Coordination with Flexible Information Acquisition

By MING YANG[*]

We study flexible information acquisition in a coordination game. "Flexible" acquisition means that players choose not only how much but also what kind of information to acquire. Information acquisition has a cost proportional to reduction of entropy. Hence, players will collect the information most relevant to their welfare but can be rationally inattentive to other aspects of the fundamental. When information is cheap, this flexibility enables players to acquire information that makes efficient coordination possible, which also leads to multiple equilibria. This result contrasts with the global game literature, where information structure is less flexible and cheap information leads to a unique equilibrium with inefficient coordination. We then go beyond the entropic information cost to set out the key aspects of flexibility and examine the way in which they drive our results.
JEL: C72 D82
Keywords: coordination game, flexible information acquisition, rational inattention

I. Introduction

This paper studies a coordination game in which players can acquire information flexibly. Coordination requires that all the players know the payoffs. In practice, payoffs may not be perfectly observable, so such common knowledge is lacking. The global game literature models this by endowing players with additive private signals on payoffs. Accordingly, researchers obtain a well-known limit unique equilibrium with inefficient coordination (e.g., (Hans Carlsson and Eric van Damme 1993), (David M. Frankel, Stephen Morris and Ady Pauzner 2003)). We re-examine this result in an environment where players can flexibly acquire information about the payoffs. Rather than fixing an additive information structure and letting players choose the precision of signals, we allow them to choose any information structures. An information structure exclusively specifies both the amount (measured by reduction of Shannon's entropy) and the substance of the information. Hence, information acquisition is flexible in the sense that the players choose not only how much but also what kind of information to acquire. This paper thus addresses the following questions: What information will be acquired under

this flexibility? And how does this flexibility affect the efficiency of coordination? These questions are interesting because a player's incentive to acquire information is shaped by his payoff structure, which further depends on his opponent's information acquisition. This strategic concern together with the flexibility of information acquisition leads to nontrivial implications.

We examine our concept of flexible information acquisition in the following game. Two players coordinate investment in a risky project whose future cash flow is driven by a randomly fluctuating fundamental. Each player decides whether to invest, but the payoff depends not only on the realized fundamental but also on the other player's action. Given the other player's choice, a player's payoff from investing increases as the fundamental improves. And for any given realization of the fundamental, a player's gain from "invest" over "not invest" is strictly greater when the other player also invests. That is, players' actions are strategic complements. Before making a decision, each player can independently purchase private information about the fundamental in the form of an information structure, i.e., the conditional distribution of the signal given the fundamental. The player then takes action according to the realized signal.

The players' information acquisition strategy is determined by two factors. The first is the effect of the fundamental: intuitively, given the other player's action, a player wants to collect information that induces investment with high (low) probability in the high (low) states of the fundamental. The second, and more interesting, factor is the player's incentive to match the other's informational choice, so as to minimize the probability of miscoordination. As a result, the strategic complementarity between actions creates a motive for the coordination of information as well. Indeed, coordination in information acquisition arises thanks to the fact that information acquisition is flexible. Especially, when the cost is low, players can coordinate to achieve approximate common knowledge of any "cutoff event" with the cutoff taking intermediate values.[1] Consequently, lowering the information cost makes possible efficient coordination in investing through efficient coordination in acquiring information; it also gives rise to multiple equilibria, insofar as there is a multiplicity of ways of coordinating information acquisition.

This paper contributes to our understanding of the impacts of information acquisition on coordination. Especially, a key feature, flexibility is highlighted. To see the point, consider an extended global game model where the players can purchase more accurate signals but cannot change the additive information structure. Intuitively, cheaper information induces the players to acquire more precise private signals, which lead to the inefficient, unique equilibrium commonly seen in standard global game models. The reason why efficient coordination is no longer sustainable and multiplicity disappears is precisely the rigidity of information acquisition, embodied in the constraint that players must pay equal attention to all possible realizations of the fundamental, as the additive information structure forces the observational error to be invariant in the fundamental. As a result, players coordinate only in choosing overall precision but cannot coordinate to achieve approximate common knowledge for any "cutoff event". This rigidity contrasts

---

[1]By approximate common knowledge, we mean the common-p belief (see (Dov Monderer and Dov Samet 1989)) with p close to 1.

sharply with the mechanism of flexible information acquisition described above.

We proceed as follows. Section II describes the model and sets out some simple facts about information acquisition behavior in equilibrium. In Section III, we characterize the equilibria and gain some initial insight through comparative statics. Section IV reports our main results, first comparing our approach with the extended global game model and exploring the origins of the difference and then comparing the welfare implications of flexible information acquisition in strategic and non-strategic settings. Finally, the section extends the concept of flexibility and discusses its essence. Section V examines the effects of public information through comparative statics with respect to the common prior. Section VI concludes with a discussion of several extensions of the benchmark model. Most of the proofs are given in the appendix.

Relation to the literature. The flexible information acquisition is modeled in the rational inattention framework of (Christopher A. Sims 2003). In applied work, rational inattention is studied chiefly in two cases: linear-quadratic (e.g., (Bartosz Mackowiak and Mirko Wiederholt 2009)) and binary-action. A prime instance of the latter is (Michael Woodford 2009). Our model too adopts the binary-action setup. Our work differs from most other rational inattention models in that we study information acquisition in a strategic environment rather than in decision problems.

Our model is closely related to the global games literature in that both study coordination under incomplete information. The global game models are characterized by additive information structures (e.g., (Stephen Morris and Hyun Song Shin 1998), (Stephen Morris and Hyun Song Shin 2004) and (Itay Goldstein and Ady Pauzner 2005)), which are less flexible than that in our setup. This difference in flexibility differentiates our model from the global games literature.

The equilibria of games with incomplete information are known to be highly sensitive to the belief environment ((Ariel Rubinstein 1989), (Atsushi Kajii and Stephen Morris 1997), and (Jonathan Weinstein and Muhamet Yildiz 2007), etc.). Here, the focus shifts from the belief environment to information acquisition, which determines the beliefs in equilibrium. So rather than analyze the effects of exogenously assumed beliefs, our model studies how the equilibria depend on the properties of information acquisition.

Another related strand of work makes information acquisition endogenous but rigid, where players can purchase Gaussian additive signals before selecting actions (e.g., (Christian Hellwig and Laura Veldkamp 2009), (David P. Myatt and Chris Wallace 2012), (Michal Szkup and Isabel Trevino 2012), and others). Both (Hellwig and Veldkamp 2009) and our model show that if players' actions are strategic complements, then so are their information choices. However, our research provides a different insight into the creation of approximate common knowledge and multiplicity. Whereas in (Hellwig and Veldkamp 2009), public noise is assumed in the signals and players have the option to access it to create approximate common knowledge, in our approach approximate common knowledge stems from the flexibility of information acquisition.

TABLE 1—PAYOFF MATRIX CONDITIONAL ON FUNDAMENTAL

	invest	*not invest*
invest	θ, θ	$\theta - r, 0$
not invest	$0, \theta - r$	$0, 0$

.

II. The Model

A. The Basic Environment

We define our game as follows. Two players[2] play a coordination game with payoffs shown by Table 1.

Here θ is a random state with support $\Theta \subset \mathbb{R}$; hereafter, we call θ "the fundamental state". The action set of player $i \in \{1, 2\}$ is $A_i = \{0, 1\}$, where 1 stands for "invest" and 0 stands for "not invest". Hence player i's payoff from $a_i \in A_i$ when the state is θ and the other player takes $a_j \in A_j$ is given by

$$u_i\left(a_i, a_j, \theta\right) = a_i \cdot \left[\theta - r \cdot \left(1 - a_j\right)\right].$$

Note that $r > 0$ is the cost of miscoordination. It measures the degree of strategic complementarity. Fundamental θ is drawn from Θ according to a common prior P, which is a probability measure over Θ. We assume that P is absolutely continuous with respect to the Lebesgue measure over $\mathbb{R}$.[3]

This game can be interpreted as a problem of coordination. The two players coordinate in investing in a project with uncertain future cash flow $\theta \in \Theta$. Player $i \in \{1, 2\}$ must decide whether to invest ($a_i = 1$) or not to invest ($a_i = 0$). If both players invest, each gets a payoff θ. If only one player invests, he gets $\theta - r$. The payoff for not investing is normalized to zero, regardless of the other player's action.

To see how this game works, first note that when information is complete and $\theta \in [0, r]$, there exist two strict Nash equilibria: (invest, invest) and (not invest, not invest). This multiplicity stems from the common knowledge of the fundamental.

Our model assumes incomplete information. Suppose player i is endowed with an information structure (S_i, q_i) that conveys information about θ. Here $S_i \subset \mathbb{R}$ is the set of realizations of player i's signal, and $q_i\left(s_i|\theta\right)$ is the conditional probability density function of that signal. We assume conditional independence between s_i and s_j given θ, in order to capture the private nature of the players' information. Given information structure (S_i, q_i), player i's strategy can be represented by a mapping σ_i from S_i to $[0, 1]$,

---

[2]The "two-player" setup is not as restrictive as it seems. All our results remain valid when there is a continuum of players if we redefine the payoff for "invest" as $\theta - r \cdot (1 - m)$, where m is the fraction of players that invest. We discuss the "n-player" case in Subsection VI.C.

[3]Note that public information affects the common prior, and hence its effects on the equilibria can be studied through comparative statics with respect to the common prior, as shown in Section V.

where $\sigma_i(s_i)$ denotes the probability of choosing 1 upon observing $s_i \in S_i$. Player i's expected payoff given (s_i, s_j, θ) becomes

$$\sigma_i(s_i) \cdot [\theta - r \cdot (1 - \sigma_j(s_j))].$$

We can then define

$$
\begin{aligned}
&U_i\left(\left((S_i, q_i), \sigma_i\right), \left((S_j, q_j), \sigma_j\right)\right) \\
(1) \quad &= \int_\theta \int_{s_i} \int_{s_j} \sigma_i(s_i)\left[\theta - r \cdot (1 - \sigma_j(s_j))\right] q_i(s_i|\theta) q_j(s_j|\theta)\, ds_i ds_j dP(\theta),
\end{aligned}
$$

as player i's expected payoff with strategy profile (σ_i, σ_j) under information structure $\left((S_i, q_i), (S_j, q_j)\right)$, $i, j \in \{1, 2\}, i \neq j$.

Now let us consider a larger game with flexible information choices: player i chooses strategy $((S_i, q_i), \sigma_i)$ according to the preference given by (1), $i \in \{1, 2\}$. More precisely, player i acquires information by choosing information structure (S_i, q_i), and then takes (mixed) action σ_i according to the signal s_i generated by (S_i, q_i). The solution concept is Nash equilibrium. Note that the players' beliefs about the fundamental and others' actions can be heterogeneous, owing to the private nature of their signals.

The conditional density $q_i(s_i|\theta)$ describes player i's information acquisition strategy. By choosing different functional forms for $q_i(s_i|\theta)$, player i can make the signal received co-vary with the fundamental in any way desired. Intuitively, if player i's welfare is sensitive to the fluctuation of the fundamental within some range $A \subset \Theta$, he would pay close attention to this event by choosing an information structure to ensure that the signal s_i is closely correlated with $\theta \in A$. In this sense, choosing an information structure can be seen as equivalent to hiring an analyst to write a report from the standpoint of one's own interests.

If information acquisition is cost-free, player i would like to establish a one-to-one mapping between s_i and θ, and thus obtain all the information about the fundamental. This makes our problem trivial, since it reduces to a coordination game with complete information. In practice, however, information acquisition is unlikely to be cost-free, so we examine the more interesting case of costly information acquisition. We associate each conditional density $q_i(s_i|\theta)$ with a cost $\mu \cdot I(q_i)$, where

$$
\begin{aligned}
I(q_i) \quad = \quad & \int_\theta \int_{s_i} q_i(s_i|\theta) \ln q_i(s_i|\theta) \cdot ds_i \cdot dP(\theta) \\
(2) \quad & - \int_{s_i} \int_\theta q_i(s_i|\theta) \cdot dP(\theta) \cdot \ln\left(\int_\theta q_i(s_i|\theta) \cdot dP(\theta)\right) \cdot ds_i
\end{aligned}
$$

is the mutual information between the two random variables s_i and θ. This variable

measures the amount[4] of information about θ conveyed by s_i.[5][6]   The marginal cost of information acquisition is $\mu > 0$.   Since information acquisition is costly, it is not advisable to seek a signal that is highly informative on all values of θ.   Players should make their signals sensitive to the events most relevant to their welfare.

