency when beauty contest incentives are strong (e.g., short-horizon trading). Moreover, the flip side is that investors may systematically gravitate toward high-clarity but low-content information. Rumors or even sunspot events with high clarity may drive prices, whereas the impact of corporate disclosures on market prices diminishes.

The strategic effect of beauty contest incentives on information acquisition offers a new perspective to understand market reactions to accounting disclosure. Prior research suggests that information disclosed in footnotes has less market impact than information recognized in financial statements (e.g., Aboody, 1996; Davis-Friday et al., 1999; Ahmed et al., 2006). Past research emphasizes the individual perspective of information processing suggesting that the difference in impact could be due to investors' cognitive biases or differences in processing costs (Schipper, 2007). Our study offers a strategic perspective to explaining this empirical regularity by emphasizing individual beliefs about others' behavior. Footnote disclosures may contain a lot of detail but often lack clarity.[6] When beauty contest incentives are strong, investors may underuse

---

[6]Morris and Shin (2007) suggest that accounting standards on recognition provide a common rule for investors to interpret recognized numbers, thereby enhancing their clarity. In contrast, there are no common standards for footnote disclosure (e.g., Schipper, 2007), and investors' interpretation of footnote

or even completely ignore footnote disclosures because they believe other investors most likely do the same. Beauty contest incentives are stronger when there are more transient institutional investors who speculate on short-term price changes (e.g., Sapra, 2010). A testable empirical prediction based on our findings is that market under-reaction to footnote disclosure is greater for firms with more transient institutional investors.

2 Theory and Predictions

We consider a quadratic beauty-contest game with two players, i and j. Players' payoffs depend on the closeness of their actions to an unobserved economic fundamental (θ) and to each other's actions (a^i and a^j). Player i's utility function is

$$u^i \left(a^i, a^j, \theta \right) = \bar{u} - (1-r)(a^i - \theta)^2 - r(a^i - a^j)^2, \tag{1}$$

where $r \in (0,1)$ is a complementarity parameter. There are two stages of the game. In the first stage, players simultaneously decide to acquire (i.e., pay attention to) signals from two information sources by paying a cost. In the second stage, having observed their own private signals, as well as other players' information acquisition decisions (but not their signal realizations), players play the beauty contest game.

Players have a common prior about the fundamental, $\theta \sim N(\bar{\theta}, \sigma_\theta^2)$. Each player i has access to two real-valued signals x_A^i and x_B^i. The signals

$$x_A^i = \theta + \eta_A + \varepsilon_A^i \text{ and}$$
$$x_B^i = \theta + \eta_B + \varepsilon_B^i \tag{2}$$

contain noise terms $\eta_A \sim N(0, \alpha_A^2)$, $\eta_B \sim N(0, \alpha_B^2)$ and $\varepsilon_A^i \sim N(0, \kappa_A^2)$, $\varepsilon_B^i \sim N(0, \kappa_B^2)$ that are mutually independently distributed. *Signal* x_A^i (player i's signal A) is generated by information *source* A and signal x_B^i (player i's signal B) is generated by information source B. A player's

disclosure can vary a lot among individuals.

choice to acquire a signal from a source A is represented by $z_A^i = 1$, and the choice to not acquire by $z_A^i = 0$. The cost of information acquisition from sources A and B is $C(z_A^i, z_B^i)$.

The noise in each signal x_A^i has two components: a common noise η_A and an idiosyncratic noise ε_A^i, which is different for each player. The common noise reflects the "sender" noise, or the noise from an information source. The idiosyncratic noise reflects the "receiver" noise, or the noise generated by the receiver's interpretation of the common signal $x_A \equiv \theta + \eta_A$. The *content* of an information source is defined as the precision of the common noise; i.e., $1/\alpha_A^2$. The *clarity* of an information source is the precision of the idiosyncratic noise; i.e., $1/\kappa_A^2$.[7] Loosely, the content is the level of detail an information source provides about the fundamental, whereas the clarity is how easily the source content is commonly interpreted (understood) by the players.

A source with higher clarity generates signals with more correlated noises, and thus, helps players coordinate better with each other. To see this, note that for any source A, the conditional correlation between players' signals is

$$\rho_A \equiv Corr(x_A^i, x_A^j|\theta) = \frac{1/\kappa_A^2}{1/\alpha_A^2 + 1/\kappa_A^2}.$$

The correlation increases with clarity, but decreases with the content of a source. The intuition is that as clarity increases, the variation of the idiosyncratic noise decreases, and thus, the variation of a signal is primarily driven by the variation of the common noise.

The information structure assumed here is similar to Myatt and Wallace (2012) but with discrete information choice and finite number of agents. This structure blurs the traditional distinction between the public and private signals in the beauty contest literature (e.g., Morris and Shin, 2002). On the one hand, a signal with imperfect content but perfect clarity is a "pure public" signal—i.e., as $\kappa_A \to 0$, $\rho_A \to 1$. On the other hand, a signal with perfect content but imperfect clarity is a "pure private" signal—i.e., as $\alpha_A \to 0$, $\rho_A \to 0$. In our information structure, each signal is characterized not only by its variance, but also by its cross-sectional correlation among the players. As such, intermediate values of "publicity" are admissible in this

---

[7]The precision of the common noise is called "accuracy" in Myatt and Wallace (2012). The precision of the idiosyncratic noise is called "transparency" in Pavan (2014).

information structure by allowing a general correlation structure.

2.1 Signal Weights

In this section, we take players' information acquisition decisions as given, and derive the equilibrium weights on each signal in their optimal action choices in the beauty contest game. We allow players to play asymmetric strategies in both coordination and information acquisition stages.

Optimality requires that a player's action satisfies

$$a^i = (1-r)\mathbb{E}^i[\theta] + r\mathbb{E}^i[a^j], \tag{3}$$

where the expectation operator $\mathbb{E}^i[.] \equiv \mathbb{E}[.|\mathcal{I}^i]$, and $\mathcal{I}^i$ denotes the information set of player i. We focus on a linear (but not necessarily symmetric) equilibrium. Suppose that each player follows the following linear action strategy:

$$a^i = \bar{\theta} + \omega_A^i(x_A^i - \bar{\theta}) + \omega_B^i(x_B^i - \bar{\theta}) \tag{4}$$

$$a^j = \bar{\theta} + \omega_A^j(x_A^j - \bar{\theta}) + \omega_B^j(x_B^j - \bar{\theta}), \tag{5}$$

where a signal weight $\omega_A^i > 0$ if $z_A^i = 1$ and $\omega_A^i = 0$ if $z_A^i = 0$. The weight on the prior θ for player i is $\omega_0^i \equiv 1 - \omega_A^i - \omega_B^i$.

To derive equilibrium signal weights, we take players' information acquisition choices as given. For brevity of notation, we denote players' information choice profile as $ab(cd)$ for any $a, b, c, d \in \{1, 0\}$ such that $z_A^i = a$, $z_B^i = b$, $z_A^j = c$, and $z_B^j = d$. Here, we derive equilibrium signal weights for an asymmetric information choice profile $11(10)$, in which player i acquires signals from both sources, but player j acquires a signal from only source A. An example of the derivation of equilibrium weights for a symmetric information choice profile is shown in the appendix.

By (4) and (5), players' actions, for the information choice profile 11(10), are

$$a^i = \bar{\theta} + \omega_A^i \left(x_A^i - \bar{\theta} \right) + \omega_B^i \left(x_B^i - \bar{\theta} \right) \tag{6}$$

$$a^j = \bar{\theta} + \omega_A^j \left(x_A^j - \bar{\theta} \right). \tag{7}$$

By the optimality condition (3) of player i, $a^i = (1-r)\mathbb{E}\left[\theta | x_A^i, x_B^i \right] + r\mathbb{E}\left[a^j | x_A^i, x_B^i \right]$, which, after replacing the value of a^j from (7), yields

$$a^i = (1-r)\mathbb{E}\left[\theta | x_A^i, x_B^i \right] + r\mathbb{E}\left[\bar{\theta} + \hat{\omega}_A^j \left(x_A^j - \bar{\theta} \right) | x_A^i, x_B^i \right], \tag{8}$$

where $\hat{\omega}_A^j$ is player i's conjecture about player j's weight on signal A. Replacing the values of $\mathbb{E}\left[\theta | x_A^i, x_B^i \right]$ and $\mathbb{E}\left[(x_A^j - \bar{\theta}) | x_A^i, x_B^i \right]$ in (8), and comparing the coefficients of $\left(x_A^i - \bar{\theta} \right)$ and $\left(x_B^i - \bar{\theta} \right)$ with (6), we have

$$\omega_A^i = \frac{(1-r)\sigma_\theta^2 \left(\alpha_B^2 + \kappa_B^2 \right)}{D} + \left[\frac{r \left\{ \alpha_A^2 \left(\alpha_B^2 + \kappa_B^2 \right) + \sigma_\theta^2 \left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2 \right) \right\}}{D} \right] \hat{\omega}_A^j \tag{9}$$

$$\omega_B^i = \frac{(1-r)\sigma_\theta^2 \left(\alpha_A^2 + \kappa_A^2 \right)}{D} + \left(\frac{r\kappa_A^2 \sigma_\theta^2}{D} \right) \hat{\omega}_A^j, \tag{10}$$

where $D \equiv \left(\alpha_A^2 + \kappa_A^2 \right) \left(\alpha_B^2 + \kappa_B^2 \right) + \sigma_\theta^2 \left(\alpha_A^2 + \kappa_A^2 + \alpha_B^2 + \kappa_B^2 \right).$

Using an analogous argument, player j's optimal action is

$$a^j = (1-r)\mathbb{E}\left[\theta | x_A^j \right] + r\mathbb{E}\left[\bar{\theta} + \hat{\omega}_A^i \left(x_A^i - \bar{\theta} \right) + \hat{\omega}_B^i \left(x_B^i - \bar{\theta} \right) | x_A^j \right], \tag{11}$$

where $\hat{\omega}_A^i$ and $\hat{\omega}_B^i$ are player j's conjecture about player i's weights on signals A and B. Substituting the values of $\mathbb{E}\left[\theta | x_A^j \right], \mathbb{E}\left[(x_A^i - \bar{\theta}) | x_A^j \right]$, and $\mathbb{E}\left[\theta | x_A^j \right]$ in (11), and comparing the coefficient of $\left(x_A^j - \bar{\theta} \right)$ with (7), we have

$$\omega_A^j = \frac{(1-r)\sigma_\theta^2 + r \left(\sigma_\theta^2 + \alpha_A^2 \right) \hat{\omega}_A^i + r\sigma_\theta^2 \hat{\omega}_B^i}{\sigma_\theta^2 + \alpha_A^2 + \kappa_A^2}. \tag{12}$$

In equilibrium, $\hat{\omega}_n^l = \omega_n^l$ for any $l \in \{i, j\}$ and any $n \in \{A, B\}$. Replacing the values of ω_A^i

and ω_B^i from (9) and (10) in (12), we have $\omega_A^j = \frac{M}{N}$, where

$$M \equiv (1-r)\sigma_\theta^2 \left[\left\{ (1+r)\,\alpha_A^2 + \kappa_A^2 \right\} \left(\alpha_B^2 + \kappa_B^2 \right) + (1+r)\,\sigma_\theta^2 \left(\alpha_A^2 + \kappa_A^2 + \alpha_B^2 + \kappa_B^2 \right) \right]$$

$$N \equiv \left(\sigma_\theta^2 + \alpha_A^2 + \kappa_A^2 \right) D - r^2 \left(\sigma_\theta^2 + \alpha_A^2 \right) \left[\alpha_A^2 \left(\alpha_B^2 + \kappa_B^2 \right) + \sigma_\theta^2 \left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2 \right) \right] - r^2 \kappa_A^2 \sigma_\theta^4.$$

Values of weights ω_A^i and ω_B^i are obtained from (9) and (10) by substituting the value of ω_A^j.

2.2 Information Acquisition

Having derived equilibrium signal weights, we now work backwards to calculate player's optimal information acquisition choices. Each player decides whether to acquire or not acquire a signal from sources A and B to maximize his ex-ante expected utility. A players' ex-ante expected utility given equilibrium weights, and the cost of information acquisition is given by

$$U^i = \bar{u} - L^i - C\left(z_A^i, z_B^i \right), \tag{13}$$

where the expected losses (see the derivation in the appendix) are

$$
\begin{aligned}
L^i \left(\sigma_\theta^2, \alpha_n^2, \kappa_n^2 \right) &= (1-r)\mathbb{E}\left[(a^i - \theta)^2 \right] + r\mathbb{E}\left[(a^i - a^j)^2 \right] \\
&= \left[(1-r)\left(\omega_0^i \right)^2 + r\left(\omega_0^i - \omega_0^j \right)^2 \right] \sigma_\theta^2 + \\
&\quad \sum_{n=A,B} \left[(1-r)\left(\omega_n^i \right)^2 + r\left(\omega_n^i - \omega_n^j \right)^2 \right] \alpha_n^2 + \\
&\quad \sum_{n=A,B} \left[(1-r)\left(\omega_n^i \right)^2 + r\left\{ \left(\omega_n^i \right)^2 + \left(\omega_n^j \right)^2 \right\} \right] \kappa_n^2. \tag{14}
\end{aligned}
$$

For any information choice profile $ab(cd)$, we denote a player's expected utility by $U_{ab}^i(cd)$, and expected losses as $L_{ab}^i(cd)$. We also denote expected benefits from acquiring one or both

signals by

$$B_{1'b}^{i}\left(cd\right) \equiv L_{0b}^{i}\left(cd\right) - L_{1b}^{i}\left(cd\right)$$

$$B_{a1'}^{i}\left(cd\right) \equiv L_{a0}^{i}\left(cd\right) - L_{a1}^{i}\left(cd\right)$$

$$B_{11}^{i}\left(cd\right) \equiv L_{00}^{i}\left(cd\right) - L_{11}^{i}\left(cd\right).$$

For example, $B_{1'0}^{i}\left(11\right) = L_{00}^{i}\left(11\right) - L_{10}^{i}\left(11\right)$ denotes player i's expected benefits from acquiring a signal from source A (note the "′" on first "1") given that he does not acquire a signal from source B, and player j acquires signals from both sources.