Taking account of information cost, player i's payoff from strategy $((S_i, q_i), \sigma_i)$ is expected future cash flow less the cost of information acquisition, i.e.,

$$V_i \left(((S_i, q_i), \sigma_i), ((S_j, q_j), \sigma_j)\right)$$

$$(3) \qquad = \ U_i \left(((S_i, q_i), \sigma_i), ((S_j, q_j), \sigma_j)\right) - \mu \cdot I(q_i), i \neq j.$$

The game is summarized in the following definition:

DEFINITION 1:   *(**Game with Flexible Information Acquisition** $G(r, \mu)$): Two players with preference (3) play the game by choosing $((S_i, q_i), \sigma_i)$, where $S_i \subset \mathbb{R}$ is the set of realizations of player i's signal, $q_i(s_i|\theta)$ is the conditional probability density of player i's signal, and σ_i is a mapping from S_i to $[0, 1]$ that defines player i's action upon receiving signal $s_i \in S_i$, $i \in \{1, 2\}$. The equilibrium concept is Nash equilibrium.*

In principle, this problem would appear to be hard to deal with, since players' possible strategies belong to a functional space, and even S_1 and S_2, the sets of realizations of the signals, are endogenous.   Fortunately, some patterns emerge from the players' optimal information acquisition behavior, which helps considerably to simplify the problem.

B. Some Simple Facts About The Equilibria

Suppose $((S_i, q_i), \sigma_i)$ is player i's equilibrium strategy.   Let

$$\begin{aligned} S_{i,I} &= \{s_i \in S_i : \sigma_i(s_i) = 1\}, \\ S_{i,N} &= \{s_i \in S_i : \sigma_i(s_i) = 0\}, \end{aligned}$$

and

$$S_{i,ind} = \{s_i \in S_i : \sigma_i(s_i) \in (0, 1)\}.$$

Note that $S_{i,I}$ ($S_{i,N}$) is the set of signal realizations such that player i certainly invests (does not invest).   Player i is indifferent when the signal belongs to $S_{i,ind}$.   Note that

---

[4]Here the unit of $I(q_i)$ is "nat". If "ln" is replaced by "$\log_2$", the unit becomes a "bit".   Knowing the result of a single toss of a coin obtains 1 bit of information.   Since 1 bit equals $\ln 2$ nats, choosing bit or nat as the unit is indifferent for our analysis.

[5]Shannon's mutual information is a natural measure of information about one random variable conveyed by another random variable. Shannon defines information as the reduction of uncertainty, reflected in the difference between the two terms in formula (2). And this is the only "right" way to measure information under some intuitive axioms.   (Thomas M. Cover and Joy A. Thomas 1991) discuss mutual information in detail.   But this specific entropic functional form is not necessary for our qualitative results. What is essential is flexibility in choosing information structures. Subsection IV.C extends our results under more general functional forms.

[6]By definition, the functional form of $q_i(s_i|\theta)$ uniquely determines $I(q_i)$, the amount of information.   Moreover, it also defines the qualitative nature of the information to acquire.   Different forms of $q_i(s_i|\theta)$ may generate the same value for $I(q_i)$; in other words, one can acquire a given amount of information from different aspects of θ.

$\{S_{i,I}, S_{i,N}, S_{i,ind}\}$ partitions S_i. Since the only use of the signal is to make a binary decision, a signal that differentiates more finely among the states only provides redundant information, wasting the players' attention. Hence player i will not discern signal realizations within any $S_{i,I}$, $S_{i,N}$ and $S_{i,ind}$. Moreover, being indifferent between the two actions upon event $S_{i,ind}$, the player would rationally devote no resources (pay no attention) to distinguishing this event from others. Hence, upon receiving their signals the players always play pure strategies and accordingly always prefer binary-signal information structures. (Michael Woodford 2008) makes a similar argument that the agent only needs to acquire a "yes/no" signal. For completeness, we prove the results in our context.

LEMMA 1: *In any equilibrium of game $G(r, \mu)$, $\#S_i = 1$ or 2, and $\Pr(S_{i,ind}) = 0$, $\forall i \in \{1, 2\}$.*

PROOF: See Appendix A.

This lemma follows from the fact that the information cost given by rational inattention conforms to Blackwell's ordering[7] of information structures: namely, more highly informative information structures are more expensive. Hence, to save information costs players only choose binary information structures.

Suppose $((S_i, q_i), \sigma_i)$ is player i's equilibrium strategy. This induces a conditional probability function m_i from Θ to $[0, 1]$, such that player i invests with probability $m_i(\theta)$ when the fundamental is equal to θ. But Lemma 1 implies that m_i also suffices to characterize strategy $((S_i, q_i), \sigma_i)$. That is, we can recover $((S_i, q_i), \sigma_i)$ from m_i. Specifically, in the trivial case

$$m_i(\theta) = 1 \ a.s. \ (\text{or } m_i(\theta) = 0 \ a.s.),$$

let

$$S_i = \{s_{i,I}\}, q_i(s_{i,I}|\theta) = 1 \ a.s., \sigma_i(s_{i,I}) = 1$$
$$(\text{or } S_i = \{s_{i,N}\}, q_i(s_{i,N}|\theta) = 0 \ a.s., \sigma_i(s_{i,N}) = 0);$$

otherwise, let

$$S_i = \{s_{i,I}, s_{i,N}\},$$
$$\forall \theta \in \Theta, q_i(s_{i,I}|\theta) = m_i(\theta), q_i(s_{i,N}|\theta) = 1 - m_i(\theta),$$
$$\sigma_i(s_{i,I}) = 1, \sigma_i(s_{i,N}) = 0.$$

Hence, conditional probability function m_i characterizes player i's strategy $((S_i, q_i), \sigma_i)$. For our equilibrium analysis, we can focus on the strategy profile (m_1, m_2).

With a slight abuse of notation but without confusion, we can rewrite player i's expected payoff as

(4) $\quad U_i(m_i, m_j) = \int m_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta), i, j \in \{1, 2\} \text{ and } i \neq j.$

[7]See (David Blackwell 1953).

This expression is derived from (1).

Lemma 1 also implies that $I(q_i)$, the amount of information acquired, is a function of m_i, so hereafter we use $I(m_i)$ instead of $I(q_i)$. By (2) and Lemma 1, $I(m_i)$ is expressed as

$$
\begin{aligned}
I(m_i) \quad = \quad & \int \left[m_i(\theta) \ln m_i(\theta) + (1 - m_i(\theta)) \ln (1 - m_i(\theta)) \right] dP(\theta) \\
& - p_{Ii} \ln p_{Ii} - (1 - p_{Ii}) \ln (1 - p_{Ii}),
\end{aligned}
$$

(5)

where

$$
p_{Ii} = \Pr(a_i = 1) = \int m_i(\theta) \, dP(\theta)
$$

is player i's unconditional probability of investing. Mutual information $I(m)$ measures function m's variability, which reflects the informativeness of actions about the fundamental. For example, when $m(\theta)$ is constant, the actions convey no information about θ and the corresponding mutual information is nil. This is because the integrand in the first term of (5) is strictly convex, so $I(m)$ is zero if and only if $m(\theta)$ is constant. A nice property of our technology of information acquisition, therefore, is that information acquisition exists if and only if $m(\theta)$ varies over θ, if and only if information cost is strictly positive. Also note that the functional form ("shape") of m determines not only the quantity but also the qualitative nature of the information. For instance, a player can concentrate his attention on some event by making $m(\theta)$ highly sensitive to θ if the event occurs. In this sense, our technology of information acquisition is flexible, inasmuch players can decide both the quantity and the qualitative nature of their information by unrestricted choice of m.[8]

Taking information cost into account, player i's overall expected payoff (in terms of m_i, m_j) is

(6) $V_i(m_i, m_j) = U_i(m_i, m_j) - \mu \cdot I(m_i), i, j \in \{1, 2\}$ and $i \neq j$.

For simplicity, in the rest of the paper we abstract from the story of information acquisition and treat the problem as a two-player game with preference (6) and strategy profile (m_1, m_2). We assume that each player's strategy space is $L^1(\Theta, P)$, i.e., the space of all P-integrable functions on Θ equipped with the norm

$$
\|m_1 - m_2\|_{L^1(\Theta, P)} = \int_{\Theta} |m_1(\theta) - m_2(\theta)| \, dP(\theta) \ .
$$

---

[8]This idea is discussed more amply in the remarks of Proposition 1 in Appendix A, with an example illustrating the key feature of flexible information acquisition.

III.   The Equilibria

A Nash equilibrium of game $G(r, \mu)$ is a strategy profile (m_1, m_2) that solves the following problem:

$$m_i \in \arg \max_{\widetilde{m}_i \in L^1(\Theta, P)} V_i \left(\widetilde{m}_i, m_j \right) = U_i \left(\widetilde{m}_i, m_j \right) - \mu \cdot I \left(\widetilde{m}_i \right)$$

$$\text{s.t.} \quad \widetilde{m}_i (\theta) \in [0, 1], \forall \theta \in \Theta,$$

where $i, j \in \{1, 2\}$ and $i \neq j$.

In order to solve this game, we look first at the general binary-action decision problem with flexible information acquisition.

A.   The Decision Problem

Consider a decision maker who has to choose an action $a \in \{0, 1\}$ and will receive a payoff $u(a, \theta)$, where $\theta \in \Theta$ and P are defined as above. Before deciding, the decision maker can acquire information in a manner specified in Section II. According to Lemma 1, we can represent this player's strategy as a function

$$m(\theta) = \Pr(\text{taking action 1} \mid \text{fundamental equals } \theta) .$$

Let

$$\Delta u(\theta) = u(1, \theta) - u(0, \theta)$$

be the payoff gain from action 1 over action 0. This payoff gain determines the incentive to acquire information by shaping strategy m, as the following proposition demonstrates.

PROPOSITION 1:   *Let* $\Pr(\Delta u(\theta) \neq 0) > 0$ *to exclude the trivial case of the decision maker always being indifferent between the two actions. Let*

$$m \in \Omega \triangleq \left\{ m \in L^1(\Theta, P) : \forall \theta \in \Theta, m(\theta) \in [0, 1] \right\}$$

be an optimal strategy and

$$p_1 = \int_\Theta m(\theta) \, dP(\theta)$$

be the corresponding unconditional probability of taking action 1. Then,
i) the optimal strategy is unique;
ii) there are three possibilities for the optimal strategy:
        a) $p_1 = 1$ *(i.e., $m(\theta) = 1$ for all $\theta \in \Theta$) if and only if*

(7)
$$\int_\Theta \exp\left(-\mu^{-1} \Delta u(\theta)\right) dP(\theta) \leq 1 ;$$

b) $p_1 = 0$ (i.e., $m(\theta) = 0$ for all $\theta \in \Theta$) if and only if

$$(8) \qquad \int_\Theta \exp\left(\mu^{-1} \Delta u(\theta)\right) dP(\theta) \leq 1 \, ;$$

c) $p_1 \in (0, 1)$ if and only if

$$(9) \qquad \int_\Theta \exp\left(\mu^{-1} \Delta u(\theta)\right) dP(\theta) > 1 \text{ and } \int_\Theta \exp\left(-\mu^{-1} \Delta u(\theta)\right) dP(\theta) > 1 \, ;$$

in this case, the optimal strategy m is characterized by

$$(10) \qquad \Delta u(\theta) = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1 - m(\theta)}\right) - \ln\left(\frac{p_1}{1 - p_1}\right) \right]$$

for all $\theta \in \Theta$.

PROOF: See Appendix A.

This proposition, restated here for completeness, is the same as Proposition 1 of (Ming Yang 2012) and Lemma 2 of (Woodford 2008). To maintain focus on the game itself, the proof and detailed discussion are relegated to Appendix A. The equilibrium analysis in the rest of the paper is based on this proposition.

B. Characterizing the Equilibria

Given player j's strategy m_j, player i's payoff gain from investing rather than not investing is

$$\Delta u_i(\theta) = \theta - r \cdot \left[1 - m_j(\theta)\right] \, .$$

As is shown in Proposition 1, this payoff gain function determines player i's incentive to acquire information. There are two motives for acquiring information: to reduce uncertainty about the fundamental and to coordinate the investment decision with the other player's by coordinating information acquisition. Given this second motive, player i should pay attention to the events to which player j pays attention. But if player j never acquires information (e.g., always invests), this second motive for player i is lacking. Moreover, if the fundamental is very likely to be positive ex ante, the first motive does not hold either. Then player i may find it optimal always to invest and never acquire information. This confirms player j's no information-acquisition strategy and so constitutes an equilibrium. Because such equilibria are trivial, we exclude them by the following assumption.