We derive two main results. First, strategic complementarity in action choices leads to strategic complementarity in information acquisition decisions. That is, if a player wants to do what others do, then he wants to know what others know. This result is reminiscent of Hellwig and Veldkamp (2009), albeit with a different information structure. Second, signals from the source with higher clarity are acquired in a greater range of parameter sets than those from the source with lower clarity, even though the latter source has higher content. Multiple equilibria in players' information acquisition choices resurface because unlike in Myatt and Wallace (2012), players' acquisition decisions in our model are discrete. We use a numerical simulation to derive these results.

Choice of parameter values. To conduct a horse race of content versus clarity, we make the total variance of each source the same, and vary their contents and clarifies. Source A has higher clarity, but lower content, whereas source B has lower clarity but higher content. The total variance (conditional on θ) of each source is the same; i.e., $\alpha_A^2 + \kappa_A^2 = \alpha_B^2 + \kappa_B^2$. To examine the differential impact of content and clarity, we manipulate the complementarity and cost of information parameters. For the complementarity parameter r, we choose one high ($r = 0.8$) and one low ($r = 0.2$) value. The cost parameter c measures the amount a player pays to receive a signal from a source. Since both sources have the same total variance, players pay the same cost for a signal from each source. We choose one high ($c = 200$) and one low ($c = 5$) value to ensure separation of players' behaviors. The variance of the prior distribution is chosen to be high

$(\sigma_\theta^2 = 200)$ to minimize its impact—and maximize the impact of the two information sources—on players' equilibrium behaviors. The complete set of parameter values are:

$$\bar{\theta} = 500, \ \sigma_\theta = 200, \ \alpha_A = \kappa_B = 30, \ \alpha_B = \ \kappa_A = 5, \ r \in \{0.2, 0.8\}, \ c \in \{5, 200\}, \ \bar{u} = 3000. \quad (15)$$

In result 1, we compare players' expected benefits with the costs of acquiring one or both signals. Players benefit from acquiring a signal that the other players also acquire, because knowing what others know help them coordinate better with each others' action choices. We show that a player's net expected benefits are higher when the other player also acquires the same signal(s).

Result 1 (Complementarity in information acquisition). *For a positive value of strategic complementarity, a player's expected benefits of acquiring a signal from an information source increase when the other player also acquires the signal from the same source. Given the parameter values in (15), the following inequalities hold*

i) $\quad B_{1'0}^i (10) > B_{1'0}^i (00) \, ; \ B_{01'}^i (01) > B_{01'}^i (00) \, ; \ B_{11}^i (11) > B_{11}^i (00)$

ii) $\quad B_{1'1}^i (11) > B_{1'1}^i (01) \, ; \ B_{11'}^i (11) > B_{11'}^i (10) \ .$

Table 1 presents the values of expected benefits for complementarity parameter values, $r \in \{0, 0.2, \ 0.8\}$. Expected benefits satisfy inequalities in Result 1 at $r = 0.2$ and $r = 0.8$, but not at $r = 0$.

[**Insert Table 1 near here**]

In result 2, we characterizes *all* information acquisition equilibria for *any* value of the cost parameter $c \in (0, \infty)$, and parameter values in (15), and $r = 0.8$. While our focus is on the two values of cost parameter (i.e., $c = 5$ and $c = 200$), we characterize equilibria for all values of c in order to have a comprehensive understanding of how information cost affects an agent's

information acquisition decisions.[8]

Result 2 (Information acquisition equilibria). *If $r = 0.8$, $c \in (0, \infty)$, and rest of the parameters follow* (15), *then*

i)   acquiring both signals is the unique *equilibrium if $c < 22$*

ii)   2 equilibria—acquiring only signal A and acquiring both signals—exist if $22 \leqslant c \leqslant 32$

iii) acquiring only signal A is the unique *equilibrium if $32 < c \leqslant 345$*

iv)   2 equilibria—acquiring only signal A and acquiring only signal B—exist if $345 < c \leqslant 1564$

v)   3 equilibria—acquiring only signal A, acquiring only signal B, and acquiring no signal—exist if $1564 < c \leqslant 7857$

vi)   2 equilibria—acquiring only signal A and acquiring no signal—exist if $7857 < c \leqslant 9032$

vii) acquiring no signal is the unique *equilibrium if $c > 9032$.*

Two points regarding Result 2 are noteworthy. One, with the exception of the highest cost range ($c > 9032$), acquiring signal A is always part of the set of equilibria. This is because signal A has higher clarity than signal B, even though A has lower content than B. Two, while we allow for asymmetric strategies for the players—in both the beauty contest and information acquisition stages—equilibria are always symmetric because of strategic complementarity. In the next section, we exploit the properties of these two results in greater detail to develop our hypotheses.

2.3   Hypotheses

We derive 5 hypotheses based on the parameter values stated in (15). Hypotheses 1-2 concern the use of signals in players' actions choices. Hypotheses 3-5 concern the acquisition of signals

---

[8]To save space, we do not show all the equilibria for $r = 0.2$. Equilibria are similar to Result 2, but ranges of cost parameter differ. For example, acquiring both signals is the unique equilibrium if $c < 290$, and acquiring no signal is the unique equilibrium if $c > 36271$. Other equilibria exist within $c \in [290, 36271]$.

14

from the 2 information sources.

Hypothesis 1 (Clarity versus content)**.** *Players place more weight on signal A than B in their action choices.*

Hypothesis 2 (Complementarity and signal weights)**.** *The weight on signal A increases, and the same on B decreases as the complementarity parameter increases.*

Hypothesis 3 (Complementarity in information acquisition)**.** *The likelihood of a player's acquiring signal B increases with her opponent's acquiring the same signal.*

Hypothesis 4 (Complementarity and likelihood of acquisition)**.** *As the complementarity parameter increases, the likelihood of a player's acquiring signal B decreases.*

Hypothesis 5 (Information cost and likelihood of acquisition)**.**

   a) *At low complementarity, information cost has no impact on the likelihood of a player's acquiring signal B;*

   b) *At high complementarity, information cost decreases the likelihood of a player's acquiring signal B.*

Hypothesis 1 follows from the fact that the ratio of the weights (derived in (A.5) and (A.6) in the appendix) is

$$\frac{\omega_A}{\omega_B} = \frac{(1-r)\alpha_B^2 + \kappa_B^2}{(1-r)\alpha_A^2 + \kappa_A^2}.$$

Thus, $\omega_A > \omega_B$ if $\frac{1}{\kappa_A^2} > \frac{1}{\kappa_B^2}$, that is, signal A is assigned more weight than signal B if the clarity of information source A is greater than the clarity of source B.

For Hypothesis 2, note that since $\alpha_A > \alpha_B$, and $\kappa_A < \kappa_B$,

$$\frac{d\omega_A}{dr} > 0 \ \text{ and } \ \frac{d\omega_B}{dr} < 0,$$

which implies that the weight on signal A increases, but the same on B decreases with the complementarity parameter r. In Table 2, last column, we show the equilibrium signal weights

for $r \in \{0.2, 0.8\}$. At $r = 0.2$, the weights are $\omega_A = 0.55$ and $\omega_B = 0.44$, whereas at $r = 0.8$, $\omega_A = 0.80$, and $\omega_B = 0.18$.[9]

[**Insert Table 2 near here**]

Hypothesis 3 follows directly from Result 1. Hypothesis 4 is based on Table 1, which shows that a player's expected benefits for a signal with higher clarity increases, and the same for a signal with lower clarity decreases with the complementarity parameter. As r increases from 0.2 to 0.8, $B_{1'1}(11)_{r=0.8} = 933 > 542 = B_{1'1}(11)_{r=0.2}$, but $B_{11'}(11)_{r=0.8} = 32 < 344 = B_{11'}(11)_{r=0.2}$.

Hypothesis 5 is based on Result 2 and Table 3 (last column),[10] which show that at low complementarity, a player acquires both signals at high as well as low information costs, whereas at high complementarity, he acquires both signals at low costs, but acquires *only* the signal with higher clarity at high costs. The equilibrium information acquisition decision is consistent with the fact that at low complementarity ($r = 0.2$), a player's expected benefits from acquiring either of the signals (conditional on the other player's acquiring both signals) are higher than the information costs under both high- and low-cost treatments (i.e., $B_{1'1}(11) = 542 > c$, and $B_{11'}(11) = 344 > c$) for any $c \in \{5, 200\}$. However, at high complementarity ($r = 0.8$), $B_{1'1}(11) = 933 > c$ for any $c \in \{5, 200\}$, whereas $B_{11'}(11) = 32 < c = 200$.

[**Insert Table 3 near here**]

---

[9]In the first four columns, we show theoretical signal weights assuming finite level of reasoning by the subjects. We discuss subjects' limited level of reasoning in Section 5.1. Signal weights under equilibrium conditions (equilibrium weights) and limited levels of reasoning are shown in the same table for the ease of comparison.

[10]In the first four columns, we show subjects' information acquisition decisions assuming limited levels of reasoning, which we discuss in detail in Section 5.1.

3 The Experiment

This section describes our experiment design and procedure.

3.1 Experimental Design

We manipulate two factors in our experiment: (i) strength of complementarity r, and (ii) information acquisition cost c. We adopt a 2×2, between-subject design (Low Complementarity/High Complementarity $\times$ Low Cost/High Cost). The parameter r is 0.2 in Low Complementarity condition, and 0.8 in High Complementarity condition. The information acquisition cost is 5 for each clue in the Low Cost condition, and 200 in the High Cost condition.

We conduct 3 sessions for each treatment and 12 sessions in total. Each session has 12 participants, divided into 2 groups of equal size. Each participant is matched with another participant in the same group. Participants interact only within their own group. In total, there are 6 independent group-level observations for each treatment. Table 4 provides a session summary.

[**Insert Table 4 here**]

Each experimental session has two stages. In the first stage, two clues, A and B, are provided to the subjects at no cost. This stage with exogenously provided signals allows subjects to learn the value of the two signals while making action choices in the beauty contest game. There are 10 rounds in this stage. After each round, subjects receive feedback about the true value of the fundamental θ, their partners' action choices, and their own payoffs. In the second stage, subjects first make information acquisition decisions, and then play the beauty contest game. In the information acquisition stage, subjects simultaneously decide whether to purchase clue A, clue B, both, or none of the clues by paying the required costs. After subjects make their own information acquisition decision, they are informed about their partners' information acquisition decisions. Participants then receive their chosen clue(s) and simultaneously make action choices

17

in the beauty contest game. There are 20 rounds in this stage. After each round, subjects receive feedback about the true value of θ, their partners' action choices and their own payoffs.

Each session has 30 rounds. Subjects are re-matched in each round within their group. We generate 30 independent values of θ, and 2 clues for each value of θ. The same sets of fundamental θ values and clues are used in all experimental sessions. The mean (standard deviation) of the realized θ values is 464 (175) and the values range from 144 to 805.

3.2 Experimental Procedure

We conducted our experiments using a computer program written in JAVA at a university computer laboratory during the 2017 spring. In total, 144 student subjects (undergraduate and graduate) participated in 12 sessions.

In each session, after reading the instructions, subjects then completed a quiz to test their understanding of the instructions. The experimenter then reviewed the instructions. Subjects participated first in the exogenous information (first) stage for 10 rounds and then in the endogenous information (second) stage for additional 20 rounds. In the second stage, subjects first made information acquisition decisions and then played the beauty contest game. At the end of the session, subjects completed a post-experiment questionnaire (the instruction, quiz, and the questionnaire are in the appendix). Screen shots of the two stages are shown in Figure 3.

Subjects were paid in cash after successfully completing the two stages. Subjects' payments were based on their performance, which was measured in a fictitious currency, "experimental francs," and then converted to U.S. dollars at the rate of $1 for every 200 francs. At the end of each session, we randomly selected two rounds from each stage and paid the subjects based on their cumulative payoffs from these four rounds. Each session lasted approximately 2 hours and the average cash payment was $26.9.

[Insert Figure 3 here]

4 Experimental Results

In this section, we first report the results of the stage with exogenously provided clues and then the stage with endogenously acquired clues.

4.1 Exogenous Clues (Stage 1)

We use the data from this stage to study how the subjects used the two clues to make action choices in the beauty contest game. Theory predicts that subjects choose actions that are weighted averages of the two clues and the prior. The higher the clue clarity, the greater the weight on that clue. We find that almost all actions fall between the minimum and maximum of the three sources of information. Ninety-nine percent (100%) of action choices fall in this interval for $r = 0.2$ ($r = 0.8$). Since the variance of the prior is much larger than the variance of the two clues, we expected subjects to put most of the weights on the two clues. We also check the percentage of actions that fall between the minimum and maximum of the two clues. Eighty-five percent (88%) of choices fall in this interval for $r = 0.2$ ($r = 0.8$). For actions within this interval, we report the percentage of actions that are closer to A, closer to B, or in the middle. As r increases, we expect a higher proportion of actions closer to A, the clue with higher clarity but lower content. About 47% (64%) of actions are closer to A for $r = 0.2$ ($r = 0.8$). Table 5 summarizes those results.

[Insert Table 5 here]

We estimate the relative weights on the two clues using subjects' decisions. In each session, subjects were divided into two groups and interacted only within their own group. We estimate the weights per group per session. In our statistical analyses, we use each of these group-specific estimates to test our hypotheses since the individual observations within each group are not independent. We run the following regression using data for individual i in each round t for each independent group:

$$a^{it} = \alpha + \beta_A x_A^{it} + \beta_B x_B^{it} + \epsilon^{it}. \tag{16}$$

The weight on clue A (B) is captured by the coefficient $\beta_A(\beta_B)$, the weight on the prior is $1 - \beta_A - \beta_B$. Panel (A) of Table 6 reports estimated average weights on the two clues for all groups in all sessions.

[Insert Table 6 here]

We estimate the weights using data from the first and second half of stage 1 separately to evaluate whether systematic trends exist over time and find no systematic changes from the first to the second half of data. We compare the average estimated weights with the theoretical predicted weights. For $r = 0.2$, the estimated weights on A is 0.5 for the first half, which is significantly smaller than the theoretical prediction 0.55 ($p = 0.02$ one-tailed Wilcoxon rank test); the estimated weights on A is 0.47 for the second half, which is also significantly smaller than the theoretical prediction ($p = 0.003$). For $r = 0.8$, the estimated weights on A is 0.61 using data from the first or second half, which are significantly smaller than the theoretical prediction 0.80 ($p = 0.0002$).

4.1.1 Hypotheses Tests

Our Hypotheses 1 and 2 are related to the relative importance of clues A and B in subjects' action choices. Results of the hypothesis tests are shown in panel (B) of Table 6. For $r = 0.2$, the estimated weights on A is not significantly different from the weights on B ($p = 0.83$). Thus, we reject H1 for $r = 0.8$. However, for $r = 0.2$, the estimated weights on A are significantly greater than the weights on B ($p = 0.0002$) using data from both first and second halves. Thus, we cannot reject H1 for $r = 0.8$.

We find evidence supporting H2. The estimated weights on A for $r = 0.8$ are significantly greater than the weights on A for $r = 0.2$ with p-value of 0.0003 and 0.0001 for the first and second half, respectively. In contrast, the estimated weights on B for $r = 0.8$ are significantly

20

smaller than the weights on B for $r = 0.2$ with p-value of 0.0003 and 0.0006, respectively.

4.2 Endogenous Information Acquisition (Stage 2)

4.2.1 Descriptive Evidence on Information Acquisition

There are four possible information acquisition strategies: buy A only, buy B only, buy both clues and buy no clues. Figure 1 plots the frequency of each strategy in each treatment over time.

[**Insert Figure 1 here**]

The frequency of buying no clues (depicted in circle) is close to 0 in all treatments. This strategy is strictly dominated by other strategies. The frequency of buying B (depicted in diamond) is also very small. This strategy is strictly dominated by buying A. The frequency of buying A is greater than buying B in all treatments. The two other strategies vary by treatments. For $r = 0.2$, the frequency of buying both clues (depicted in triangle) is the highest. For $r = 0.8$ and $c = 200$, the frequency of buying A (depicted in square) is the highest. For $r = 0.8$ and $c = 5$, the strategies of buying A and buying both have similar frequencies.

Table 7 reports the frequency of each strategy observed in each independent group in each session, using data from the first (second) half of each session.

[**Insert Table 7 here**]

There is no evidence of a time trend. The strategy is similar between the first ten rounds and last ten rounds. We use the last ten rounds for our following discussion. For $r = 0.2$, the frequency of buying both clues is the greatest among all four strategies and this frequency is 86% when the cost is 5 and 64% when the cost is 200. For $r = 0.8$, when the cost is 200, the frequency of buying A only is 77%, the highest among the four strategies; when the cost is 5, the frequency of buying A only is similar to the frequency of buying both clues. For $r = 0.8/c = 5$, among the six group-level observations, one group converges towards buying A only with a frequency approaching 90% in the last ten rounds. Two groups converge towards buying both clues with a

frequency approaching 83% and 75%, respectively. For the other three groups, both strategies of buying A only and buying both clues are selected with high frequency.

To evaluate the stability of information acquisition strategy over time, we calculate the transition matrix of information acquisition strategies for all the four treatments using data from all rounds. Results are reported in Table 8.

[Insert Table 8 here]

For $r = 0.2$, buying both clues is the absorbent state. The likelihood of continuing to buy both clues is 98.31% when the cost is 5 and 89.37% when the cost is 200. For $r = 0.8$ and $c = 5$, buying A only and buying both clues are both absorbent states. The likelihood of continuing to buy A only is 84.08% and the likelihood of continuing to buy both clues is 85.55%. For $r = 0.8$ and $c = 200$, most observations are buying A only and the likelihood of continuing this strategy is 89.6%.

At the subject level, we observe that most individuals do not change their information acquisition decisions. Only 16 out of 144 subjects changed their information acquisition decision more than 5 times during stage 2. Nine subjects changed their information acquisition decision more than 10 times. We classify these type of subjects as "indeterminate." Table 9 reports the percentage of subjects adopting each information acquisition strategy.

[Insert Table 9 here]

4.2.2 Hypotheses Tests

Our Hypotheses 3, 4, and 5 are related to the acquisition of clue B because the acquisition of clue A does not vary with our manipulations. We plot the frequency of buying A and B separately in Figure 2.

[Insert Figure 2 here]

22

Figure 2 Panel A plots the frequency of buying A. There are no systematic differences in the frequency of buying A across various treatments. The frequency of buying A converges toward 100% in all treatments. Panel B of Figure 2 plots the frequency of buying B, which varies by treatments systematically. For $r = 0.2$, the frequency of buying B approaches 100% for the treatment with $c = 5$, and this frequency is reduced when the cost is 200. For $r = 0.8$, the frequency of buying B is the lowest in the treatment with $c = 200$, and this frequency is higher when $c = 5$. The frequency of buying B is always higher when $r = 0.2$ than when $r = 0.8$.

We run a Probit regression to test our hypotheses on the acquisition of clue B. The dependent variable, Buy_B, is a dummy variable which is 1 if a subject buys clue B and 0 otherwise. Table 10 reports the regression results. Standard errors are clustered at the subject level.

[Insert Table 10 here]

The explanatory variable $Partner\_BuyB$ is a dummy variable, which is 1 if the partner buys clue B and 0 otherwise. H3 predicts that a subject is more likely to acquire clue B if the matched partner also acquires clue B; therefore, the coefficient on $Partner\_BuyB$ is expected to be positive. Consistent with H3, we find that the coefficient on $Partner\_BuyB$ is 0.905, which is significantly positive ($p < 0.001$).

Hypothesis 4 predicts that the likelihood of buying B decreases with r. Our explanatory variable includes r, a dummy variable, which is 1 if $r = 0.8$ and 0 if $r = 0.2$. Consistent with H4, we find that the coefficient on r is -1.576, which is significantly negative ($p < 0.001$).

Hypothesis 5a predicts that cost of information acquisition has no effect on the likelihood of acquiring clue B for $r = 0.2$. In contrast, Hypothesis 5b predicts that cost of information acquisition reduces the likelihood of acquiring clue B for $r = 0.8$. We include a dummy variable Cost, which is 1 if $c = 200$ and 0 if $c = 5$. In addition, we include an interaction between r and Cost. We find that the coefficient on Cost is -0.945, which is significantly negative ($p < 0.01$). There is a significant main effect of cost on the acquisition of clue B. However, the interaction term between r and $Cost$ is not statistically significant, which suggests that the effect of cost on the acquisition of clue B does not differ between high and low r treatments. Our evidence

supports H5b, but rejects H5a. Higher cost of information acquisition reduces acquisition of clue B in both high and low r treatments.

5 Discussions

5.1 Limited Levels of Reasoning

In this section, we examine whether the incongruity between the observed and theoretical signal weights, and information acquisition decisions can be explained by subjects' limited levels (depths) of reasoning (level-k model) in the sense of Nagel (1995) and Stahl and Wilson (1994, 1995). The equilibrium weights we have developed in Section 2.1 are based on strong assumptions of players' rationality—i.e., they can make the fixed-point argument in solving Nash equilibria. Experimental evidence suggests that subjects' responses can systematically deviate from Nash equilibrium (e.g., Crawford et al., 2013). Subjects tend to avoid the fixed-point argument as well as the indefinitely iterated dominance reasoning required in equilibrium definition.

A level-k model is based on the assumption that subjects' behaviors can be classified by different depths of reasoning. A level-0 player chooses her behavior randomly without forming any beliefs about her opponent. A level-1 player believes that her opponent is a level-0 player, and best responds to this belief. Similarly, a level-2 player best responds to a level-1 opponent, a level-3 player to a level-2 opponent, and so on. In general, a level-k player best responds to a level-k-1 opponent by taking a level-k action. However, no player can anticipate how an opponent with the same or higher level (i.e., equally smart or smarter opponent) will act. This feature of a level-k model circumvents the need for a player to be able to use the fixed-point argument to solve a Nash equilibrium (e.g., Stahl and Wilson, 1994). Theoretically, as $k \to \infty$, a level-k action corresponds to the Nash equilibrium action.

Prior literature on beauty-contest games have used level-k models to explain subjects' nonequilibrium behavior (Cornand and Heinemann, 2013; Baeriswyl and Cornand, 2016). Unlike these studies, we have an additional stage of information acquisition. A level-k player in our setup

has a 2-dimensional strategy $\langle (\omega_{A,k}, \omega_{B,k}), (z_{A,k}, z_{B,k}) \rangle$, where $(\omega_{A,k}, \omega_{B,k})$ are optimal signal weights, and $(z_{A,k}, z_{B,k})$ are optimal information acquisition decisions. One innovation in our theory development is to derive information acquisition equilibria at limited levels of reasoning.

Consistent with the spirit of a level-k model, we start by assuming that a level-0 action has a uniform distribution over the action space. That is, a level-0 player's choice of signal weights and information acquisition decision are random and do not depend on information costs. We then recursively derive optimal signal weights and information acquisition decisions for level-1 upwards. Derivation of level-k signal weights and information acquisition equilibria are shown in the appendix.

Table 2 summarizes signal weights at level-1 to level-4, and at equilibrium for $r \in \{0.2, 0.8\}$. Comparing estimated group-specific weights reported in Table 6 with weights assuming limited levels of reasoning summarized in Table 2, the observed average weights are close to the weights of level-2 reasoning for $r = 0.8$. Using the second half data, the average estimated weights on signal A is 0.61 for $r = 0.8$, which is not significantly different from the weights of level-2 reasoning 0.68 ($p = 0.146$). The average estimated weights on signal B is 0.39, which is not significantly different from the weights of level-2 reasoning 0.30 ($p = 0.774$). For $r = 0.2$, however, the observed weights are close to the weights of level-1 reasoning. Using the second half data, the average estimated weights on A is 0.47 for $r = 0.8$, which is not significantly different from the weights of level-1 reasoning 0.49 ($p = 0.146$). The average estimated weights on B is 0.53, which is not significantly different from the weights of level-1 reasoning 0.49 ($p = 0.388$). The estimated weight on A is significantly smaller than the predicted weight on A of level-2 reasoning 0.54 ($p = 0.03$), and the estimated weights on B is significantly greater than the weights on A of level-2 reasoning 0.45 ($p = 0.019$). Our results are similar to the findings of Cornand and Heinemann (2013), which show that subjects' level of reasoning is higher for higher values of r.

Table 3 presents optimal information acquisition decisions at level-1 to level-4, and at equilibrium for $r \in \{0.2, 0.8\}$. It shows that at $r = 0.2$, the optimal information acquisition decision is acquiring both the signals at any level of reasoning. At $r = 0.8$, and $c = 200$, acquiring only signal one is the unique optimal decision at all levels of reasoning above level-1. However, at

$r = 0.8$, and $c = 5$, acquiring both signals is the unique optimal decision at all levels except level-3, where the optimal decision is acquiring only signal one with higher clarity. The intuition is that, while high complementarity increases a subject's likelihood of acquiring only signal A, low information cost enhances the odds of acquiring both signals. At level-3, the complementarity effect dominates the cost effect. In other levels, low costs overwhelm the complementarity effect. Table 11 summarizes the values of expected benefits at level-1 to level-4, and at equilibrium for $r \in \{0.2, 0.8\}$.

[**Insert Table 11 near here**]

5.2 Acquisition Equilibria with Estimated Weights

To further examine the difference between observed and predicted information acquisition decisions for the treatment $r = 0.8/c = 5$, we replace equilibrium signal weights at the beauty contest stage by their estimated weights, and then work backward to calculate optimal information acquisition decision. Specifically, we replace equilibrium signal weights for information profiles 11(11), 11(10), and 10(11) by their estimated weights from Table 6. For example, in information profile 11(11), we replace equilibrium signal weights of $\omega_A = 0.81$, $\omega_B = 0.18$ (Table 2) by their estimated weights of $\omega_A = 0.61$, $\omega_B = 0.39$ (last 10 rounds). Similarly, for information profiles 11(10) and 10(11), we replace equilibrium signal weights of $\omega_A = 0.87$, $\omega_B = 0.11$ (untabulated) by their estimated weights of $\omega_A = 0.51$, $\omega_B = 0.49$ (untabulated).[11] All other weights are kept at their equilibrium levels. The resulting information acquisition equilibria are: 10(10) and11(11), which are consistent with the experimental data.