Assumption: $\int_\Theta \exp\left(-\mu^{-1}\theta\right) dP(\theta) > 1$ *and* $\int_\Theta \exp\left(\mu^{-1}\theta\right) dP(\theta) > e^{\mu^{-1} r}$.[9]

---

[9] In Proposition 9 and 10 of Appendix B, we prove that there exists a pooling equilibrium "$m_i(\theta) = m_j(\theta) = 1$ almost everywhere" ("$m_i(\theta) = m_j(\theta) = 0$ almost everywhere") if and only if $\int_\Theta \exp\left(-\mu^{-1}\theta\right) dP(\theta) > 1$ ($\int_\Theta \exp\left(\mu^{-1}\theta\right) dP(\theta) > e^{\mu^{-1} r}$). Hence, this assumption excludes all equilibria with no information acquisition.

This assumption is analogous to the "limit dominance" assumption of the global games approach.[10] The basic intuition here is that common prior P should not be concentrated within the interval $[0, r]$, insofar as if it is, the players are both confident of event $\theta \in [0, r]$. Once a player always invests (or does not invest), the other's payoff gain is very likely to be positive (negative) ex ante, eliminating the incentive to acquire information. Note that if the common prior is $N\left(t, \sigma^2\right)$, this assumption is equivalent to

$$\sigma^2 > r \cdot \mu \text{ and } t \in \left(r - \mu^{-1}\sigma^2/2, \mu^{-1}\sigma^2/2\right) ;$$

or if the common prior is a uniform distribution over interval $[-A, r + A]$, then the assumption holds when A is large enough. All the following results are derived under this assumption, unless otherwise noted.

PROPOSITION 2: *In equilibrium, player i's strategy is characterized by*

$$(11) \quad \theta - r \cdot \left[1 - m_j\left(\theta\right)\right] = \mu \cdot \left[\ln\left(\frac{m_i\left(\theta\right)}{1 - m_i\left(\theta\right)}\right) - \ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right)\right] \text{ for all } \theta \in \Theta,$$

where

$$(12) \qquad\qquad p_{Ii} = \int m_i\left(\theta\right) dP\left(\theta\right) \in (0, 1)$$

is player i's unconditional probability of investing and $i, j \in \{1, 2\}, i \neq j$.

PROOF: Note that

$$\Delta u_i\left(\theta\right) = \theta - r \cdot \left[1 - m_j\left(\theta\right)\right]$$

and

$$\int \exp\left(\mu^{-1}\Delta u_i\left(\theta\right)\right) dP\left(\theta\right)$$
$$= \int \exp\left(\mu^{-1}\left(\theta - r \cdot \left[1 - m_j\left(\theta\right)\right]\right)\right) dP\left(\theta\right)$$
$$\geq \int \exp\left(\mu^{-1}\left(\theta - r\right)\right) dP\left(\theta\right)$$
$$> 1 ,$$

where the last inequality follows from the assumption $\int \exp\left(\mu^{-1}\theta\right) dP\left(\theta\right) > e^{\mu^{-1}r}$.

---

[10]See the survey of global game models by (Stephen Morris and Hyun Song Shin 2001).

Note further that

$$\int \exp\left(-\mu^{-1}\Delta u_i\left(\theta\right)\right) dP\left(\theta\right)$$

$$= \int \exp\left(-\mu^{-1}\left(\theta - r \cdot \left[1 - m_j\left(\theta\right)\right]\right)\right) dP\left(\theta\right)$$

$$\geq \int \exp\left(-\mu^{-1}\theta\right) dP\left(\theta\right)$$

$$> 1,$$

where the last inequality is simply the assumption. Therefore, (11) and (12) are direct implications of Proposition 1 (ii) (c).

This result is a direct implication of Proposition 1. It states that a pair $(m_1, m_2) \in \Omega \times \Omega$ constitutes an equilibrium if and only if it satisfies (11) and (12). Given player j's strategy $m_j\left(\theta\right)$, the left-hand side of (11) is player i's marginal benefit from increasing the conditional probability of "invest". Since $\mu > 0$ is the marginal cost of acquiring an extra bit of information and

$$\left[\ln\left(\frac{m_i\left(\theta\right)}{1 - m_i\left(\theta\right)}\right) - \ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right)\right]$$

is the "derivative" of the amount of information with respect to $m_i\left(\theta\right)$, the right-hand side of (11) is player i's marginal cost of increasing $m_i\left(\theta\right)$. Hence (11) states that the marginal cost must equal the marginal benefit. Also note that

$$\ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right)$$

is player i's average odds of investing (ratio of "invest" to "not invest"), while

$$\ln\left(\frac{m_i\left(\theta\right)}{1 - m_i\left(\theta\right)}\right)$$

is the odds conditional on θ. Therefore (11) indicates that player j's strategy m_j shapes player i's marginal benefit

$$\theta - r \cdot \left[1 - m_j\left(\theta\right)\right],$$

which in turn determines the deviation from player i's average odds of investing.

Since the payoff matrix (Table 1) is symmetric and the players' actions are strategic complements, it is natural to expect symmetric equilibria.

PROPOSITION 3: *All the equilibria in game $G\left(r, \mu\right)$ are symmetric, i.e., $m_1\left(\theta\right) = m_2\left(\theta\right)$ almost everywhere.*

PROOF: See Appendix B.

The strategic complementarity between the players' actions creates the motive for coordination in the acquisition of information. Given the private nature of their information acquisition, coordination is possible only if they choose the same information structure (i.e., the same $m(\cdot)$). By this proposition, from now on we can use a single function m to represent the equilibrium.

COROLLARY 1: (m_1, m_2) *is an equilibrium of game* $G(r, \mu)$ *if and only if there exists an* $m \in \Omega$, *such that* $m_1(\theta) = m_2(\theta) = m(\theta)$ *and*

$$(13) \qquad \theta - r \cdot [1 - m(\theta)] = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1 - m(\theta)} \right) - \ln\left(\frac{p_I}{1 - p_I} \right) \right]$$

for all $\theta \in \Theta$, *where*

$$p_I = \int m(\theta) \cdot dP(\theta).$$

PROOF: This corollary follows directly from Proposition 3.

This corollary is an immediate implication of Propositions 2 and 3. In fact, equation (13) summarizes all the previous derivations and is sufficient, as well as necessary, for all the equilibria. The equilibrium analysis of the next two sections is conducted through this equation.

It is easy to verify that the graph

$$\left\{ (\theta, m) \,|\, \theta - r \cdot (1 - m) = \mu \cdot \left[\ln\left(\frac{m}{1 - m} \right) - \ln\left(\frac{p_I}{1 - p_I} \right) \right] \right\}$$

is central-symmetric in the $\theta \sim m$ plane about the point $(\theta_0, 1/2)$, where

$$(14) \qquad \theta_0 = r/2 - \mu \cdot \ln\left(\frac{p_I}{1 - p_I} \right).$$

Combining (13) and (14) gives

$$\theta - \theta_0 = \mu \cdot \ln\left(\frac{m(\theta)}{1 - m(\theta)} \right) + r \cdot \left(\frac{1}{2} - m(\theta) \right).$$

Hence any solution to (13) has an expression $m(\theta - \theta_0)$ and can be indexed by θ_0, i.e.,

$$(15) \qquad \theta - \theta_0 = \mu \cdot \ln\left(\frac{m(\theta - \theta_0)}{1 - m(\theta - \theta_0)} \right) + r \cdot \left(\frac{1}{2} - m(\theta - \theta_0) \right).$$

In other words, any solution is a translation of function $m(\theta)$, which is implicitly defined by

$$\theta = \mu \cdot \ln\left(\frac{m(\theta)}{1 - m(\theta)} \right) + r \cdot \left(\frac{1}{2} - m(\theta) \right).$$

As a result, a solution to (13) is jointly determined by its position θ_0 and its "shape"
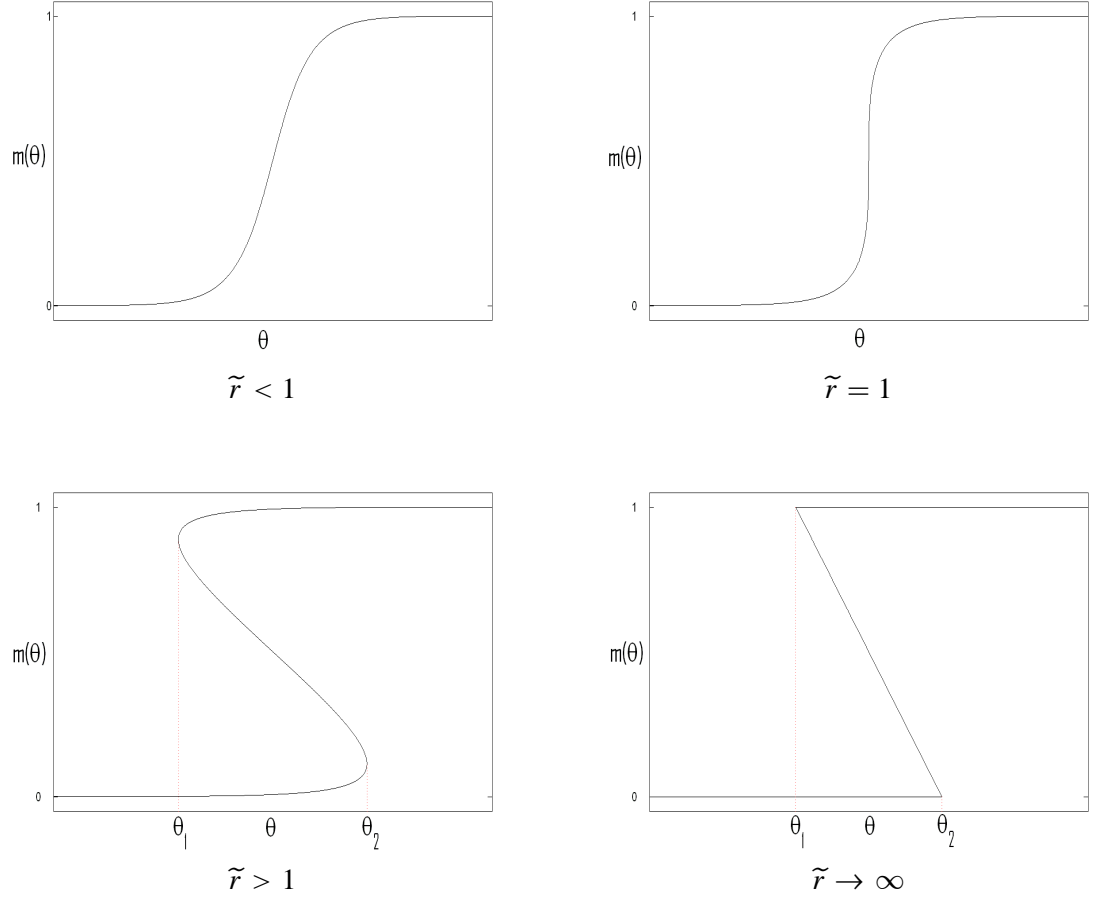
FIGURE 1. EVOLUTION OF THE SHAPE OF EQUILIBRIUM

$m(\theta)$. Note, however, that not every $\theta_0 \in \mathbb{R}$ is sufficient to make $m(\theta - \theta_0)$ a solution. The position θ_0 is endogenously determined in equilibrium.

We first analyze the "shape" of the equilibrium. The "shape" $m(\theta)$ is determined by $\widetilde{r} \triangleq \frac{r}{4 \cdot \mu}$, the ratio of strategic complementarity r to the marginal cost of information acquisition μ. Figure 1 shows how $m(\theta)$ varies as $\widetilde{r}$ increases.

What information is acquired in equilibrium? According to Lemma 1, we can recover the equilibrium information structure from $m(\theta)$. Let $S_i = \{0, 1\}$ be the set of realizations for player i's signal s_i. Player i invests if $s_i = 1$ and does not invest otherwise. This information structure is characterized by conditional probability $\Pr(s_i = 1|\theta) = m(\theta)$. Since the probability of investing is highly sensitive to θ where the slope $\left| \frac{dm(\theta)}{d\theta} \right|$ is steep,

$\left| \frac{dm(\theta)}{d\theta} \right|$ reflects player i's attentiveness around θ[11]. Under this interpretation, Figure 1 reveals that players actively acquire information for intermediate values of the fundamental but are rationally inattentive to tail values. This coincides with our intuition. When θ is too high (low), the players should invest (not invest) anyway. Hence the information about θ at tail values is largely irrelevant to their payoffs. When θ takes intermediate values, each player's gain from investing depends crucially on the value of θ and its implication concerning the other player's action. Therefore, the information about θ in the intermediate region is payoff-relevant and attracts most of their attention.