Our theoretical analysis assuming that subjects use the experimentally estimated weights suggests that both buy A only and buy both clues are equilibria for the treatment $r = 0.8$ and $c = 5$. Results shown in Table 7 on the information acquisition by groups indicate the presence of two equilibria. The highest frequency of information acquisition strategy is to buy A only for

---

[11]For information choice profiles, 11(01) and 01(11), we left signal weights at their equilibrium values because we could not derive estimated weights due to the small sample size (N=9).

three out of the six groups, and to buy both clues for the other three groups. When there are multiple equilibria, subjects need to learn over time to coordinate on one of the two equilibria. The convergence toward these two equilibria appears to be incomplete in most groups. Subjects indicated in their post-experiment questionnaire that they were more likely to buy B if their partners bought B in the past round(s). To examine such a dynamic adaptation process, we run Probit regressions on subjects' decisions to buy B. The explanatory variables include own decision and partners' decision to acquire B in the past round. Lag_BuyB is 1 if a subject acquires B in the previous round and 0 otherwise. $Lag\_Partner\_BuyB$ is 1 if a subject's partner in the previous round acquires B and 0 otherwise. Our regression uses the data from the last ten rounds and standard errors are clustered at the group level. Results are shown in Table 12.

[Insert Table 12 near here]

One interesting observation in the Probit regression (Table 12) is that the $Lag\_Partner\_BuyB$ is significant and positive only in the treatment $r = 0.8/c = 5$. In other treatments, a player's decision to acquire signal B does not depend on his partner's past acquisition of B. For the treatment $r = 0.8/c = 5$, the regression coefficient on $Lag\_Partner\_BuyB$ is 0.552, which is significantly positive ($p = 0.022$). This evidence confirms the dynamics indicated in the post-experiment questionnaire. That is, subjects were more likely to buy B if their partners bought B in the previous round. This indicates that subjects tried to coordinate on one of the two equilibria, buy A only or buy both.

5.3   Clue Uses

Prior research in psychology has shown that people tend to use information more when it comes at a cost compared to when it is given for free. Thaler (1980) notes that "paying for the right to use a good or service will increase the rate at which the good will be utilized, ceteris paribus. This hypothesis will be referred to as the sunk cost effect." If this sunk cost effect is present in our setting, we then expect the use of information to differ when clues are exogenously given versus when clues are acquired endogenously. Here, we study whether such a sunk cost effect exists in

27

our setting. Specifically, we compare the weights subjects place on clues A and B when the clues are provided exogenously (Table 6) free of costs with those under endogenous information choices (Table 13) when subjects pay for clues. In Table 13, we estimate the weights on clue A and B using the regression model (16) for each treatment conditional on partners' clue choices. We test whether the weight on A is different from that on B. We pooled data from all sessions in each treatment because the number of observations in each session is small.

[**Insert Table 13 here**]

We separate our analysis using the first ten rounds and last ten rounds data so that this result can be compared with a similar analysis using data in stage 1 where both clues are exogenously given (see Table 6). For $r = 0.2$, the estimated weights on A and B are not significantly different from each other. For $r = 0.8$, only a small number of observations exist when the cost is 200; when the cost is 5, we find the estimated weight on A is significantly greater than the estimated weight on B using the last ten rounds data ($p = 0.04$). The estimated weights on A is 0.61 and on B is 0.38, which are similar to results in stage 1 where clues are exogenously given (Table 6). Overall, subjects' use of information under endogenous information choice is similar to that under exogenously provided information, implying the absence of the sunk cost effect.

6   Concluding Remarks

We study how beauty contest incentives affect players' acquisition and use of multiple information sources with different contents and clarities. We provide experimental evidence that strong beauty contest incentives lead players to completely ignore an information source with high content if it has low clarity. Instead, players acquire information from a source with higher clarity despite its lower content. Our theory predicts that, when information is exogenously provided, players' desire to coordinate induces them to under-weight information with low clarity and high content. Our experimental results suggest that the under-use of information with low clarity is weaker than theoretical predictions assuming an infinite level of higher order beliefs, but closer to level-2

reasoning in a cognitive behavioral model.

When players acquire information endogenously, our theory predicts that the coordination motive can induce players to completely ignore information with low clarity despite its high content. Our experimental evidence suggests that the under-acquisition of information with low clarity is even stronger than our theoretical predictions. In anticipation of the over-use of the less-clear information at the beauty contest stage, subjects underacquire the less-clear information more often than equilibrium predictions.

These results have important practical implications for a firm's disclosure choices when coordination motives are present. Investors can ignore a firm's disclosure if it is not sufficiently clear, even though the disclosure has high content with all the technical details. More damagingly, to gather information, investors focus their attention to outside non-firm information sources, which can be arguably more clear, but have lower content, than firm disclosures. In a post-truth age with so-called "fake news" and "alternative facts," this is a concern for firms as well as regulators and standard setters.

Our paper takes the first step to understanding the differential impacts of content and clarity on the use and acquisition decisions in situations in which coordination motives predominate. There are several avenues for future research. We assume information choice to be discrete; i.e., acquire or not acquire. Future studies may examine how a subject's attention choice changes continuously with a source's content and clarity. For example, conditional on acquiring a signal, does a subject's attention to that signal increase monotonically with the source content and clarity, or the attention changes non-monotonically? Another potential avenue for future research is to endogenize the beauty contest motive by examining investors' use and acquisition behaviors in short-horizon trading in financial markets.

APPENDIX

A Derivations

A.1 Equilibrium Signal Weights

Here, we derive signal weights for a symmetric information choice profile 11(11). Since both players acquire signals from both information sources, we conjecture that their linear action strategies follow (4) and (5). By (3), player i's optimal action is

$$a^i = (1-r)\mathbb{E}\left[\theta | x_A^i, x_B^i\right] + r\mathbb{E}\left[\bar{\theta} + \hat{\omega}_A^j\left(x_A^j - \bar{\theta}\right) + \hat{\omega}_B^j\left(x_B^j - \bar{\theta}\right) | x_A^i, x_B^i\right], \qquad \text{(A.1)}$$

where "hat" represents conjectured value. After replacing the values of $\mathbb{E}\left[\theta | x_A^i, x_B^i\right]$, $\mathbb{E}\left[x_A^j - \bar{\theta} | x_A^i, x_B^i\right]$, and $\mathbb{E}\left[x_B^j - \bar{\theta} | x_A^i, x_B^i\right]$, player i's optimal action is

$$
\begin{aligned}
a^i = (1-r)&\left[\bar{\theta} + \left\{\frac{\sigma_\theta^2\left(\alpha_B^2 + \kappa_B^2\right)}{D}\right\}\left(x_A^i - \bar{\theta}\right) + \left\{\frac{\sigma_\theta^2\left(\alpha_A^2 + \kappa_A^2\right)}{D}\right\}\left(x_B^i - \bar{\theta}\right)\right] + r\bar{\theta} \\
&+ r\hat{\omega}_A^j\left[\left\{\frac{\alpha_A^2\left(\alpha_B^2 + \kappa_B^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2\right)}{D}\right\}\left(x_A^i - \bar{\theta}\right) + \left\{\frac{\kappa_A^2\sigma_\theta^2}{D}\right\}\left(x_B^i - \bar{\theta}\right)\right] \\
&+ r\hat{\omega}_B^j\left[\left\{\frac{\kappa_B^2\sigma_\theta^2}{D}\right\}\left(x_A^i - \bar{\theta}\right) + \left\{\frac{\alpha_B^2\left(\alpha_A^2 + \kappa_A^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_A^2\right)}{D}\right\}\left(x_B^i - \bar{\theta}\right)\right]. \qquad \text{(A.2)}
\end{aligned}
$$

Comparing coefficients of $\left(x_A^i - \bar{\theta}\right)$ and $\left(x_B^i - \bar{\theta}\right)$ in (4) and (A.2), we have

$$\omega_A^i = \frac{(1-r)\sigma_\theta^2\left(\alpha_B^2 + \kappa_B^2\right) + r\hat{\omega}_A^j\left\{\alpha_A^2\left(\alpha_B^2 + \kappa_B^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2\right)\right\} + r\hat{\omega}_B^j\kappa_B^2\sigma_\theta^2}{D} \qquad \text{(A.3)}$$

$$\omega_B^i = \frac{(1-r)\sigma_\theta^2\left(\alpha_A^2 + \kappa_A^2\right) + r\hat{\omega}_A^j\kappa_A^2\sigma_\theta^2 + r\hat{\omega}_B^j\left\{\alpha_B^2\left(\alpha_A^2 + \kappa_A^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_A^2\right)\right\}}{D}. \qquad \text{(A.4)}$$

In equilibrium, $\hat{\omega}_n^l = \omega_n^l$ for any $l \in \{i, j\}$ and any $n \in \{A, B\}$. For a symmetric equilibrium, $\omega_A^i = \omega_A^i = \omega_A$ and $\omega_B^i = \omega_B^j = \omega_B$, which with (A.3) and (A.4) lead to two equations and two

unknowns. After some algebra, the equilibrium weights are:

$$\omega_A = \frac{\frac{1}{(1-r)\alpha_A^2+\kappa_A^2}}{\frac{1}{(1-r)\sigma_\theta^2} + \frac{1}{(1-r)\alpha_A^2+\kappa_A^2} + \frac{1}{(1-r)\alpha_B^2+\kappa_B^2}} \tag{A.5}$$

$$\omega_B = \frac{\frac{1}{(1-r)\alpha_B^2+\kappa_B^2}}{\frac{1}{(1-r)\sigma_\theta^2} + \frac{1}{(1-r)\alpha_A^2+\kappa_A^2} + \frac{1}{(1-r)\alpha_B^2+\kappa_B^2}}. \tag{A.6}$$

A.2 Expected Losses

Note that for any $n \in \{A, B\}$

$$a_i - \theta = \omega_0^i \zeta_\theta + \sum_{n=A,B} \omega_n^i \left(\eta_n + \varepsilon_n^i \right),$$

and

$$a_i - a_j = - \left(\omega_0^i - \omega_0^j \right) \zeta_\theta + \sum_{n=A,B} \left(\omega_n^i - \omega_n^j \right) \eta_n + \sum_{n=A,B} \left(\omega_n^i \varepsilon_n^i - \omega_n^j \varepsilon_n^j \right)$$

where $\omega_0^i = 1 - \omega_A^i - \omega_B^i$ and $\omega_0^j = 1 - \omega_A^j - \omega_B^j$. Thus

$$\mathbb{E}\left[(a_i - \theta)^2 \right] = \left(\omega_0^i \right)^2 \sigma_\theta^2 + \sum_{n=A,B} \left(\omega_n^i \right)^2 \alpha_n^2 + \sum_{n=A,B} \left(\omega_n^i \right)^2 \kappa_n^2 \tag{A.7}$$

$$\mathbb{E}\left[(a_i - a_j)^2 \right] = \left(\omega_0^i - \omega_0^j \right)^2 \sigma_\theta^2 + \sum_{n=A,B} \left(\omega_n^i - \omega_n^j \right)^2 \alpha_n^2 + \sum_{n=A,B} \left[\left(\omega_n^i \right)^2 + \left(\omega_n^j \right)^2 \right] \kappa_n^2. \tag{A.8}$$

A.3 Limited Levels of Reasoning

A.3.1 Level-k Signal Weights

Here, we derive the optimal level-k weights for the information choice profile 11(10). The weights for the rest of the cases can be derived using an analogous method. Like Nagel (1995) and Stahl and Wilson (1994), we assume that an level-0 action has a uniform distribution over the action space, which, in our case, is the reals. Following Cornand and Heinemann (2013); Baeriswyl and Cornand (2016), we define a level-1 action to be the first order expectation of the state conditional

31

on the player's information set:

$$a_1^i = \mathbb{E}\left[\theta | \mathcal{I}^i\right]. \tag{A.9}$$

Given the information choice profile 11(10), suppose that, for any $k > 0$, player i believes that player j attaches weight $\omega_{A,k}^j$ to the signal from A such that player j's level-k action is

$$a_k^j = \bar{\theta} + \omega_{A,k}^j(x_A^j - \bar{\theta}).$$

By the optimality condition (3), player i's best response (level-k+1 action) to j's action is

$$
\begin{aligned}
a_{k+1}^i &= (1-r)\mathbb{E}\left[\theta | x_A^i, x_B^i\right] + r\mathbb{E}\left[\bar{\theta} + \omega_{A,k}^j(x_A^j - \bar{\theta}) | x_A^i, x_B^i\right] \\
&= \bar{\theta} + \left[\frac{(1-r)\sigma_\theta^2\left(\alpha_B^2 + \kappa_B^2\right) + r\omega_{A,k}^j\left\{\alpha_A^2\left(\alpha_B^2 + \kappa_B^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2\right)\right\}}{D}\right]\left(x_A^i - \bar{\theta}\right) \\
&\quad + \left[\frac{(1-r)\sigma_\theta^2\left(\alpha_A^2 + \kappa_A^2\right) + r\omega_{A,k}^j\kappa_A^2\sigma_\theta^2}{D}\right]\left(x_B^i - \bar{\theta}\right),
\end{aligned}
$$

where

$$D \equiv \left(\alpha_A^2 + \kappa_A^2\right)\left(\alpha_B^2 + \kappa_B^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \kappa_A^2 + \alpha_B^2 + \kappa_B^2\right). \tag{A.10}$$

Thus, player i's level-k+1 signal weights are:

$$\omega_{A,k+1}^i = \frac{(1-r)\sigma_\theta^2\left(\alpha_B^2 + \kappa_B^2\right) + r\omega_{A,k}^j\left\{\alpha_A^2\left(\alpha_B^2 + \kappa_B^2\right) + \sigma_\theta^2\left(\alpha_A^2 + \alpha_B^2 + \kappa_B^2\right)\right\}}{D} \tag{A.11}$$

$$\omega_{B,k+1}^i = \frac{(1-r)\sigma_\theta^2\left(\alpha_A^2 + \kappa_A^2\right) + r\omega_{A,k}^j\kappa_A^2\sigma_\theta^2}{D}. \tag{A.12}$$

Player j believes that player i attaches weight $\omega_{A,k}^i$ to the signal from A, and $\omega_{B,k}^i$ to the signal from B such that player i's Lk action is

$$a_k^i = \bar{\theta} + \omega_{A,k}^i(x_A^i - \bar{\theta}) + \omega_{B,k}^i(x_B^i - \bar{\theta}). \tag{A.13}$$

Player j's best response to player i's action is

$$
\begin{aligned}
a^j_{k+1} &= (1-r)\mathbb{E}\left[\theta|x^j_A\right] + r\mathbb{E}\left[\bar{\theta} + \omega^i_{A,k}(x^j_A - \bar{\theta}) + \omega^i_{B,k}(x^j_B - \bar{\theta})|x^j_A\right] \\
&= \bar{\theta} + \left[\frac{(1-r)\sigma^2_\theta + r\omega^i_{A,k}\left(\sigma^2_\theta + \alpha^2_A\right) + r\omega^i_{B,k}\sigma^2_\theta}{\sigma^2_\theta + \alpha^2_A + \kappa^2_A}\right]\left(x^j_A - \bar{\theta}\right).
\end{aligned}
$$

Thus, j's level-k+1 signal weight is

$$
\omega^j_{1,k+1} = \frac{(1-r)\sigma^2_\theta + r\omega^i_{A,k}\left(\sigma^2_\theta + \alpha^2_A\right) + r\omega^i_{B,k}\sigma^2_\theta}{\sigma^2_\theta + \alpha^2_A + \kappa^2_A}. \tag{A.14}
$$

As defined in (A.9), player i's level-1 action is

$$
a^i_1 = \mathbb{E}\left[\theta|x^i_A, x^i_B\right] = \bar{\theta} + \left[\frac{\sigma^2_\theta\left(\alpha^2_B + \kappa^2_B\right)}{D}\right]\left(x^i_A - \bar{\theta}\right) + \left[\frac{\sigma^2_\theta\left(\alpha^2_A + \kappa^2_A\right)}{D}\right]\left(x^i_B - \bar{\theta}\right),
$$

such that player i's level-1 signal weights are

$$
\omega^i_{A,1} = \frac{\sigma^2_\theta\left(\alpha^2_B + \kappa^2_B\right)}{D}; \ \omega^i_{B,1} = \frac{\sigma^2_\theta\left(\alpha^2_A + \kappa^2_A\right)}{D}, \tag{A.15}
$$

where D is defined in (A.10). Similarly, player j's level-1 signal weight is

$$
\omega^j_{A,1} = \frac{\sigma^2_\theta}{\sigma^2_\theta + \alpha^2_A + \kappa^2_A}. \tag{A.16}
$$

Player i's level-2 signal weights, $\omega^i_{A,2}$ and $\omega^i_{B,2}$, are derived by replacing j's level-1 weights from (A.16) in (A.11)-(A.12). Similarly, use (A.15) in (A.14) to derive j's level-2 signal weight, $\omega^j_{A,2}$, which, in turn, is used to derive i's level-3 signal weights, $\omega^i_{A,3}$ and $\omega^i_{B,3}$, and so on.

A.3.2 Level-k Information Acquisition

For calculating the ex-ante expected utility for player i, we use Lk signal weights of player i, and level-k-1 weights of player j. Specifically, for any $k \geqslant 1$, i's level-k ex-ante expected utility is

$$U_k^i = \bar{u} - L_k^i - C\left(x^i\right), \tag{A.17}$$

where, for any $n \in \{A, B\}$, the expected loss function at level k is:

$$L_k^i = \left[(1-r)\left(\omega_{0,k}^i\right)^2 + r\left(\omega_{0,k}^i - \omega_{0,k-1}^j\right)^2\right]\sigma_\theta^2 \tag{A.18}$$

$$+ \sum_{n=1}^{2}\left[(1-r)\left(\omega_{n,k}^i\right)^2 + r\left(\omega_{n,k}^i - \omega_{n,k-1}^j\right)^2\right]\alpha_n^2$$

$$+ \sum_{n=1}^{2}\left[(1-r)\left(\omega_{n,k}^i\right)^2 + r\left\{\left(\omega_{n,k}^i\right)^2 + \left(\omega_{n,k-1}^j\right)^2\right\}\right]\kappa_n^2.$$

Using a similar notational convention as before, we denote a player's level-k ex-ante expected utility as

$$U_{k,ab}^i(cd) = U_{k,z_{A,k}^i=a,z_{B,k}^i=b}^i(z_{A,k-1}^j = c, z_{B,k-1}^i = d).$$

A level-k player uses these utilities to derive his optimal information acquisition decision.

For the assumed parameter set in (15), players' optimal information acquisition decisions at level-1-level-4 are shown in Table 3. For example, consider the case of $r = 0.8$ and $c = 5$. At level-1, a player's optimal acquisition decision is $(z_{A,1}^i = 1, z_{B,1}^i = 1)$. The methodology used for the derivation is the following. A level-1 player's ex-ante expected utilities (i.e., $U_{1,ab}^i(cd)$) are calculated using the values of $\{\omega_{A,1}^i, \omega_{B,1}^i\}$ (derived by using (A.9)). For a level-1 player, his opponent is a level-0 player, who places equal probability to any of the four possible information acquisition choices, $\{(1,1), (1,0), (0,1), (0,0)\}$. A level-1 player computes his expected payoffs

over the distribution of his level-0 opponent's choices as

$$U^i_{1,ab} \equiv \frac{1}{4} \left[U^i_{1,ab}(11) + U^i_{1,ab}(10) + U^i_{1,ab}(01) + U^i_{1,ab}(00) \right],$$

for any $a, b \in \{1, 0\}$. Thus, a level-1 optimal information acquisition decision is

$$(z^i_{A,1} = a^*, z^i_{B,1} = b^*) = \arg\max_{a,b} U^i_{1,ab},$$

that is, he chooses an information profile (a^*, b^*) that corresponds to the maximum of $U^i_{1,11}$, $U^i_{1,10}$, $U^i_{1,01}$ and $U^i_{1,00}$. The optimal acquisition decision turns out to be $(z^i_{A,1} = 1, z^i_{B,1} = 1)$.

A level-2 player uses his level-2 signal weights ($\omega^i_{A,2}$ and $\omega^i_{B,2}$), and his opponent's level-1 signal weights ($\omega^j_{A,1}$ and $\omega^j_{B,1}$) to derive his ex-ante expected utilities using (A.17) and (A.18). Since a level-1 opponent's optimal acquisition decision is $(z^j_{A,1} = 1, z^j_{B,1} = 1)$, a level-2 player's optimal information acquisition choice is given by

$$(z^i_{A,2} = a^*, z^i_{B,2} = b^*) = \arg\max_{a,b} U^i_{2,ab}(11),$$

that is, he chooses an information profile (a^*, b^*) that corresponds to the maximum of $U^i_{2,11}(11)$, $U^i_{2,10}(11)$, $U^i_{2,01}(11)$ and $U^i_{2,00}(11)$. As shown in Table 3, the level-2 optimal acquisition decision for $r = 0.8$ and $c = 5$ is $(z^i_{A,2} = 1, z^i_{B,2} = 1)$.

Similarly, a level-3 player uses his level-3 signal weights ($\omega^i_{A,3}$ and $\omega^i_{B,3}$), and his opponent's level-2 signal weights ($\omega^j_{A,2}$ and $\omega^j_{B,2}$) to derive his ex-ante expected utilities, and best responds to his opponent's level-2 optimal acquisition decision ($z^j_{A,2} = 1, z^j_{B,2} = 1$). As shown in Table 3, the level-3 optimal acquisition decision for $r = 0.8$ and $c = 5$ is $(z^i_{A,3} = 1, z^i_{B,3} = 0)$. Optimal information choices at level-4 and upwards are derived in a similar way.

Table 1: Expected Benefits of Information Acquisition

	$B_{1'0}(10)$	$B_{1'0}(00)$	$B_{01'}(01)$	$B_{01'}(00)$	$B_{11}(11)$	$B_{11}(00)$	$B_{1'1}(11)$	$B_{1'1}(01)$	$B_{11'}(11)$	$B_{11'}(10)$
$r = 0$	39096	39096	39096	39096	39543	39543	447	447	447	447
$r = 0.2$	36271	25021	35940	25021	36618	25307	542	437	344	290
$r = 0.8$	9032	1564	7857	1564	9065	1582	933	345	32	22

This table presents expected benefits of a player when he acquires one, two or none of the signals for parameter values in (15).

Table 2: Signal Weights

		level-1	level-2	level-3	level-4	...	Equilibrium Weights
$r = 0.8$							
	ω_A	0.49	0.68	0.75	0.78	...	0.80
	ω_B	0.49	0.30	0.23	0.20	...	0.18
$r = 0.2$							
	ω_A	0.49	0.54	0.55	0.55	...	0.55
	ω_B	0.49	0.45	0.44	0.44	...	0.44

This table shows theoretical weights on signals A and B—for exogenous information structure—for equilibrium as well as level-k models for parameter values in (15).

Table 3: Information Acquisition Decisions

		level-1	level-2	level-3	level-4	...	Equilibrium
$r = 0.8$							
	$c = 200$	$(1,1)$	$(1,0)$	$(1,0)$	$(1,0)$	...	$(1,0)$
	$c = 5$	$(1,1)$	$(1,1)$	$(1,0)$	$(1,1)$	...	$(1,1)$
$r = 0.2$							
	$c = 200$	$(1,1)$	$(1,1)$	$(1,1)$	$(1,1)$	...	$(1,1)$
	$c = 5$	$(1,1)$	$(1,1)$	$(1,1)$	$(1,1)$	...	$(1,1)$

This table shows information acquisition decisions for equilibrium and level-k models for parameter values in (15). Each decision shows player i's information acquisition choices for signals A and B, ($z_{A,k}^i \in \{1,0\}, z_{B,k}^i \in \{1,0\}$), at level-k reasoning. A player at level-k best responds to a player at level-k-1.

Table 4: Experimental Design

	Low Cost ($c = 5$)	High Cost ($c = 200$)
Low Coordination ($r = 0.2$)	HR/HC (3 sessions)	HR/LC (3 sessions)
High Coordination ($r = 0.8$)	LR/HC (3 sessions)	LR/LC (3 sessions)

This table summarizes the experimental design and the number of sessions. We adopt a $2X2$, between-subjects design that manipulates information acquisition cost (High Cost or Low Cost) and the coordination incentives parameter (High Coordination or Low Coordination). We conducted three experimental sessions for each treatment. Each session includes 30 rounds. In Round 1-10 (Stage 1), clue A and clue B are provided to subjects at no cost. In Round 11-30 (Stage 2), subjects make information acquisition decisions.