How does information acquisition affect coordination? First, the equilibrium strategy curve becomes flatter as μ increases. Higher cost of information directly weakens the players' ability to acquire it. Hence players' responses will display more idiosyncratic errors. Moreover, expecting their counterparts to react in a noisier fashion, the players now have less incentive to coordinate. Thus the equilibrium strategy becomes even less decisive.

Second, multiple equilibria might emerge[12] as $\widetilde{r} = \frac{r}{4 \cdot \mu}$ exceeds unity. As the lower-left subgraph of Figure 1 shows, there exist $\theta_1 < \theta_2$ such that multiple values of $m(\theta)$ satisfy (15) for all θ within $[\theta_1, \theta_2]$. Note that while strategic complementarity r measures players' motive for coordination, the cost of coordination is given by μ, the marginal cost of information acquisition, as this is a prerequisite to coordinating investment decisions. Hence, the condition $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$ means that when coordination motive dominates coordination cost, the players have multiple ways of coordinating information acquisition, which leads to approximate common knowledge and thus multiplicity.

Third, the Monotonic Likelihood Ratio Property (MLRP), a frequent assumption in applied models with incomplete information, could be violated by our players when $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$.[13] When coordination motive exceeds coordination cost, a player has both the incentive and the ability to coordinate with the other player's non-MLRP strategy. Consequently, our approach offers a condition for evaluating the fit of MLRP.

Finally, as the lower-right subgraph of Figure 1 shows, when the information cost goes to zero this equilibrium approximates the switching strategy. This result coincides with the equilibria of coordination games with complete information.

When $\widetilde{r} = \frac{r}{4 \cdot \mu} \leq 1$, there is a unique shape of $m(\theta)$ that can satisfy (15). But when $\widetilde{r}$ exceeds unity there are infinitely many shapes that satisfy (15) when. Figure 2 illustrates four benchmark shapes.

Define the set of possible equilibrium shapes as

$$\mathbf{M}(r, \mu) \triangleq \left\{ m \in \Omega : \theta = \mu \cdot \ln \left(\frac{m(\theta)}{1 - m(\theta)} \right) + r \cdot \left(\frac{1}{2} - m(\theta) \right) \right\}.$$

---

[11] $\left| \frac{dm(\theta)}{d\theta} \right| \triangleq \infty$ when $m(\theta)$ is discontinuous at θ.

[12] We prove this multiplicity later.

[13] We say a player's strategy satisfies MLRP if the conditional probability of investing increases in the fundamental; that is, if the information structure is more likely to suggest investing when the fundamental is higher.
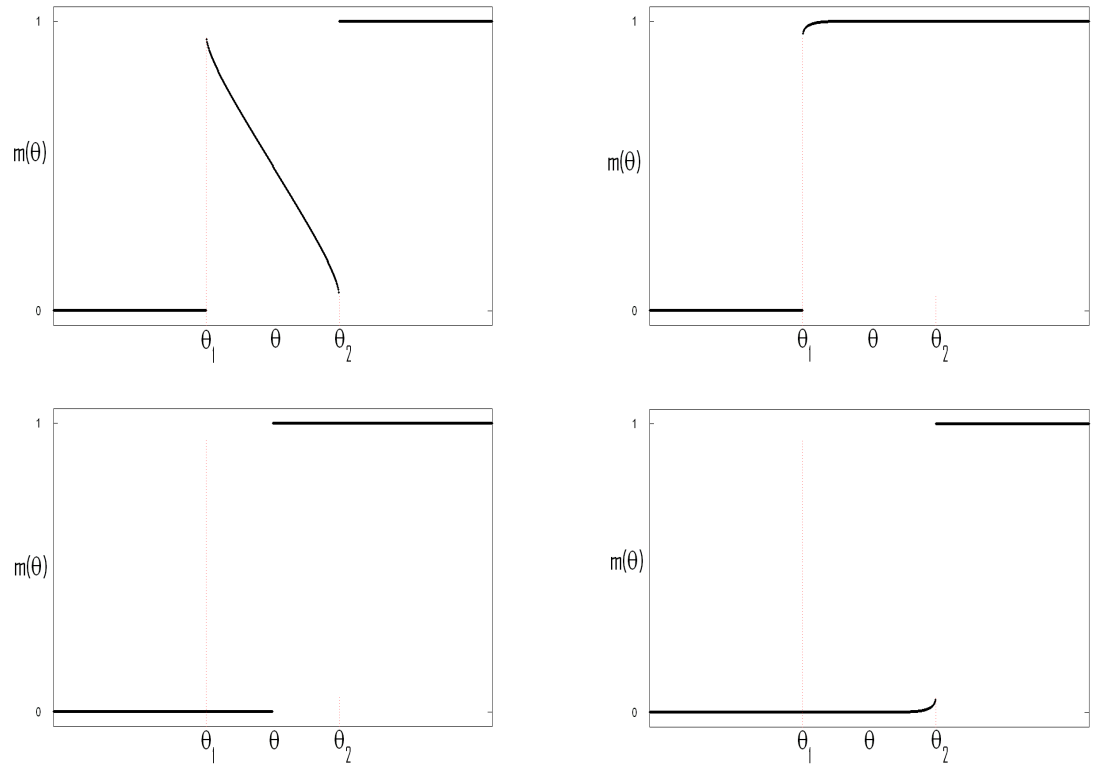
FIGURE 2. BENCHMARK SHAPES OF EQUILIBRIA

Note that

$$\#\mathbf{M}\,(r,\mu) = \begin{cases} 1 & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} \le 1 \\ \infty & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} > 1 \end{cases}.$$

Given r and μ, an equilibrium $m\,(\theta - \theta_0)$ is determined by its shape $m \in \mathbf{M}\,(r,\mu)$ and by its position θ_0. According to (14), the equilibrium condition for θ_0 is

$$\theta_0 = r/2 - \mu \cdot \ln\left(\frac{\int m\,(\theta - \theta_0)\cdot dP\,(\theta)}{1 - \int m\,(\theta - \theta_0)\cdot dP\,(\theta)}\right).$$

Hence, searching for an equilibrium with any given shape $m \in \mathbf{M}\,(r,\mu)$ is equivalent to seeking a fixed point θ_0 in the following mapping:

$$(16) \qquad g\,(\theta_0, m) \triangleq r/2 - \mu \cdot \ln\left(\frac{\int m\,(\theta - \theta_0)\cdot dP\,(\theta)}{1 - \int m\,(\theta - \theta_0)\cdot dP\,(\theta)}\right).$$

Since public information is summarized in common prior P, equation (16) also shows that public information affects the equilibrium only by changing its position θ_0, while leaving its shape unaffected.

As $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$ allows multiple shapes, a natural question, explored in Section IV, is whether this leads to multiple equilibria.

IV.   Private Information Acquisition: Rigidity versus Flexibility

This section sets out our main results.   Multiple equilibria emerge when strategic complementarity dominates information cost.   We contrast this result to the result of an extended global game model, showing why the roles of rigid and flexible information acquisition differ so greatly.   We go on to show how flexible information acquisition allows more efficient coordination.   Finally, we go beyond the entropic information cost to examine the essence of flexibility and the key aspect that drives our results.

LEMMA 2:   *For any possible shape $m \in \mathbf{M}\,(r,\mu)$, there exists $\theta_0 \in \mathbb{R}$ such that $m\,(\theta - \theta_0)$ is an equilibrium.*

PROOF: See Appendix C.
   This lemma demonstrates the existence of the equilibrium and also provides a sufficient condition for multiple equilibria.

PROPOSITION 4:   *If $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$, then game $G\,(r,\mu)$ has infinitely many equilibria.*

PROOF: As shown in Section III, $\#\mathbf{M}\,(r,\mu) = \infty$ when $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$.   Hence, this proposition is a direct implication of Lemma 2.
   This result is consistent with our previous intuition.   Since strategic complementarity exceeds information cost, multiple modes of information acquisition are sustainable in equilibrium.   Flexibility plus relatively low information cost enables players to achieve approximate common knowledge of payoffs, which leads to multiplicity.

In addition to the condition for multiplicity, we also enrich our results in Subsection VI.A from the standpoint of supermodularity. This property enables us to calculate the extreme equilibria and show the emergence of multiplicity as the cost of information varies.

We can now compare flexible and rigid information acquisition.

A.   An Extended Global Game Model

To illustrate the indispensable mechanism of flexibility, it is instructive to contrast the role of information acquisition in our benchmark model from that in an extended global game model, where the players can purchase more accurate signals but cannot change any other aspect of the information structure. Specifically, let two players play the game with payoff matrix (Table 1). The common prior about fundamental θ is P. Player $i \in \{1, 2\}$ takes action $a_i \in \{0, 1\}$ after observing the private signal $x_i = \theta + \beta_i^{-1/2} \cdot \varepsilon_i$, where ε_i is distributed according to a density function f with full support, $\mathbf{E}\varepsilon_i = 0$ and $Var(\varepsilon_i) < \infty$. Here β_i represents the precision of player i's private information. The cost of acquiring information of precision β is $c \cdot h(\beta)$, where $c > 0$ is an exogenous parameter controlling the difficulty of information acquisition and h is continuous and increasing with $h(0) = 0$ and $\lim_{\beta \to \infty} h(\beta) = \infty$. The information structure is rigid in that the additive nature of the signal generating process is not adjustable.

Each player's strategy involves simultaneously choosing a precision $\beta_i \in [0, +\infty)$ and an action rule $s_i : \mathbb{R} \to [0, 1]$, which means that player i chooses 1 with probability $s_i(x_i)$ upon observing x_i. We write $G(c)$ for the game with cost parameter c.

PROPOSITION 5: *Let $\left(\beta_1(c), \beta_2(c)\right)$ be the precision pair chosen in an equilibrium of $G(c)$. Then for any $\beta > 0$, there exists $\bar{c} > 0$, such that for all $c < \bar{c}$, $\beta_i(c) > \beta$, $i \in \{1, 2\}$.*

PROOF: See Appendix C.

This proposition indicates that players would like to acquire information of arbitrarily great precision if the cost is arbitrarily low. A well known result in the global game literature is that uniqueness is guaranteed if private information is sufficiently accurate relative to public information (e.g., Morris and Shin (2004)). Proposition 5 allows us to retrieve the standard global game result in this extended model with information acquisition.

COROLLARY 2: *For any $\delta > 0$, there exists $\bar{c} > 0$, such that for all $c < \bar{c}$, if strategy $s : \mathbb{R} \to \{0, 1\}$ survives the iterated deletion of strictly dominated strategies in game $G(c)$, then $s(x) = 0$ for all $x \leq r/2 - \delta$ and $s(x) = 1$ for all $x \geq r/2 + \delta$.*

PROOF: The proof is a direct application of Proposition 2.2 in (Morris and Shin 2001) together with Proposition 5. By Proposition 2.2 in (Morris and Shin 2001), $\forall \delta > 0$, $\exists \bar{\beta} > 0$, such that the above statement holds for all $\beta > \bar{\beta}$. Then Proposition 5 shows the existence of $\bar{c} > 0$ such that the players acquire information of precision at least $\bar{\beta}$.

According to Corollary 2, when to cost goes to zero all equilibria become approximately the unique switching strategy

$$s\left(x\right) = \begin{cases} 0 & \text{if } x \leq r/2 \\ 1 & \text{if } x > r/2 \end{cases}$$

This is consistent with the standard global game arguments. That is, lowering informational cost induces more accurate private signals, which undermines common knowledge and so facilitates the uniqueness. But by Proposition 4, our model with flexible information acquisition generates the opposite prediction: lowering the cost of information enhances approximate common knowledge and facilitates multiplicity.

How should we understand this radical discrepancy? Strategic complementarity of actions produces the motive for coordination in acquiring information. This motive evolves into actual coordination in our benchmark model with flexibility, especially when the cost of information is lower. Hence, we recreate approximate common knowledge of the payoffs, resulting in multiplicity.

In the global games approach, where private noise is additive to the fundamental, the players are constrained to pay equal attention to all possible values of θ, in the sense that the distribution of the observational error $\beta^{-1/2} \cdot \varepsilon$ is invariant with respect to θ. As a result, they coordinate only in choosing overall precision, while the potential motive for coordinating attention allocation for different levels of θ cannot materialize. This mechanism of rigidity contrasts diametrically with the mechanism at work when information acquisition is flexible.

B. *Welfare Implications: Rigidity versus Flexibility*

Rigid and flexible information acquisition also entail different welfare implications. In the extended global game model, the limit unique equilibrium (when $c \to 0$ and thus $\beta \to \infty$) is inefficient. Both players would have had higher payoffs if they had committed to the most efficient strategy

$$\widetilde{s}\left(x\right) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases} .$$

However, this strategy is not sustainable in equilibrium. A player with a signal just above zero will rationally assign a large probability to the other player's signal being negative, and thus be reluctant to invest for fear of miscoordination. This problem could be overcome if the players could commit to ignoring the exact values of their signals and considering only the signs. For example, if there were a third party who observed the signals and told the players (privately) only the signs of their own signals, each player would find it optimal to invest if and only if the signal were positive. That is, players achieve efficient coordination by discarding information through a commitment device. In a trivial sense, insofar as they map continuous signals to binary actions players always discard some information. But the point is that while, some information is not reflected in players' actions, it is still used in making inferences about others' beliefs

and consequent actions. This is the way in which information matters, which sets it apart from all other economic resources. Therefore, "discarding information" means committing to forget it, i.e., refraining from making inferences based on it.

By contrast, in our benchmark model with flexible information acquisition, $\forall \widehat{\theta} \in [0, r]$,

$$m(\theta) = \left\{ \begin{array}{ll} 0 & \text{if } \theta \leq \widehat{\theta} \\ 1 & \text{if } \theta > \widehat{\theta} \end{array} \right.$$

is an equilibrium when informational cost μ vanishes. Hence the most efficient strategy with cutoff $\widehat{\theta} = 0$ can be supported in equilibrium. Here, flexible and costly information acquisition helps players to acquire only information that is valuable for efficient coordination and to ignore information that is detrimental to it even if its cost goes to zero; they could have a certain welfare gain over the case of rigid information acquisition. Unlike the extended global game model, where players can only discard harmful information by an explicit commitment device, our model with flexibility helps them choose the quantity and the qualitative nature of their information, whereby they act as if they had committed to discarding the information harmful to coordination. In other words, we can interpret this as an implicit commitment device inhabiting the flexibility of information acquisition. It is worth underscoring that this contrast was not discernible in one-person decision problems with information acquisition, and not only because by definition they do not entail coordination but also because where cost is not a consideration more information is always more desirable, no matter whether acquisition is flexible or rigid.

C. Flexibility: General Information Cost

This subsection goes beyond entropic information cost to explore the essence of flexibility and the essential aspect driving our results. Here it needs to be underscored that different forms of information acquisition could be exclusively captured by different schemes of information cost. For example, if the information structure is exogenously endowed, it could be posited that only this endowed structure can be acquired at no cost while all others incur an infinite cost. If information acquisition is endogenous but rigid, as in the extended global game model, only information structures following the form of $x = \theta + \beta^{-1/2} \cdot \varepsilon$ are associated with finite cost. In our benchmark model with rational inattention, any information structure is associated with a cost proportional to the consequent reduction of entropy. All in all, analysis of information cost schemes covers all possible considerations concerning forms of information acquisition. On this basis, we go beyond rational inattention to establish a condition concerning information cost schemes that is sufficient to deliver the same qualitative results as our benchmark model.

DEFINITION 2: *Information acquisition is flexible if information cost strictly respects Blackwell's ordering and satisfies Lipschitz continuity over* Ω.

(Blackwell 1953) introduces a partial ordering to reflect the informativeness of information structures. Structure A is less informative than structure B if A can be obtained

from B by the addition of garbling noise[14]. Here an information cost respects Black-well's ordering if it assigns lower cost to less informative structures. Under such an information cost scheme, both players find it optimal to consider only binary informa-tion structures, precluding all structures with more than two signal realizations, which therefore contain redundant and even harmful information. As a result, we only need to specify an information cost over Ω, the set of binary information structures.

An information cost $c : \Omega \to \mathbb{R}_+$ satisfies Lipschitz continuity if there exists $K > 0$ such that

$$\forall m, \widetilde{m} \in \Omega, \ |c(m) - c(\widetilde{m})| \leq K \cdot \int |m(\theta) - \widetilde{m}(\theta)| \, dP(\theta) \ ,$$

where P is the common prior. The Lipschitz continuity prevents the information cost from varying wildly and hence guarantees the availability of all potentially valuable in-formation structures.

Consider a coordination game with payoffs given by Table 1 and information cost $\mu \cdot c$, where $\mu > 0$ and $c : \Omega \to \mathbb{R}_+$.

PROPOSITION 6: *If information acquisition is flexible, then for any strict Nash equi-librium of the game with complete information (except the two extreme equilibria $m(\theta) = 1_{\{\theta > 0\}}$ and $m(\theta) = 1_{\{\theta > r\}}$), there exists $\overline{\mu} > 0$ such that for all $\mu \in [0, \overline{\mu}]$, this strict Nash equilibrium is still a Nash equilibrium of the game with information cost $\mu \cdot c$.*

PROOF: See Appendix D.

This proposition is analogous to Proposition 4 and bears the same interpretation. It supports our thesis that what drives our result is flexibility, not the entropic functional form. Our benchmark model incorporates entropic information cost not only for its meaningful interpretation in information theory but also because it allows us to obtain a clear condition $r > 4\mu$, which intuitively inspires our thinking on the current problem. The standard global game model and the extended global game model discussed above are two typical setups where information acquisition is not flexible: they violate both of the conditions in our definition of flexibility.

Moreover, it should be observed that it is the second rather than the first condition in our definition of flexibility that is essential. Consider information costs respecting Blackwell's ordering, so that we can focus on Ω, the set of binary information structures. In this case, almost any strategy $m^* \in \Omega$ can be supported in equilibrium by a suitable choice of an information cost that violates the second condition. Specifically, choose a subset $S \subset \Omega$ such that $m^* \in S$ is the unique equilibrium of our coordination game when players can choose freely from S at no cost but have no choice outside S. Define an information cost such that $c(m) = \infty$ for all $m \in \Omega \backslash S$ and $c(m^*) \leq c(m) < \infty$ for all $m \subset S$. Then for all $\mu \in \mathbb{R}_+$, m^* is the unique equilibrium of the coordination game with information cost $\mu \cdot c(m)$.[15] The point is that restricting the players to subset S

---

[14]See (Blackwell 1953) for detailed discussion.
[15]A trivial construction is to let $S = \{m^*\}$, then we have a trivial game with m^* as the unique equilibrium.

physically precludes some information that is potentially valuable. Therefore, the failure of the second condition leads to uniqueness.

V. The Impact of Public Information

In our benchmark model, players acquire private information at a cost. We assume that public information is directly observable and cost-free. That is, public information is common knowledge. It affects players' decisions by changing the common prior about the fundamental. Here we conduct a comparative static analysis with respect to common prior P to study the differing impact of public information. Broadly, if the common prior is concentrated in the intermediate region $[0, r]$, both players are confident that the event $\{\theta \in [0, r]\}$ will happen with high probability. As a result, their coordination motive dominates concern over fluctuations in the fundamental, and multiple equilibria emerge regardless of the information cost.

PROPOSITION 7: *For any $r > 0$ and $\mu > 0$, game $G(r, \mu)$ has multiple equilibria if $\mathbf{E}e^{-\mu^{-1}\theta} \leq 1$ and $\mathbf{E}e^{\mu^{-1}(\theta-r)} \leq 1$, where expectations are formed according to common prior P.*

PROOF: See Appendix D.

Under the condition of this proposition, at least "always invest" (i.e., $m(\theta) = 1$ for all $\theta \in \Theta$) and "never invest" (i.e., $m(\theta) = 0$ for all $\theta \in \Theta$) are both equilibria[16]. The players find it optimal not to acquire information and to perfectly coordinate their investment decisions. We can gain some insight into this situation from a Gaussian common prior $N(t, \sigma^2)$. In this case, it is easy to verify that condition

$$\mathbf{E}e^{-\mu^{-1}\theta} \leq 1 \text{ and } \mathbf{E}e^{\mu^{-1}(\theta-r)} \leq 1$$

is equivalent to

$$\sigma^2 \leq r \cdot \mu \text{ and } t \in \left[\mu^{-1}\sigma^2/2, r - \mu^{-1}\sigma^2/2\right] .$$

That is, the common prior should have low dispersion and its probability peak should be close to $r/2$. Proposition 7 is strong in the sense that the criterion

$$\mathbf{E}e^{-\mu^{-1}\theta} \leq 1 \text{ and } \mathbf{E}e^{\mu^{-1}(\theta-r)} \leq 1$$

is uniform for all common priors. For comparability with the standard global game results, let us establish the following corollary:

COROLLARY 3: *Let $p(\theta)$ be a probability density function. Then, for any $r > 0$, $\mu > 0$ and $y \in (0, r)$, there exists $\overline{\beta} > 0$ such that for all $\beta > \overline{\beta}$, game $G(r, \mu)$ with common prior $\beta^{1/2} p\left(\beta^{1/2}(\theta - y)\right)$ (density function) has multiple equilibria.*

---

[16]We prove this result in Proposition 9 and 10 in Appendix B.

PROOF: First note that

$$\lim_{\beta \to \infty} \beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right) = \delta (\theta - y) \ ,$$

where $\delta (\cdot)$ is the Dirac delta function. Hence

$$
\begin{aligned}
&\lim_{\beta \to \infty} \int_{\Theta} e^{-\mu^{-1}\theta} \cdot \beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right) d\theta \\
&= \int_{\Theta} e^{-\mu^{-1}\theta} \cdot \delta (\theta - y) \, d\theta \\
&= e^{-\mu^{-1}y} \\
&< 1 \ ,
\end{aligned}
$$

where the inequality follows from the condition $y \in (0, r)$. Since

$$\int_{\Theta} e^{-\mu^{-1}\theta} \cdot \beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right) d\theta$$

is continuous in β, there exists $\overline{\beta}_1 > 0$ such that for all $\beta > \overline{\beta}_1$,

$$\int_{\Theta} e^{-\mu^{-1}\theta} \cdot \beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right) d\theta \ < 1 \ .$$

By a symmetric argument, we can find a $\overline{\beta}_2 > 0$ such that for all $\beta > \overline{\beta}_2$,

$$\int_{\Theta} e^{\mu^{-1}(\theta - r)} \cdot \beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right) d\theta \ < 1 \ .$$

Let $\overline{\beta} = \max \left(\overline{\beta}_1, \overline{\beta}_2 \right)$; then according to Proposition 7, the game has multiple equilibria for all $\beta > \overline{\beta}$.

Here β represents the precision of public information. Suppose players have a uniform common prior before any public information. They then observe a public signal

$$y = \theta + \beta^{-1/2} \cdot \varepsilon \ ,$$

where ε is distributed according to a density function p. This public signal results in an updated common prior[17] with density function $\beta^{1/2} p \left(\beta^{1/2} (\theta - y) \right)$. Therefore, Corollary 3 means that providing highly precise public information leads to multiple equilibria. This is consistent with the well known result in the global game literature.

Another well known result is that uniqueness is guaranteed if private signals are sufficiently accurate relative to public signals (e.g., (Morris and Shin 2004)). This means that as regards uniqueness, the effects of greater precision of public signals can be offset

---

[17]We call it a prior since it is formed before players' private information acquisition.

by increasing the precision of private signals. In the framework of our extended global game model, Corollary 2 implies that the effect of more precise public signals can be offset by lowering the cost of acquiring private information. In our benchmark model with flexible information acquisition, however, Proposition 4 states that there are always infinitely many equilibria when $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, regardless of the precision of public information. That is, the effects of public information and private information acquisition are disentangled, because when information cost is low, players have sufficient freedom to coordinate their private information acquisition. This freedom is totally divorced from public information, implying that the entanglements of global game models also depend on the rigidity they implicitly impose on the information structure.

VI. Discussion

First we enrich our results by utilizing the supermodularity of the game. Then we discuss a related game in which players have the capacity to acquire a fixed amount of information at no cost. Finally, we address two extensions of the benchmark model: i) more than two players, $n \geq 2$; and ii) discontinuous payoff gain with respect to the fundamental and to the proportion of players taking a specific action.