Table 5: Guesses relative to Clues (Stage 1)

	$r = 0.2$	$r = 0.8$
Inside $[Min(x_A, x_B), Max(x_A, x_B)]$	0.85	0.88
Closer to clue A	0.47	0.64
Middle	0.15	0.10
Closer to clue B	0.38	0.26
Inside $[Min(x_A, x_B, 500), Max(x_A, x_B, 500)]$	0.99	1.00
Number of Obs.	720	720

Table 6: Estimated Clue Weights by Group (Stage 1)

Panel A: Group specific weights on clues

Treatment	Session.Group	First Half		Second Half	
		Clue A	Clue B	Clue A	Clue B
$r = 0.2/c = 5$	1.1	0.56	0.42	0.42	0.59
	1.2	0.53	0.46	0.5	0.51
	2.1	0.47	0.53	0.4	0.63
	2.2	0.46	0.52	0.42	0.6
	3.1	0.52	0.45	0.5	0.5
	3.2	0.53	0.47	0.52	0.46
$r = 0.2/c = 200$	4.1	0.5	0.51	0.51	0.5
	4.2	0.51	0.48	0.49	0.49
	5.1	0.52	0.48	0.49	0.53
	5.2	0.6	0.4	0.58	0.42
	6.1	0.46	0.5	0.5	0.42
	6.2	0.31	0.66	0.28	0.69
	Average	0.5	0.49	0.47	0.53
	Theoretical	0.55	0.44	0.55	0.44
$r = 0.8/c = 5$	7.1	0.58	0.4	0.54	0.46
	7.2	0.69	0.31	0.69	0.31
	8.1	0.59	0.4	0.59	0.43
	8.2	0.69	0.3	0.7	0.27
	9.1	0.58	0.43	0.54	0.48
	9.2	0.62	0.38	0.64	0.38
$r = 0.8/c = 200$	10.1	0.55	0.44	0.57	0.41
	10.2	0.56	0.43	0.56	0.43
	11.1	0.56	0.43	0.58	0.42
	11.2	0.74	0.26	0.74	0.26
	12.1	0.58	0.41	0.56	0.44
	12.2	0.61	0.39	0.58	0.4
	Average	0.61	0.38	0.61	0.39
	Theoretical	0.8	0.18	0.8	0.18

Panel B: p-values for Wilcoxon signed rank tests (one-tailed)

	Null	Alternative	First Half	Second Half
H1 $r = 0.2$	$\omega_A = \omega_B$	$\omega_A > \omega_B$	0.39	0.83
$r = 0.2$	$\omega_A = 0.55$	$\omega_A < 0.55$	0.02	0.003
H1 $r = 0.8$	$\omega_A = \omega_B$	$\omega_A > \omega_B$	0.0002	0.0002
$r = 0.8$	$\omega_A = 0.81$	$\omega_A < 0.81$	0.0002	0.0002
H2	$\omega_A^{r=0.2} = \omega_A^{r=0.8}$	$\omega_A^{r=0.2} < \omega_A^{r=0.8}$	0.0003	0.0001
H2	$\omega_B^{r=0.2} = \omega_B^{r=0.8}$	$\omega_B^{r=0.2} > \omega_B^{r=0.8}$	0.0003	0.0006

Table 7: Clue Acquisitions by Group

Treatment	Session.Group	First Half				Second Half			
		Buy A	Buy B	Both	None	Buy A	Buy B	Both	None
$r = 0.2/c = 5$	1.1	3%	7%	88%	2%	0%	15%	85%	0%
	1.2	0%	0%	100%	0%	0%	0%	100%	0%
	2.1	23%	10%	67%	0%	25%	8%	67%	0%
	2.2	5%	0%	95%	0%	0%	0%	100%	0%
	3.1	0%	2%	98%	0%	0%	0%	100%	0%
	3.2	8%	22%	70%	0%	0%	33%	67%	0%
	Average	7%	7%	86%	0%	4%	9%	86%	0%
	Theoretical	0%	0%	100%	0%	0%	0%	100%	0%
$r = 0.2/c = 200$	4.1	50%	27%	17%	7%	70%	20%	10%	0%
	4.2	15%	37%	47%	2%	5%	25%	70%	0%
	5.1	7%	18%	67%	8%	3%	18%	78%	0%
	5.2	12%	3%	80%	5%	0%	0%	100%	0%
	6.1	37%	13%	50%	0%	33%	2%	65%	0%
	6.2	33%	13%	53%	0%	23%	17%	60%	0%
	Average	26%	19%	52%	4%	23%	14%	64%	0%
	Theoretical	0%	0%	100%	0%	0%	0%	100%	0%
$r = 0.8/c = 5$	7.1	53%	2%	42%	3%	58%	0%	42%	0%
	7.2	57%	0%	43%	0%	65%	2%	33%	0%
	8.1	53%	0%	47%	0%	38%	0%	62%	0%
	8.2	30%	0%	70%	0%	17%	0%	83%	0%
	9.1	23%	0%	77%	0%	25%	0%	75%	0%
	9.2	78%	8%	13%	0%	90%	5%	5%	0%
	Average	49%	2%	49%	1%	49%	1%	50%	0%
	Theoretical	0%	0%	100%	0%	0%	0%	100%	0%
$r = 0.8/c = 200$	10.1	67%	7%	18%	8%	88%	8%	3%	0%
	10.2	87%	0%	2%	12%	100%	0%	0%	0%
	11.1	48%	2%	47%	3%	42%	13%	43%	2%
	11.2	77%	2%	7%	15%	82%	0%	18%	0%
	12.1	60%	27%	8%	5%	77%	23%	0%	0%
	12.2	78%	20%	0%	2%	73%	27%	0%	0%
	Average	69%	9%	14%	8%	77%	12%	11%	0%
	Theoretical	100%	0%	0%	0%	100%	0%	0%	0%

Table 8: Transition Matrix of Information Acquisition Decisions

Panel A: $r = 0.2/c = 5$

Acquisition Strategy in t	Transition Probability from t to $t+1$				Obs.
	None	Buy A	Buy B	Buy Both	
None	0%	100%	0%	0%	1
Buy A	0%	70.27%	10.81%	18.92%	37
Buy B	1.79%	5.36%	83.93%	8.93%	56
Both	0%	0.68%	1.02%	**98.31**%	590

Panel B: $r = 0.2/c = 200$

Acquisition Strategy in t	Transition Probability from t to $t+1$				Obs.
	None	Buy A	Buy B	Buy Both	
None	30.77%	23.08%	15.38%	30.77%	13
Buy A	1.23%	70.55%	14.72%	13.5%	163
Buy B	3.54%	15.04%	59.29%	22.12%	113
Both	0%	6.58%	4.05%	**89.37**%	395

Panel C: $r = 0.8/c = 5$

Acquisition Strategy in t	Transition Probability from t to $t+1$				Obs.
	None	Buy A	Buy B	Buy Both	
None	50%	50%	0%	0%	2
Buy A	0.3%	**84.08**%	1.80%	13.81%	333
Buy B	0%	60.00%	20.00%	20.00%	10
Both	0%	13.86%	0.59%	**85.55**%	339

Panel D: r=0.8/c=200

Acquisition Strategy in t	Transition Probability from t to $t+1$				Obs.
	None	Buy A	Buy B	Buy Both	
None	35.71%	53.57%	3.57%	7.14%	28
Buy A	1.80%	**89.60**%	4.40%	4.20%	500
Buy B	2.78%	26.39%	66.67%	4.17%	72
Both	3.57%	21.43%	3.57%	71.43%	84

Table 9: Classification of individual by acquisition strategy

	Buy A	Buy B	Both	None	Indeterminate
$r = 0.2/c = 5$	3%	11%	86%	0%	0%
$r = 0.2/c = 200$	19%	14%	58%	0%	8%
$r = 0.8/c = 5$	42%	0%	50%	0%	8%
$r = 0.8/c = 200$	75%	6%	11%	0%	8%

Table 10: Probit Regression on Decisions to Acquire Clue B

Explanatory Variables	Dependent Variable	
	Buy_B	Buy_B
r	$-1.576 ***$	
	(0.386)	
Cost	$-0.945 **$	
	(0.403)	
Cost*r	0.186	
	(0.510)	
Partner_BuyB		$0.905 ***$
		(0.247)
Intercept	$1.593 ***$	-0.259
	(0.314)	(0.181)
Observations	2880	2880
R-squared	0.25	0.09

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors (in parentheses) are clustered at the group level.

Table 11: Expected Benefits for Limited Depths of Reasoning

	$r = 0.8$					$r = 0.2$				
	level-1	level-2	level-3	level-4	Eqm.	level-1	level-2	level-3	level-4	Eqm.
$B_{1'0}(10)$	39096	39058	9032	9032	9032	39096	39086	36271	36271	36271
$B_{1'0}(00)$	39096	1564	1564	1564	1564	39096	25021	25021	25021	25021
$B_{01'}(01)$	39096	37732	7748	7776	7857	39096	38753	35940	35940	35940
$B_{01'}(00)$	39096	1564	1564	1564	1564	39096	25021	25021	25021	25021
$B_{11}(11)$	39543	39247	9019	9051	9065	39543	39456	36618	36618	36618
$B_{11}(00)$	39543	1582	1582	1582	1582	39543	25307	25307	25307	25307
$B_{1'1}(11)$	447	843	951	972	933	447	534	541	541	542
$B_{1'1}(01)$	447	413	387	380	345	447	438	437	437	437
$B_{11'}(11)$	447	170	-14	18	32	447	365	343	344	344
$B_{11'}(10)$	447	22	22	22	22	447	290	290	290	290

This table shows a player's expected benefits for different levels of reasoning.

Table 12: Probit Regression of the Effect of Past Behavior on Decision to Acquire B. (Last 10 Rounds)

	r=0.2/c=5	r=0.2/c=200	r=0.8/c=5	r=0.8/c=200
Lag_BuyB	3.621***	2.283***	2.391***	2.332***
	(0.348)	(0.607)	(0.243)	(0.184)
Lag_Partner_BuyB	0.231	0.489	0.552**	−0.128
	(0.233)	(0.344)	(0.241)	(0.163)
Intercept	−1.318***	−1.097***	−1.493***	−1.489***
	(0.211)	(0.143)	(0.106)	(0.233)
R square	0.71	0.46	0.5	0.44
Number of Obs.	360	360	360	360

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors (in parentheses) are clustered at the group level.

Table 13: Estimated Clue Weights (Stage 2) Conditional on Acquiring Both Clues

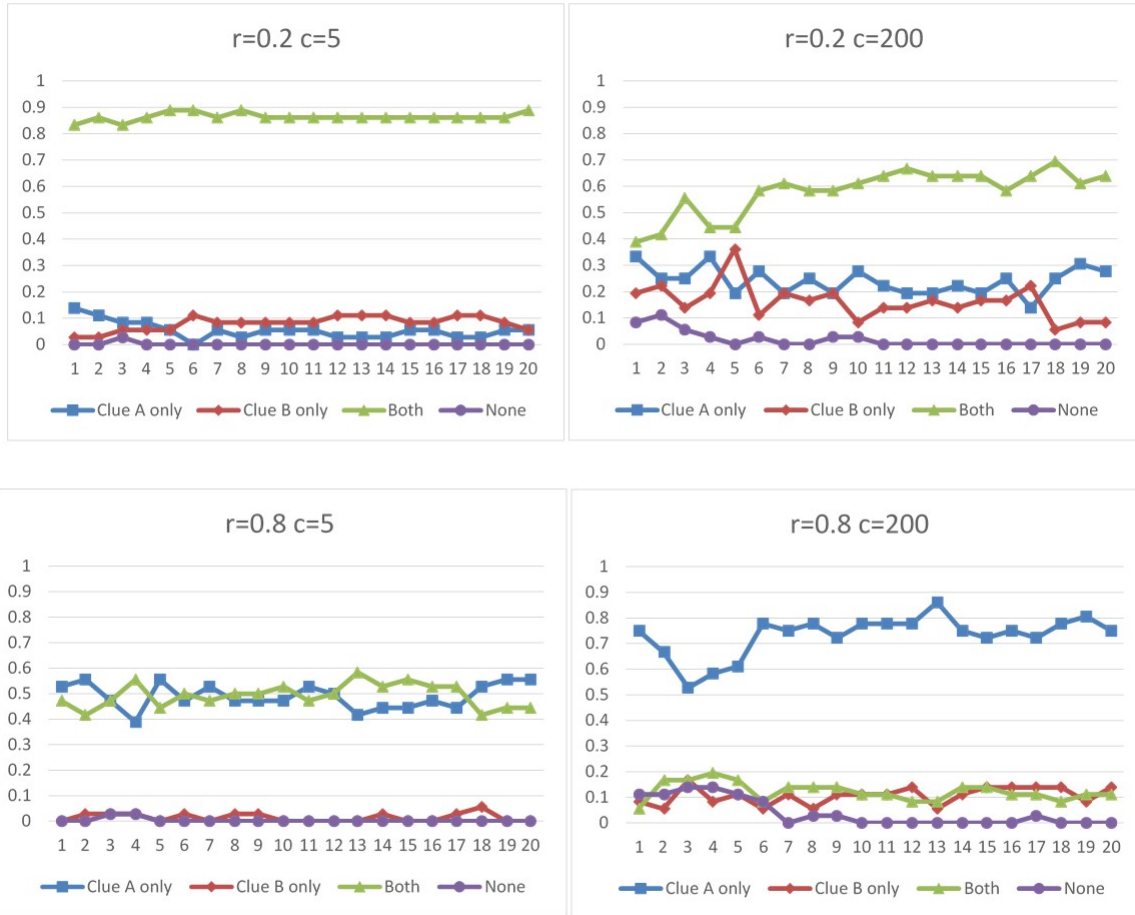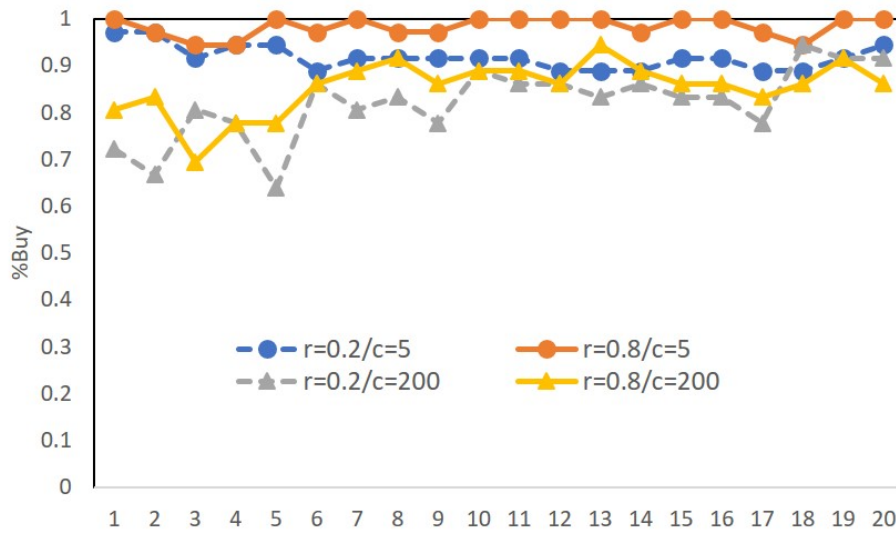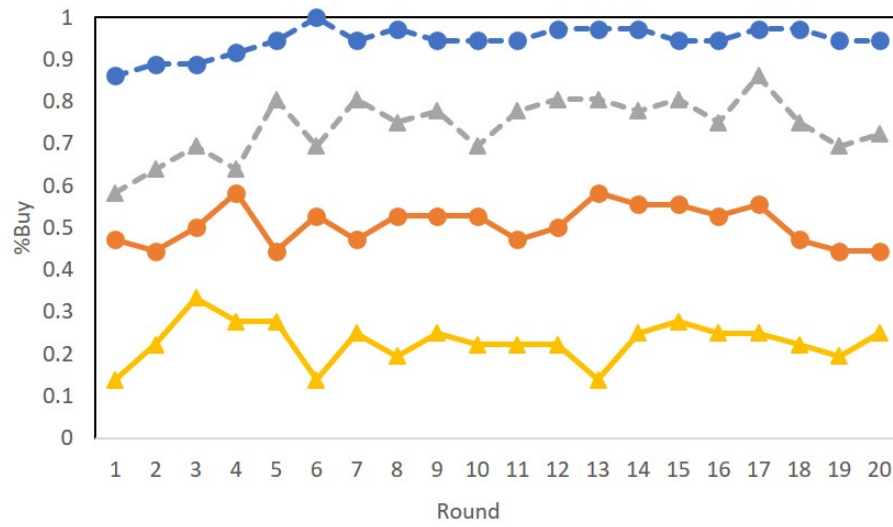|	First Half				Second Half			
	clue A	clue B	Diff. Sig.	N	clue A	clue B	Diff. Sig.	N
$r = 0.2/c = 5$	0.53	0.45	$p = 0.44$	246	0.50	0.49	$p = 0.84$	294
$r = 0.2/c = 200$	0.49	0.50	$p = 0.87$	98	0.52	0.46	$p = 0.76$	180
$r = 0.8/c = 5$	0.52	0.47	$p = 0.65$	74	0.61	0.38	$p = 0.04$	112
$r = 0.8/c = 200$	0.49	0.51	$p = 0.758$	12	0.50	0.50	$p = 0.99$	12