A. Supermodularity and Extreme Equilibria

The strategy space Ω endowed with the natural pointwise order "$\geq$" becomes a complete lattice. Specifically, for any $m_1, m_2 \in \Omega$, the order is defined as

$$m_2 \geq m_1 \text{ if and only if } m_2(\theta) \geq m_1(\theta) \text{ for all } \theta .$$

Note that the mutual information $I(\cdot)$ is submodular over strategy space Ω, thus our game is supermodular. The set of equilibria of a supermodular game forms a complete lattice, the supremum and infimum of which correspond to game's greatest and smallest Nash equilibria. See (Donald M. Topkis 1979), (Paul Milgrom and John Roberts 1990) and (Xavier Vives 1990) for more details about supermodular games. Here, we utilize this property to calculate the extreme equilibria of our game. Figure 3 shows how multiple equilibria emerge as the marginal cost of information μ acquisition shrinks.

These graphs show numerical solutions to the game with strategic complementarity $r = 1$ and uniform common prior over the interval $[-0.6, 1.6]$. We examine how the equilibria evolve as the value of the marginal cost of information acquisition μ varies. In the two graphs of the first row, r is greater than 4μ, so we have multiple shapes for the equilibria. It is clear that there are multiple equilibria, because the greatest and smallest equilibria differ. When μ increases from 0.1 to 0.2, the two extreme equilibria get closer, suggesting a tendency to uniqueness. In the southwest graph, r equals 4μ, so we have a unique shape. We still have multiple equilibria (same shape, different positions that satisfy (16)), although it is hard to distinguish them in the graph. In the southeast graph, r is smaller than 4μ and we have a unique equilibrium.
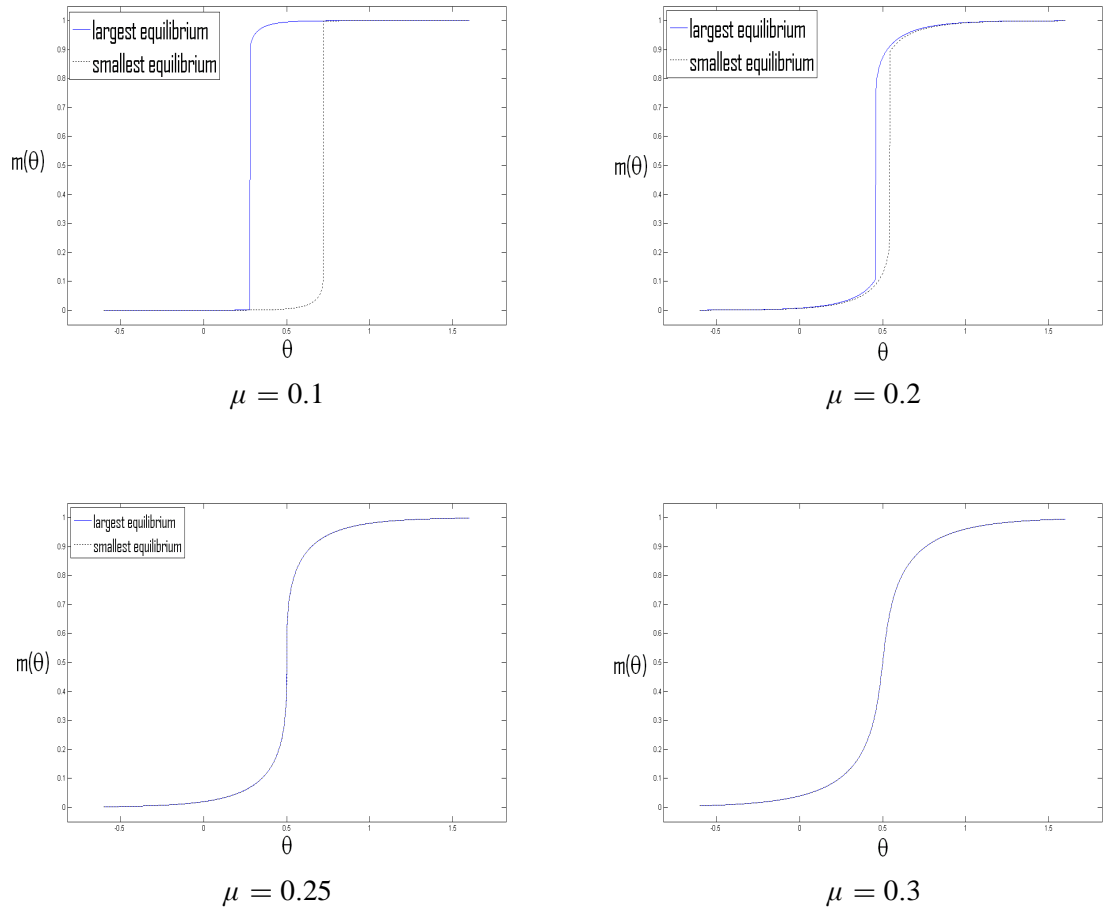
FIGURE 3. EXTREME EQUILIBRIA

B.   Constrained Information Acquisition Game

In the conventional rational inattention models, decision makers are capacity constrained. That is, they can acquire information at no cost but only up to a given amount. This case is discussed here.   Let $\kappa > 0$ denote the maximum amount of information players can acquire. By an argument similar to that of Lemma 1, player i's equilibrium strategy is characterized by a function $m_i \in L^1(\Theta, P)$, and an equilibrium is a pair (m_1, m_2) that solves the the following problem:

$$m_i \in \arg \max_{\widetilde{m}_i \in L^1(\Theta, P)} U_i\left(\widetilde{m}_i, m_j\right)$$

$$\text{s.t. } I\left(\widetilde{m}_i\right) \leq \kappa,$$

where $i, j \in \{1, 2\}, i \neq j$.

Since payoffs are symmetric and players have the same information acquisition capacity, all equilibria are symmetric.[18]   When solving for the equilibrium, the capacity constraint multiplier plays a role analogous to that of μ, the marginal cost of information acquisition in our benchmark model, but with two differences.   First, the multiplier is endogenous, its value varying in different equilibria. This complicates comparative static analysis in this setup. Second, switching strategies could be supported in equilibrium when κ is large enough (e.g., $\kappa > \ln 2$ nats). Since a binary decision problem requires at most $\ln 2$ nats of information, for large κ the capacity constraint does not bind. This case corresponds to our benchmark model with zero marginal information cost and thus has multiple equilibria.

C.   Multiple Players

So far we have considered a 2-player game, but the arguments also hold for games with more players.   Suppose there are $n \geq 2$ players.   Let $N \triangleq \{1, 2, \cdots, n\}$ denote the set of players. All the other assumptions remain the same, except that player $i \in N$ enjoys a payoff

$$\theta - r \cdot \left(1 - \frac{n'}{n-1}\right)$$

from choosing "invest" if the fundamental is θ and n' other players also choose "invest". Obviously, Lemma 1 is still applicable, so player i's strategy is again characterized by

$$m_i(\theta) \triangleq \Pr(\text{player } i \text{ invest} \mid \text{fundamental} = \theta).$$

Given fundamental θ, player i's expected payoff from investing is

$$(17) \qquad\qquad \theta - r \cdot \left(1 - \frac{\sum_{j \neq i} m_i(\theta)}{n-1}\right).$$

---

[18]The argument is analogous to that of Proposition 3 and is omitted here.

An equilibrium of this n-player flexible information acquisition game is an *n*-tuple $(m_1, \cdots m_n)$ that solves the the following problem:

$$m_i \quad \in \quad \arg \max_{\widetilde{m}_i \in L^1(\Theta, P)} V_i(\widetilde{m}_i, m_{-i}) = U_i(\widetilde{m}_i, m_{-i}) - \mu \cdot I(\widetilde{m}_i)$$

$$\text{s.t.} \quad \widetilde{m}_i(\theta) \quad \in \quad [0, 1] \text{ for all } \theta \in \Theta.$$

As in Proposition 3, we can show that all equilibria are symmetric. Hence any equilibrium can be represented by a single function m, and all the remaining arguments in the benchmark model still work.

D. Discontinuous Payoff Gain Function

In our model, the payoff gain from investing over not investing is continuous with respect both to the fundamental and to the opponent's probability of investing. In many important applications of global game theory, however, the payoff gain is discontinuous (e.g., (Morris and Shin 1998)). This subsection presents an example to illustrate this case.

Following the notation in (Morris and Shin 2001), let $\pi(m, \theta)$ denote the payoff gain for choosing "invest" when the fundamental is θ and the opponent chooses "invest" with probability m.[19] When $\pi(m, \theta)$ is discontinuous with respect to m and θ, $\frac{\partial \pi}{\partial m} = \infty$ for some (m, θ), which implies infinite strategic complementarity. The insight developed in our benchmark model suggests multiple equilibria no matter how large μ is, as is shown in the following example. Our basic scenario is the currency attack model of (Morris and Shin 1998).

There is a continuum of players in a flexible information acquisition game. Their payoff gain from choosing "attack" is defined as

$$\pi(m, \theta) = \begin{cases} -1 & \text{if } m < \theta \\ 1 & \text{if } m \geq \theta \end{cases}.$$

This payoff gain can be interpreted as follows. When the currency (the fundamental) is weak and too many speculators are attacking, the government has to abandon the currency peg and each attacker gains one dollar. If the currency is strong, the speculative attack fails and each attacker loses a dollar. In the case of non-attack, the speculator receives zero.

For simplicity, we assume a uniform prior over $[-A, 1 + A]$, where $A \geq 0$. As before, $\mu > 0$ is the marginal cost of information acquisition. We focus on the symmetric equilibria, which can be characterized by a mapping $m : [-A, 1 + A] \to [0, 1]$. Since we are interested in equilibria with information acquisition, we assume

(18)
$$\frac{A}{1 + 2A} e^{\mu^{-1}} + \frac{1 + A}{1 + 2A} e^{-\mu^{-1}} > 1.$$

---

[19] For example, in our benchmark model $\pi(m, \theta) = \theta - r \cdot (1 - m)$, and $\frac{\partial \pi}{\partial m} = r$ represents the strategic complementarity.

Then by Proposition 1, an equilibrium is characterized by

$$(19) \qquad \forall \theta \in [-A, 1+A], \, \pi \, (m, \theta) = \mu \cdot \left[\ln \left(\frac{m \, (\theta)}{1 - m \, (\theta)} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right],$$

where $p_I = \frac{1}{2 \cdot A + 1} \cdot \int_{-A}^{1+A} m \, (\theta) \cdot d\theta$.

Since $\pi \, (m, \theta)$ can take only two values, then according to (19) the same applies to $m \, (\theta)$. Hence the equilibrium strategy can be represented by two numbers $\underline{m}, \overline{m} \in [0, 1]$. Let $S_I \triangleq \{\theta \in [-A, 1+A] : m \, (\theta) = \overline{m}\}$ and $S_N \triangleq \{\theta \in [-A, 1+A] : m \, (\theta) = \underline{m}\}$ denote the region of "attack" with "high" and "low" probability, respectively. By definition, we have $S_I \subset [-A, \overline{m}]$ and $S_N \subset (\underline{m}, 1+A]$. A symmetric equilibrium is therefore characterized by $\underline{m}, \overline{m}, S_I \subset [-A, \overline{m}]$ and $S_N \subset (\underline{m}, 1+A]$ such that

$$(20) \qquad 1 = \mu \cdot \left[\ln \left(\frac{\overline{m}}{1 - \overline{m}} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right],$$

$$(21) \qquad -1 = \mu \cdot \left[\ln \left(\frac{\underline{m}}{1 - \underline{m}} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right],$$

and

$$(22) \qquad p_I = \frac{1}{2 \cdot A + 1} \cdot \left[\Pr (S_I) \cdot \overline{m} + \Pr (S_N) \cdot \underline{m} \right].$$

PROPOSITION 8:   *This game with discontinuous payoff gain has infinitely many equilibria for all $\mu > 0$.*

PROOF: See Appendix D.

This proposition confirms the basic insight of our benchmark model. Discontinuous payoff gain generates infinite strategic complementarity, and as a consequence the coordination motive always dominates information cost, so that an infinite number of ways of coordinating information acquisition can be supported in equilibrium.

<div align="center">

REFERENCES

</div>

Blackwell, David. 1953. "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics*, 24(2): pp. 265–272.

Carlsson, Hans, and Eric van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica*, 61(5): 989–1018.

Cover, Thomas M., and Joy A. Thomas. 1991. *Elements of information theory.* New York, NY, USA:Wiley-Interscience.

Frankel, David M., Stephen Morris, and Ady Pauzner. 2003. "Equilibrium selection in global games with strategic complementarities." *Journal of Economic Theory*, 108(1): 1 – 44.

Goldstein, Itay, and Ady Pauzner. 2005. "Demand-Deposit Contracts and the Probability of Bank Runs." *The Journal of Finance*, 60(3): pp. 1293–1327.

Hellwig, Christian, and Laura Veldkamp. 2009. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies*, 76(1): 223–251.

Kajii, Atsushi, and Stephen Morris. 1997. "The Robustness of Equilibria to Incomplete Information." *Econometrica*, 65(6): pp. 1283–1309.

Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3): 769–803.

Milgrom, Paul, and John Roberts. 1990. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica*, 58(6): 1255–1277.

Monderer, Dov, and Dov Samet. 1989. "Approximating Common Knowledge with Common Beliefs." *Games and Economic Behavior*, 1(2): 170 – 190.

Morris, Stephen, and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *The American Economic Review*, 88(3): pp. 587–597.

Morris, Stephen, and Hyun Song Shin. 2001. "Global Games: Theory and Applications." Cowles Foundation for Research in Economics, Yale University Cowles Foundation Discussion Papers 1275R.

Morris, Stephen, and Hyun Song Shin. 2004. "Coordination risk and the price of debt." *European Economic Review*, 48(1): 133 – 153.

Myatt, David P., and Chris Wallace. 2012. "Endogenous Information Acquisition in Coordination Games." *The Review of Economic Studies*, 79(1): 340 – 374.

Rubinstein, Ariel. 1989. "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'." *American Economic Review*, 79(3): 385–91.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of Monetary Economics*, 50(3): 665 – 690. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

Szkup, Michal, and Isabel Trevino. 2012. "Costly Information Acquisition and Transparency in Global Games." New York University Unpublished working paper.

Topkis, Donald M. 1979. "Equilibrium Points in Nonzero-Sum n-Person Submodular Games." *SIAM Journal on Control and Optimization*, 17(6): 773–787.

Vives, Xavier. 1990. "Nash equilibrium with strategic complementarities." *Journal of Mathematical Economics*, 19(3): 305–321.

Weinstein, Jonathan, and Muhamet Yildiz. 2007. "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements." *Econometrica*, 75(2): 365–400.

Woodford, Michael. 2008. "Inattention as a source of randomized discrete adjustment." Columbia University Unpublished working paper.

Woodford, Michael. 2009. "Information-constrained state-dependent pricing." *Journal of Monetary Economics*, 56(Supplement 1): S100 – S124. Supplement issue: December 12-13, 2008 Research Conference on 'Monetary Policy under Imperfect Informa-

Yang, Ming. 2012. "Optimality of Debt under Flexible Information Acquisition." Duke
University Working Paper 2012.

APPENDIX A

Proof of Lemma 1.

PROOF: Suppose $((S_i, q_i), \sigma_i)$ is player i's equilibrium strategy. Construct a new strategy $((\widetilde{S}_i, \widetilde{q}_i), \widetilde{\sigma}_i)$ with $\widetilde{S}_i = \{s_{i,I}, s_{i,N}\}$ such that

$$\forall \theta \quad \in \quad \Theta,$$
$$\widetilde{q}_i(s_{i,I}|\theta) \quad = \quad \int_{S_{i,I} \cup S_{i,ind}} q_i(s_i|\theta)\, ds_i,$$
$$\widetilde{q}_i(s_{i,N}|\theta) \quad = \quad \int_{S_{i,N}} q_i(s_i|\theta)\, ds_i,$$

and

$$\widetilde{\sigma}_i(s_{i,I}) \quad = \quad 1,$$
$$\widetilde{\sigma}_i(s_{i,N}) \quad = \quad 0.$$

It is straightforward that

$$U_i\left(((S_i, q_i), \sigma_i), ((S_j, q_j), \sigma_j)\right) = U_i\left(((\widetilde{S}_i, \widetilde{q}_i), \widetilde{\sigma}_i), ((S_j, q_j), \sigma_j)\right).$$

However, if $\#(S_i) > 2$, $(\widetilde{S}_i, \widetilde{q}_i)$ is a strict garbling of (S_i, q_i) in the sense of Blackwell's ordering and thus incurs strictly less information cost since mutual information strictly respects Blackwell's ordering. Then

$$v_i\left(((S_i, q_i), \sigma_i), ((S_j, q_j), \sigma_j)\right) - v_i\left(((\widetilde{S}_i, \widetilde{q}_i), \widetilde{\sigma}_i), ((S_j, q_j), \sigma_j)\right)$$
$$= \quad U_i\left(((S_i, q_i), \sigma_i), ((S_j, q_j), \sigma_j)\right) - \mu \cdot I(q_i)$$
$$\quad -U_i\left(((\widetilde{S}_i, \widetilde{q}_i), \widetilde{\sigma}_i), ((S_j, q_j), \sigma_j)\right) + \mu \cdot I(\widetilde{q}_i)$$
$$= \quad \mu \cdot \left[I(\widetilde{q}_i) - I(q_i)\right] < 0,$$

i.e., $((S_i, q_i), \sigma_i)$ is suboptimal and cannot be an equilibrium strategy.

Proof of Proposition 1.

PROOF: Suppose m is an optimal strategy. Let ε be any *feasible* perturbation function. The payoff from the perturbed strategy $m + \alpha \cdot \varepsilon$ is

$$V^*(m + \alpha \cdot \varepsilon)$$
$$= \quad \mathbb{E}\left[(m(\theta) + \alpha \cdot \varepsilon(\theta)) \cdot \Delta u(\theta)\right]$$
$$\quad -\mu \cdot \left[\mathbb{E}g(m(\theta) + \alpha \cdot \varepsilon(\theta)) - g(\mathbb{E}[m(\theta) + \alpha \cdot \varepsilon(\theta)])\right],$$

where the expectation operator $\mathbb{E}(\cdot)$ is with respect to θ under common prior P. $\alpha \in \mathbb{R}$, and ε is feasible with respect to m if $\exists \alpha > 0$, s.t. $\forall \theta \in \Theta$, $m(\theta) + \alpha \cdot \varepsilon(\theta) \in [0, 1]$. Then the first order variation is

$$\left. \frac{dV^*(m + \alpha \cdot \varepsilon)}{d\alpha} \right|_{\alpha=0} = \mathbb{E}\left(\varepsilon(\theta) \cdot \left[\Delta u(\theta) - \mu \cdot \left(g'(m(\theta)) - g'(p_1)\right)\right]\right).$$

Note that

$$\Delta u(\theta) - \mu \cdot \left(g'(m(\theta)) - g'(p_1)\right)$$

is the Fréchet derivative of $V^*(\cdot)$ at m. Hence the tangent hyperplane at m can be expressed as

$$\left\{(\widetilde{m}, v) \in M \times \mathbb{R} : v - V^*(m) = \mathbb{E}\left(\left[\Delta u(\theta) - \mu g'(m(\theta)) + \mu g'\left(\int_\Theta m(\theta) \, dP(\theta)\right)\right](\widetilde{m}(\theta) - m(\theta))\right)\right\}.$$

An important observation: since $V^*(\cdot)$ is a concave functional on M, V^* is upper bounded by any hyperplane tangent at any $m \in M$, i.e., $\forall m, \widetilde{m} \in M$,

$$V^*(\widetilde{m}) - V^*(m)$$
$$\leq \mathbb{E}\left(\left[\Delta u(\theta) - \mu \cdot g'(m(\theta)) + \mu \cdot g'\left(\int_\Theta m(\theta) \, dP(\theta)\right)\right](\widetilde{m}(\theta) - m(\theta))\right).$$

This inequality is strict when

$$m \in M^o \triangleq M \setminus \{m \in M : m(\theta) \text{ is a constant } a.s.\}$$

and $\Pr(\widetilde{m}(\theta) \neq m(\theta)) > 0$, since $V^*(\cdot)$ is strictly concave on M^o. This observation is helpful later in our proof.

The optimality of m requires $\left. \frac{dV^*(m + \alpha \cdot \varepsilon)}{d\alpha} \right|_{\alpha=0} \leq 0$ for all feasible perturbation ε. Hence we must have

$$(A1) \qquad \Delta u(\theta) - \mu \cdot \left(g'(m(\theta)) - g'(p_1)\right) \begin{cases} \geq 0 & \text{if } m(\theta) = 1 \\ = 0 & \text{if } m(\theta) \in (0, 1) \\ \leq 0 & \text{if } m(\theta) = 0 \end{cases}.$$

Note that $\Pr(m(\theta) = 1) > 0$ implies $\Pr(m(\theta) = 1) = 1$. Otherwise,

$$p_1 = \mathbb{E}m(\theta) < 1$$

implies

$$\Delta u(\theta) - \mu \cdot \left(g'(m(\theta)) - g'(p_1)\right) = -\infty$$

for all θ in set

$$B = \{\theta \in \Theta : m(\theta) = 1\}.$$

Thus $\varepsilon(\theta) = -1_B$ is a feasible perturbation and

$$\frac{dV^*(m + \alpha \cdot \varepsilon)}{d\alpha}\bigg|_{\alpha=0}$$

$$= \int_B (-\infty) \cdot (-1) \, dP(\theta)$$

$$= +\infty,$$

which contradicts the optimality of m. Hence we know that $\Pr(m(\theta) = 1) > 0$ if and only if $\Pr(m(\theta) = 1) = 1$. The same argument suggests that $\Pr(m(\theta) = 0) > 0$ if and only if $\Pr(m(\theta) = 0) = 1$. Therefore, the optimal strategy m must be one of the three scenarios: a) $p_1 = 1$, i.e., $m(\theta) = 1$ $a.s.$; b) $p_1 = 0$, i.e., $m(\theta) = 0$ $a.s.$; c) $p_1 \in (0, 1)$ and $m(\theta) \in (0, 1)$ $a.s.$.

We first search for the sufficient condition for scenario c). According to (A1), $m(\theta) \in (0, 1)$ $a.s.$ implies

(A2) $\Delta u(\theta) - \mu \cdot \left(g'(m(\theta)) - g'(p_1)\right) = 0$ $a.s.$.

By definition,

$$g'(x) = \ln \frac{x}{1-x},$$

thus (A2) implies

$$m(\theta) = \frac{p_1}{p_1 + (1 - p_1) \cdot \exp\left(-\mu^{-1}\Delta u(\theta)\right)}.$$

Let

(A3) $M_1 = \left\{ m(\theta, p) = \dfrac{p}{p + (1 - p) \cdot \exp\left(-\mu^{-1}\Delta u(\theta)\right)} : p \in [0, 1] \right\}$

and

$$J(p) = \mathbb{E}m(\theta, p),$$

then there exists $p_1 \in [0, 1]$ such that $m(\cdot, p_1) \in M_1$ is an optimal strategy. Note that $J(p_1) = p_1$ is a necessary condition for the optimality of $m(\cdot, p_1)$.

Since $m(\cdot, p_1) \in M_1 \subset M$, the original problem is reduced to

$$\max_{p \in [0,1]} V^*(m(\cdot, p)) = \mathbb{E}\left[\Delta u(\theta) \cdot m(\theta, p)\right] - c(m(\cdot, p)).$$

The first order derivative with respect to p is

$$\frac{dV^*\left(m\left(\cdot,p\right)\right)}{dp}$$

$$= \mathbb{E}\left(\left[\Delta u\left(\theta\right) - \mu \cdot g'\left(m\left(\theta,p\right)\right) + \mu \cdot g'\left(J\left(p\right)\right)\right] \cdot \frac{\partial m\left(\theta,p\right)}{\partial p}\right).$$

By definition,

$$\Delta u\left(\theta\right) - \mu \cdot g'\left(m\left(\theta,p\right)\right) = -\mu \cdot g'\left(p\right),$$

thus

$$\frac{dV^*\left(m\left(\cdot,p\right)\right)}{dp} = \mu \cdot \left[g'\left(J\left(p\right)\right) - g'\left(p\right)\right] \cdot \mathbb{E}\frac{\partial m\left(\theta,p\right)}{\partial p}.$$

Since

$$\frac{\partial m\left(\theta,p\right)}{\partial p} > 0$$

for all $\theta \in \Theta$,

$$\frac{dV^*\left(m\left(\cdot,p\right)\right)}{dp} \geq 0$$

if and only if

$$g'\left(J\left(p\right)\right) - g'\left(p\right) \geq 0.$$

Since g' is strictly increasing, we have

$$\frac{dV^*\left(m\left(\cdot,p\right)\right)}{dp} \geq 0$$

if and only if

$$J\left(p\right) \geq p.$$

In order to be a global maximum, $m\left(\cdot,p_1\right)$ must first be a local maximum within M_1. This requires

(A4) $$J\left(p_1\right) = p_1.$$

But (A4) is not sufficient. The sufficient condition for $m\left(\cdot,p_1\right)$ to be a local maximum within M_1 is

$$\exists \text{ neighborhood } \left(p_1 - \beta, p_1 + \beta\right),$$

$$\text{s.t. } J\left(p\right) \geq p \text{ for all } p \in \left(p_1 - \beta, p_1\right]$$
$$\text{and } J\left(p\right) \leq p \text{ for all } p \in \left[p_1, p_1 + \beta\right).$$

Note that

$$J\left(0\right) = 0, J\left(1\right) = 1,$$

$$\frac{dJ}{dp}\bigg|_{p=0} = \mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right)$$

and

$$\frac{dJ}{dp}\bigg|_{p=1} = \mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) .$$

We proceed by discussing four possible cases.