Figure 1: Information Acquisition Strategy by Treatment and Round

(A) Frequency of Clue A Acquired (Include Strategy of Buy A and Buy Both)

(B) Acquisition of Clue B by Treatment (Include Buy B and Buy Both)

Figure 2: Frequency of Clue A and B Acquired by Treatments

Experimental Instructions

[General Introduction]

This is an experiment in decision making under uncertainty. You will be paid in cash at the end of the experiment based on your performance.

If you have any questions during the experiment, please raise your hand and wait for an experimenter to come to you. Please do not talk or try to communicate with other participants during the experiment.

You are part of a group of 12 persons participating in this experiment and each of you is allocated into two groups of 6 persons. You will only interact with the other 5 persons in your group, but not interact with anyone in the other group. The two groups are completely independent.

This experiment has two parts. In each part, you play a guessing game. You play the same game repeatedly. In each round, you will be randomly matched with one of the other 5 persons in your group. In part one, there are 10 rounds. In part two, there are 20 rounds.

Your earnings are measured in francs, and your objective is to maximize your francs. **At the end of the experiment, the computer will randomly select four rounds, and the sum of the francs in the four rounds selected will be used to determine your cash payoff.** We will convert your francs into dollars at **500 francs = 1 US dollar**. Please read the instructions carefully and make your decisions wisely.

[Part 1]

There is an unknown number, Y. This unknown number is different in each round but identical for all the participants in the same group.

In each round, you are asked to make a guess. Your payoff in each stage is described below.

$$3000 - 0.8 \times (Your\ guess - Y)^2 - 0.2 \times (Your\ guess - Partner's\ guess)^2$$

Your payoff is higher the closer your guess is to the unknown number Y, and the closer your guess is to your partner's guess (about Y). To maximize your payoff, your guess has to be as close as possible to Y as well as to your partner's guess. However, note that your losses from deviating from Y are four times of your losses from deviating from your partner's guess. Thus, it is more important to be close to the unknown number Y than to the guess of your partner.

Example 1: Suppose the true value of Y is 400 and your partner's guess is 450. How much is your payoff in francs if your guess is 410?

Answer: 3000-0.8*(410-400)^2-0.2*(410-450)^2=3000-0.8*100-0.2*1600=3000-80-320=2600.

Example 2: Suppose that the true value of Y is 400 and that your partner's guess is 450, how much is your payoff in francs if your guess is 440?

Answer: 3000-0.8*(440-400)^2-0.2*(440-450)^2=3000-0.8*1600-0.2*100=3000-1280-20=1700.

As you can see from the examples above, you make more francs if your guess is closer to the true value of Y than if your guess is closer to the guess of your partner because your losses from the deviation from Y are four times that of your losses from the deviation from your partner's guess.

In this experiment, you and your partner do not know Y, but both of you will receive two clues about Y, which will inform you about Y. In the rest of the instructions, we first explain how Y is generated. Then we explain how clues are generated.

➤ **Unknown Number Y**

For each round, the computer will randomly select a number Y. **Y is drawn from a normal distribution with a mean of 500 and standard deviation of 200.** Below is a picture of the likelihood of drawing each possible Y:
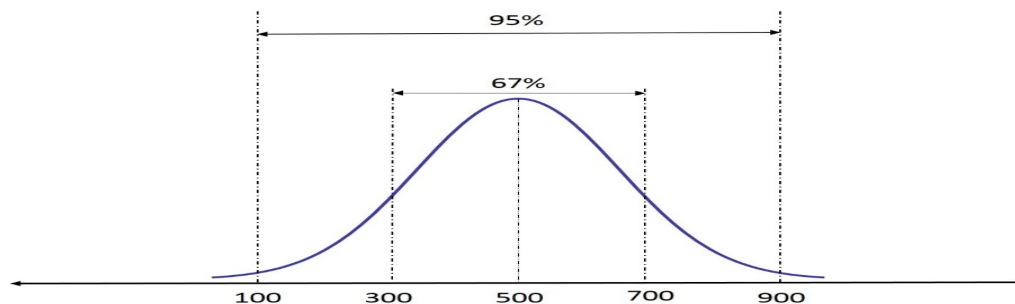
 As you can see from the picture, the number that is most likely to be drawn is the mean (500) and the chance of the other numbers being drawn decreases as they move further away from the mean. The standard deviation tells you how dispersed (i.e., spread out) the possible payoffs are. With approximately 67% chance, Y will be within one standard deviation (200) around the mean. This indicates that Y is between 300 (500–200) and 700 (500+200) about two-thirds of the time. With approximately 95% chance, Y will be within two standard deviations (200×2=400) around the mean 500. This indicates that Y is between 100 (500–400) and 900 (500+400) about 95% of the time.

In each round, the computer randomly draws a number according to the distribution outlined above. Different values of the number Y are drawn independently in each round, which means that the value of Y in one round is entirely unrelated to the values of Y in all of the other rounds.

➢ **Clues about Y**

Each player will receive two clues about Y. The two clues are Clue A and Clue B. The computer generates each of the two clues by adding two errors terms to Y. One error is a ***common error***, which is the same for you and your partner. The other error is a ***private error***, which is different between you and your partner. Thus, your clue A is likely to be different from your partner's clue A, and your clue B is likely to be different from your partner's clue B. Below we first explain how the computer generate two clue As, one for you and one for your partner.

Clue A

After the computer draws the unknown number Y, the computer generates Clue A in the following way.

First, the computer draws one common error. **This common error is randomly selected from a normal distribution with a mean 0 and a standard deviation 30**. The chance of the common error being 0 is the highest. With approximately 67% chance, the error will be within one standard deviation (30) around the mean (0), that is, between -30 (0–30) and 30 (0+30). With approximately 95% chance, the error will be within two standard deviations (30×2=60) around the mean 0, that is, between -60 and 60.

Second, the computer draws a private error for each participant independently. **The private errors are also randomly selected from a normal distribution with mean 0, but the standard deviation is 5**. **The computer draws the two private errors independently. Thus, the private error for you is likely different from the private error for your partner.** The chance of the private error being 0 is the highest. With approximately 67% chance, private errors will be within one standard deviation, i.e., between -5 and 5. With approximately 95% chance, private errors will be within two standard deviations, i.e., between -10 and 10.

Third, the computer calculates Clue A by first adding the common error to Y and then adding individual specific private error.

Your Clue A = Y + Common error A (std. dev. 30) + Your Private error A (std. dev. 5)

Partner Clue A = Y + Common error A (std. dev. 30) + Partner Private error A (std. dev. 5)

Note that the Clue A that you receive is likely different from the Clue A that your partner receives, because the private errors are different between you and your partner.

Example of Clue A:

Suppose Y drawn by the computer is 400. First, the computer draws a common error from the normal distribution with mean 0 and standard deviation 30. The computer adds this common error to Y to get Y', therefore Y' will be between 340 (400-2x30) and 460 (400+2x30) with 95% chance. Suppose the common error is -25. Y' =400-25=375.

Next the computer draws two private errors from the normal distribution with mean 0 and standard deviation 5. With 95% chance, the private errors drawn will be within two standard deviations, that is,

-10 and 10. The computer adds the private error to Y' to get Clue A. Suppose the private error drawn for you is -3; the Clue A you receive will be 375-3=372. Suppose the private error drawn for your partner is +5; the Clue A your partner receives will be 375+5=380.

Your clue A: 372   Your partner clue A: 380
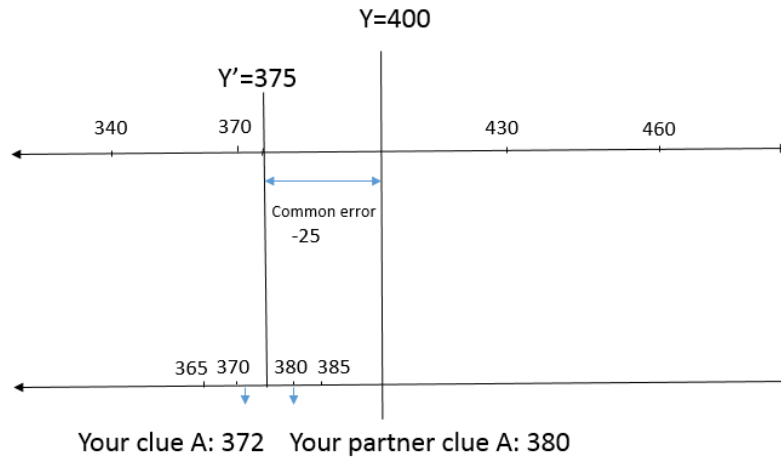
In this example, your clue A is 372 and your partner clue A is 380. The true Y is 400. Both clues are different from Y because of errors contained in the clues. Your clue A is different from your partner's clue A because the private errors are different. However, the private error is relatively small, so your clue A and your partner clue A are not that far apart.

The table below provides 10 examples of Clue As for a given Y. Please see how your clue is different from Y and different from your partner's Clue A.

Example of signal draws (Clue A)

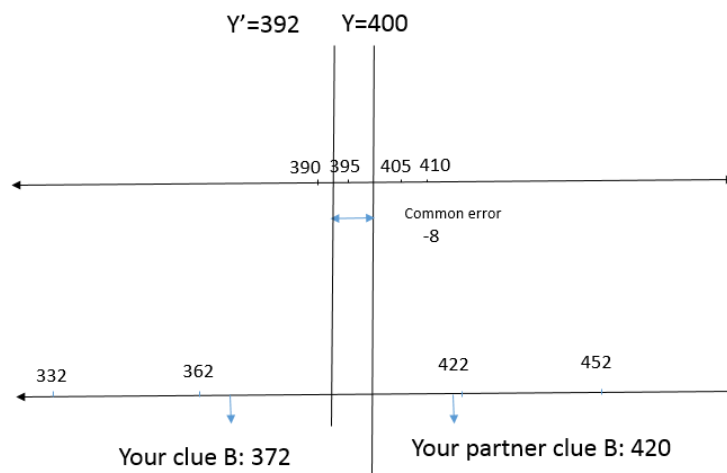| Y	Common Error (c)	Your Private Error (p)	Your partner's Private Error (p')	Your Clue A (Y+c+p)	Your partner's Clue A (Y+c+p')
386	16	-1	-1	401	401
620	39	-4	-1	655	658
670	-8	4	-2	666	660
120	1	7	8	128	129
874	13	0	-6	887	881
622	-12	1	3	611	613
483	30	6	0	519	513
627	9	-6	1	630	638
189	64	-2	3	251	256
342	-49	6	-1	299	292

Clue B

Clue B is generated in the same way as Clue A, except that the standard deviations of the errors are different. **For Clue B, the standard deviation of the common error is 5, and the standard deviation of the private error is 30.** The procedure to draw Clue B is the same as Clue A.

Example of Clue B:

Suppose Y drawn by the computer is 400. First, the computer draws a common error from the normal distribution with mean 0 and standard deviation 5. Suppose the common error is -8. Y''=400-8=392.

Next the computer draws two private errors from the normal distribution with mean 0 and standard deviation 30. Suppose the private error drawn for you is -20; the Clue B you receive will be 392-20=372. Suppose the private error drawn for your partner is +28; the Clue B your partner receives will be 392+28=420.

In this example, your clue B is 372 and your partner's clue B is 420. The true Y is 400. Your clue B is different from your partner clue B because the private errors are different. Since the private error in clue B is relatively large compared to clue A, your clue B and your partner clue B are likely to differ more from each other than your clue A and your partner clue A.

The table below provides 10 examples of Clue Bs for a given Y. Please see how your Clue B deviates from Y and from your partner's Clue B.

Example of signal draws (Clue B)

Y	Common Error (c)	Your Private Error (p)	Your partner's Private Error (p')	Your Clue A (Y+c+p)	Your partner's Clue A (Y+c+p')
386	-9	22	-6	399	371
620	-8	14	-10	626	602
670	1	-5	8	666	679
120	6	40	4	166	130
874	-1	9	12	882	885
622	4	45	-13	671	613
483	-1	7	-53	489	429
627	9	2	-17	638	619
189	2	24	63	215	254
342	9	21	-7	372	344

Clue Summary

Each player has two clues about Y: Clue A and Clue B. The computer generates each clue independently by adding to Y two error terms, one common error and one private error.

Your Clue A = Y + Common error A (std. dev. 30) + Private error A (std. dev. 5)

Your Clue B = Y + Common error B (std. dev. 5) + Private error B (std. dev. 30)

On average, you expect the size of the total error including common and private errors to be the same for Clue A and Clue B. Thus, you expect both clues to be equally informative about the unknown number Y.

Similarly, the computer generates two clues about Y for your partner.

Partner Clue A = Y + Common error A (std. dev. 30) + Partner Private error A (std. dev. 5)

Partner Clue B = Y + Common error B (std. dev. 5) + Partner Private error B (std. dev. 30).

Both the Clue A and Clue B are most likely to be different between you and your partner, because private errors are independently drawn. **Since the private errors for Clue B is more dispersed (standard error is 30) than the private errors for Clue A (standard error is 5), you expect to see that your Clue B and your partner's Clue B are farther apart from each other than your Clue A and your Partner's Clue A.**

After all participants submit their decisions for the guessing game, you will receive the feedback information including the Y, your partner's clues, your partner's guess, and your payoff.

[Part 2]

Your task is the same as Part I except that you have to now purchase the two clues at a cost. Clue A and Clue B are available to you only if you pay a cost to buy them. Before you play the guessing game, you have to decide whether you want to purchase the clues. **If you buy one clue, you pay a cost of 5 francs. The cost is the same for Clue A and Clue B. If you buy both the clues, you pay a total cost of 10 francs.** The costs are the same for you and your partner.

You select among one of the 4 choices: **buy none, buy only Clue A, buy only Clue B, or buy both clues**. Similarly, your partner also has to decide whether to purchase clues at cost and choose among the above 4 choices. You and your partner make clue purchasing decisions simultaneously. After you and your partner submit clue purchasing decisions, you will be informed about each other's clue purchasing decisions. The computer will send you the clues you purchased, and send your partner the clues he/she purchased. Finally, both of you make your decisions in the guessing game.

After all participants submit their decisions for the guessing game, you will receive the feedback information including the Y, your partner's clue purchase decision, the clues, your partner's guess, and your payoff.

Quiz

Q1. The random number Y is equally likely to be a number between 100 and 900.

_____ True   _____ False

Answer: F

Q2. If the Y in the current round is 700, then Y in the next round is likely to be close to 700.

_____ True   _____ False

Answer: F

Q3. Your clue A is the same as your partner's clue A. Your clue B is the same as your partner's clue B.

_____ True   _____ False

Answer: F

Q4. The common error in Clue A is the same for both players, but the private errors are different for you and your partner.

_____ True   _____ False

Answer: T

Q5. The common error for Clue A has mean of 0 and standard deviation of 30. On average, the error is expected to be 0. In 95% of cases, the error will be within two standard deviations. That is, the common error will be between _____and _____in 95% of cases.

Answer: -60 and +60

Q6. The common error for Clue B has mean of 0 and standard deviation of 5. On average, the error is expected to be 0. In 95% of cases, the error will be within two standard deviations. That is, the common error will be between _____and _____in 95% of cases.

Answer: -10 and 10

Q7. The computer draws the private errors for you and your partner independently. It is most likely that your private error is different from your partner's private error.

_____ True   _____ False

Answer: T

Q8. On average, you expect that Clue A and Clue B are equally informative about the true value Y.

_____ True   _____ False

Answer: T

Q9. On average, you expect that your Clue A is closer to your partner's Clue A than your Clue B to your partner's Clue B because the private errors of Clue A have smaller dispersions than the private errors of Clue B.

_____ True   _____ False

Answer: T

Q10: In Part 1, you are given Clue A and Clue B. Suppose that the true value of Y is 600 and that your partner's guess is 650, how much is your payoff in francs if your guess is 640?_____

Answer: 3000-0.8*(640-600)^2-0.2*(640-650)^2 =3000-0.8*1600-0.2*100 = 1700.

Q11. In part 2, you have the option to buy clues at a cost. How much do you have to pay to buy one clue?_____How much do you have to pay to buy two clues?_____

Answer: 5; 10

Q12: In part 2, suppose that the true value of Y is 600 and that your partner's guess is 650, how much is your payoff in francs if your guess is 640 and you bought both clue A and B?_____

Answer: 3000-0.8*(640-600)^2-0.2*(640-650)^2 – 10 =3000-0.8*1600-0.2*100 – 10 = 1690.

Q13. In part 2, you are not able to know which clues your partner has purchased before you submit your guess.

_____ True   _____ False

Answer: F

Post-Experimental Questionnaire

Major_____

Age_____

Gender_____

1. In Part 1, you are given two clues. How did you use the two clues? Did you use both clues in the same way or differently?

2. In Part 2, you have options to buy two clues. How did you make the clue-buying decisions? Did your clue-buying decision change over time after you saw other players' decision (e.g., other players' clue-buying decisions and guessing decisions)?

3. In Part 2, how did you use the clue or clues you buy to make your decisions in the guessing game?

4. Were the instructions clear? If not, which parts were not clear?

(A) Stage 1 (exogenous information)

Instructions | Quiz | **Experiment** | Questionnaire | Submission | Finish

Information

Part: 1
Round: 2
Payoff Function:
3000-.2*(Your guess-Y)^2-.8*(Your guess-Partner's guess)^2
Note: Your payoff is higher the closer your guess is to the unknown number Y,and the closer your guess is to your partner's guess (about Y). To maximize your payoff,your guess has to be as close as possible to Y as well as to your partner's guess. However,note that your losses from deviating from your partner's guess are four times of your losses from deviating from Y. Thus,it is more important to be close to the guess of your partner than to the unknown number Y.

731 (Clue A)   783 (Clue B)

My Guess: [____] Confirm

History Information

Part #	Round #	Y	My Clue A	My Clue B	Partner Clue A	Partner Clue B	My Guess	Partner Guess	My Points
1	1	693	689	610	591	677	600.0	650.0	990.20

(B) Stage 2 (endogenous information)

Instructions | Quiz | **Experiment** | Questionnaire | Submission | Finish

Information

Part: 2
Round: 3
Payoff Function:
3000-.2*(Your guess-Y)^2-.8*(Your guess-Partner's guess)^2-the cost of the clue(s) you purchased.
Note: Your task is the same as Part 1 except that you have to now purchase the two clues at a cost. If you buy one clue,you pay a cost of 200 francs. The cost is the same for Clue A and Clue B. If you buy both the clues,you pay a total cost of 400 francs.

531 (Clue B)

My Guess: [540] Confirm

Option

Note:
If you buy 1 clue it will **cost 200**. If you buy 2 clues it will **cost 400**.

○ Buy Clue A Only
◉ Buy Clue B Only
○ Buy Both
○ Buy Neither

Your partner bought Clue A

Confirm

History Information

Part #	Round #	Y	My Clue A	My Clue B	Partner Clue A	Partner Clue B	My Guess	Partner Guess	My Points
2	2	433	444	N/A	449	N/A	450.0	500.0	742.20
2	1	408	374	N/A	N/A	426	400.0	410.0	2707.20
1	10	590	564	582	545	667	560.0	600.0	1540.00
1	9	447	442	424	461	442	430.0	450.0	2622.20
1	8	438	372	443	383	433	400.0	400.0	2711.20
1	7	371	338	370	337	364	360.0	340.0	2831.80
1	6	660	681	699	691	623	700.0	650.0	680.00
1	5	805	787	805	788	812	800.0	800.0	2996.00
1	4	317	286	333	293	311	300.0	300.0	2942.20
1	3	246	282	284	293	292	280.0	300.0	996.80
1	2	774	736	764	731	783	780.0	750.0	2272.80
1	1	693	591	677	589	610	650.0	600.0	350.20

Figure 3: Screenshots of two stages of the game.

References

Aboody, D. (1996). Recognition versus disclosure in the oil and gas industry. *Journal of Accounting Research*, pages 21–32.

Ahmed, A. S., Kilic, E., and Lobo, G. J. (2006). Does recognition versus disclosure matter? evidence from value-relevance of banks' recognized and disclosed derivative financial instruments. *The Accounting Review*, 81(3):567–588.

Allen, F., Morris, S., and Shin, H. S. (2006). Beauty contests and iterated expectations in asset markets. *The Review of Financial Studies*, 19(3):719–752.

Anctil, R. M., Dickhaut, J., Kanodia, C., and Shapiro, B. (2004). Information transparency and coordination failure: Theory and experiment. *Journal of Accounting Research*, 42(2):159–195.

Arya, A. and Mittendorf, B. (2016). On the synergy between disclosure and investment beauty contests. *Journal of Accounting and Economics*, 61(2):255–273.

Baeriswyl, R. and Cornand, C. (2016). The predominant role of signal precision in experimental beauty contests. *The BE Journal of Theoretical Economics*, 16(1):267–301.

Banerjee, S. and Maier, M. (2016). Public information precision and coordination failure: An experiment. *Journal of Accounting Research*, 54(4):941–986.

Chamley, C. (2007). Complementarities in information acquisition with short-term trades. *Theoretical Economics*, 2(4):441–467.

Chen, Q., Huang, Z., and Zhang, Y. (2014). The effects of public information with asymmetrically informed short-horizon investors. *Journal of Accounting Research*, 52(3):635–669.

Chen, Q., Lewis, T. R., Schipper, K., and Zhang, Y. (2017). Uniform versus discretionary regimes in reporting information with unverifiable precision and a coordination role. *Journal of Accounting Research*, 55(1):153–196.

Cornand, C. and Heinemann, F. (2013). Measuring agents' reaction to private and public information in games with strategic complementarities. *Experimental Economics*, (1):1–17.

Cornand, C. and Heinemann, F. (2014). Measuring agents reaction to private and public information in games with strategic complementarities. *Experimental Economics*, 17(1):61–77.

Cornand, C. and Heinemann, F. (2015). Limited higher order beliefs and the welfare effects of public information. *Journal of Economic Studies*, 42(6):1005–1028.

Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1):5–62.

Davis-Friday, P. Y., Folami, L. B., Liu, C.-S., and Mittelstaedt, H. F. (1999). The value relevance of financial statement recognition vs. disclosure: Evidence from sfas no. 106. *The Accounting Review*, 74(4):403–423.

Engelberg, J., Sasseville, C., and Williams, J. (2012). Market madness? the case of mad money. *Management Science*, 58(2):351–364.

Gao, P. (2008). Keynesian beauty contest, accounting disclosure, and market efficiency. *Journal of Accounting Research*, 46(4):785–807.

Hellwig, C. and Veldkamp, L. (2009). Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies*, 76(1):223–251.

Holthausen, R. W. and Verrecchia, R. E. (1990). The effect of informedness and consensus on price and volume behavior. *Accounting Review*, pages 191–208.

Indjejikian, R. J. (1991). The impact of costly information interpretation on firm disclosure decisions. *Journal of Accounting Research*, pages 277–301.

Kim, O. and Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of accounting and economics*, 17(1):41–67.

KPMG (2011). Disclosure overload and complexity: Hidden in plain sight. *KPMG, Delaware.*

Liang, P. J. and Zhang, G. (2014). Information objectivity and accuracy in a bank run model.

Monga, V. and Chasan, E. (2015). The 109,894-word annual report: As regulators require more disclosures, 10-ks reach epic lengths; how much is too much. *Wall Street Journal June*, 2.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Morris, S. and Shin, H. S. (2007). Optimal communication. *Journal of the European Economic Association*, 5(2-3):594–602.

Myatt, D. P. and Wallace, C. (2012). Endogenous information acquisition in coordination games. *Review of Economic Studies*, 79(1):340–374.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326.

Pavan, A. (2014). *Attention, Coordination, and Bounded Recall.* Centre for Economic Policy Research.

Qu, H. (2013). How do market prices and cheap talk affect coordination? *Journal of Accounting Research*, 51(5):1221–1260.

Sapra, H. (2010). Discussion of expected mispricing: The joint influence of accounting transparency and investor base. *Journal of Accounting Research*, 48(2):383–391.

Schipper, K. (2007). Required disclosures in financial reports. *The Accounting Review*, 82(2):301–326.

Stahl, D. O. and Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, 25(3):309–327.

Stahl, D. O. and Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254.

Szkup, M. and Trevino, I. (2012). Costly information acquisition in a speculative attack: Theory and experiments. Technical report, mimeo.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.