Case i):

$$\mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) > 1$$

and

$$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) > 1 .$$

In this case, $J(p) > p$ for p close enough to 0 and $J(p) < p$ for p close enough to 1. Since $J(p)$ is continuous, the set $\{p \in (0, 1) : J(p) = p\}$ is non-empty. For any $p_1 \in \{p \in (0, 1) : J(p) = p\}$, the Fréchet derivative at $m(\cdot, p_1)$ is

$$\Delta u\left(\theta\right) - \mu \cdot g'\left(m\left(\theta, p_1\right)\right) + \mu \cdot g'\left(J\left(p_1\right)\right) = 0$$

and thus $m(\cdot, p_1)$ is a critical point of functional $V^*(\cdot)$. Since $m(\cdot, p_1) \in M^o$, the observation mentioned above implies

$$V^*\left(\widetilde{m}\right) - V^*\left(m\left(\cdot, p_1\right)\right) < 0$$

for all $\widetilde{m} \in M$ such that $\Pr\left(\widetilde{m}\left(\theta\right) \neq m\left(\theta, p_1\right)\right) > 0$. Hence, $V^*\left(m\left(\cdot, p_1\right)\right)$ is strictly higher than the values achieved at any other $\widetilde{m} \in M$, i.e.,

$$\{p \in (0, 1) : J(p) = p\} = \{p_1\}$$

and $m(\cdot, p_1)$ is the *unique* global maximum. This actually proves (9).

Case ii):

(A5) $$\mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) > 1$$

and

(A6) $$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) \leq 1 .$$

Note that this case can be summarized by (A6) alone, since (A5) is a direct implication of (A6) according to Jensen's inequality.

First, Inequality (A5) implies $J(p) > p$ for p close enough to 0.

Second, we show that $J(p) > p$ for p close enough to 1. This is obvious if Inequality (A6) holds strictly. Otherwise,

(A7) $$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) = 1 ,$$

which implies

$$\mathbb{E}\exp\left(-2\mu^{-1}\Delta u\left(\theta\right)\right) > \left[\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right)\right]^2 = 1$$

due to Jensen's inequality and the assumption that $\Delta u\left(\theta\right)$ is not constant. Thus we have

(A8) $$\left.\frac{d^2 J}{dp^2}\right|_{p=1} = -2\cdot\left[\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) - \mathbb{E}\exp\left(-2\mu^{-1}\Delta u\left(\theta\right)\right)\right] > 0.$$

Together with (A7), (A8) implies $J\left(p\right) > p$ for p close enough to 1.

Third, we claim that $J\left(p\right) > p$ for all $p \in \left(0, 1\right)$. Suppose this is not true and let

$$p_1 = \sup\left\{p \in \left(0, 1\right) : J\left(p\right) \le p\right\}.$$

Since $J\left(p\right) > p$ for p close enough to 0 and 1, $p_1 \in \left(0, 1\right)$ and thus $m\left(\cdot, p_1\right) \in M^o$. On the one hand, the continuity of $J\left(p\right)$ implies $J\left(p_1\right) = p_1$. Hence $m\left(\cdot, p_1\right)$ is a critical point of functional $V^*\left(\cdot\right)$. By the same argument as in Case i), we know that $m\left(\cdot, p_1\right)$ is the unique global maximum. On the other hand, by the construction of p_1, $J\left(p\right) > p$ for all $p \in \left(p_1, 1\right)$. Then $V^*\left(m\left(\cdot, p\right)\right) > V^*\left(m\left(\cdot, p_1\right)\right)$ for all $p \in \left(p_1, 1\right)$ since $V^*\left(m\left(\cdot, p\right)\right)$ is strictly increasing in p when $J\left(p\right) - p > 0$. This contradicts the unique optimality of $m\left(\cdot, p_1\right)$. Therefore, $J\left(p\right) > p$ for all $p \in \left(0, 1\right)$ and the optimal strategy cannot be an interior point of M (i.e., it cannot be the case $p_1 \in \left(0, 1\right)$.) Then according to our previous discussion, only scenarios a) that $p_1 = 1$ and scenario b) that $p_1 = 0$ are possible. Since $J\left(p\right) > p$ for all $p \in \left(0, 1\right)$, we know that

$$V^*\left(m\left(\cdot, 1\right)\right) > V^*\left(m\left(\cdot, 0\right)\right).$$

Hence, $p_1 = 1$, i.e., $m\left(\theta\right) = 1$ *a.s.* is the *unique* optimal strategy. This actually proves (7).

case iii):

(A9) $$\mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) \le 1$$

and

(A10) $$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) > 1.$$

Note that this case can be summarized by (A9) alone, since (A10) is a direct implication of (A9) according to Jensen's inequality. In this case, by the same argument as in case ii), $m\left(\theta\right) = 0$ *a.s.* is the *unique* optimal strategy. This actually proves (8).

case iv):

(A11) $$\mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) \le 1$$

and

(A12) $$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) \le 1 \,.$$

Case iv) cannot exists since Jensen's inequality implies

$$\mathbb{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) \ge \left[\mathbb{E}\exp\left(\mu^{-1}\Delta u\left(\theta\right)\right)dP\left(\theta\right)\right]^{-1} \,,$$

which suggests (A11) and (A12) hold with equality. This is true only if $\Delta u\left(\theta\right) = 0$ almost surely, a trivial case excluded by our assumption.

Since cases i), ii) and iii) exhaust all possibilities, for each case, the corresponding conditions are not only sufficient but also necessary.

The uniqueness of the optimal strategy is proved in each case.

Remarks to Proposition 1.

These results are intuitive. Since the information cost is convex, the decision maker's objective is concave, which gives rise to the uniqueness of the optimal strategy.

In case a), condition (7) holds if action 1 is very likely the ex ante best action and the cost of acquiring information is sufficiently high. Hence the decision maker just takes action 1 without acquiring any information. Similarly, case b) implies that if action 0 is ex ante very likely to dominate action 1 and the information cost is sufficiently high, the decision maker always takes action 0. In this two cases, marginal benefit of acquiring information is less than the marginal cost. Hence the decision maker chooses not to acquire any information.

In case c), as captured by the two inequalities, neither action 1 nor action 0 is ex ante dominant, thus there is information acquisition and $m\left(\cdot\right)$ is no longer a constant.

In order to get some intuition, consider an extreme case where action 1 is dominant, i.e., the payoff gain $\Delta u\left(\theta\right) > 0$ for all $\theta \in \Theta$. It is obvious that the decision maker will always take action 1 regardless of μ, the marginal cost of information acquisition.

When neither action is dominant, i.e.,

$$\Pr\left(\Delta u\left(\theta\right) > 0\right) > 0 \text{ and } \Pr\left(\Delta u\left(\theta\right) < 0\right) > 0 \,,$$

the marginal cost of information acquisition μ plays a role.

On the one hand,

$$\lim_{\mu \to \infty} \int \exp\left(\pm\mu^{-1}\Delta u\left(\theta\right)\right)dP\left(\theta\right) = 1 \,.$$

Hence Proposition 1 predicts that no information is acquired if μ is high enough.

On the other hand, since

$$
\lim_{\mu \to 0} \frac{d}{d\mu^{-1}} \int \exp\left(\mu^{-1} \Delta u\left(\theta\right)\right) dP\left(\theta\right)
$$

$$
= \lim_{\mu \to 0} \int \exp\left(\mu^{-1} \Delta u\left(\theta\right)\right) \Delta u\left(\theta\right) dP\left(\theta\right)
$$

$$
= \lim_{\mu \to 0} \int_{\Delta u(\theta)>0} \exp\left(\mu^{-1} \Delta u\left(\theta\right)\right) \Delta u\left(\theta\right) dP\left(\theta\right)
$$

$$
+ \Pr\left(\Delta u\left(\theta\right) = 0\right) + \lim_{\mu \to 0} \int_{\Delta u(\theta)<0} \exp\left(\mu^{-1} \Delta u\left(\theta\right)\right) \Delta u\left(\theta\right) dP\left(\theta\right)
$$

$$
= +\infty + \Pr\left(\Delta u\left(\theta\right) = 0\right) + 0
$$

$$
= +\infty \, ,
$$

we have

$$
\lim_{\mu \to 0} \int \exp\left(\mu^{-1} \Delta u\left(\theta\right)\right) dP\left(\theta\right) > 1 \, .
$$

A similar argument leads to

$$
\lim_{\mu \to 0} \int \exp\left(-\mu^{-1} \Delta u\left(\theta\right)\right) dP\left(\theta\right) > 1 \, .
$$

Therefore, Proposition 1 reads that there must exist information acquisition if the marginal cost of information is sufficiently low. This interpretation coincides with our intuition that the decision maker rationally decides whether to acquire information through comparing the cost to the benefit of information acquisition.

When neither action is dominant and the marginal cost of information acquisition takes intermediate values, the decision maker finds it optimal to acquire some information to make her action (partially, in a random manner) contingent on θ. This is the case specified by condition (9). Since g' is strictly increasing, (10) implies that $m\left(\theta\right)$, the conditional probability of choosing action 1, is increasing with respect to payoff gain $\Delta u\left(\theta\right)$. This is intuitive. The left hand side of (10) represents the marginal benefit of increasing $m\left(\theta\right)$, while the right hand side of (10) is the marginal cost of information when increasing $m\left(\theta\right)$. Therefore, if deciding to acquire information, the decision maker will equate her marginal benefit with her marginal cost of doing so.

AN EXAMPLE

The following example provides some intuition behind the decision maker's information acquisition strategy.

Let θ distribute according to $N\left(t, 1\right)$ and

$$
\Delta u\left(\theta\right) = \theta \, .
$$

It is easy to verify that the decision maker always chooses action 1 (action 0) if and only

if $t \geq \mu^{-1}/2$ ($t \leq -\mu^{-1}/2$). In this case, action 1 (action 0) is superior to action 0 (action 1) ex ante (i.e., $|t|$ is large) and the cost in acquiring information is relatively high (i.e., μ is large). Hence it is not worth acquiring any information at all.

Let $t = 0$, then the decision maker finds it optimal to acquire some information. According to (10), the optimal information acquisition strategy $m(\theta)$ satisfies

$$
\text{(A13)} \qquad \theta/\mu = g'(m(\theta)) - g'\left(\int_{\Theta} m(\theta)\, dP(\theta)\right),
$$

where

$$
g'(m) = \ln \frac{m}{1-m} .
$$

Since prior $N(0,1)$ is symmetric about the origin and payoff gain $\Delta u(\theta)$ is an odd function, the decision maker is indifferent on average, i.e.,

$$
\int_{\Theta} m(\theta)\, dP(\theta) = 1/2 .
$$

Hence

$$
g'\left(\int_{\Theta} m(\theta)\, dP(\theta)\right) = 0
$$

and (A13) becomes

$$
\theta/\mu = \ln \frac{m(\theta)}{1-m(\theta)} .
$$

Therefore,

$$
\text{(A14)} \qquad m(\theta) = \frac{1}{1+\exp(-\theta/\mu)} .
$$

First note that

$$
\lim_{\mu \to 0} m(\theta) = a(\theta) \triangleq \begin{cases} 1 & \text{if } \theta \geq 0 \\ 0 & \text{if } \theta < 0 \end{cases} .
$$

Step function $a(\theta)$ captures the decision maker's choice under complete information. In this case, the decision maker can observe the exact value of θ. When $\mu > 0$, the best response is characterized by (A14). Since information is no longer free, the decision maker has to allow some mistake in her response. The conditional probability of mistake is given by

$$
|m(\theta) - a(\theta)| ,
$$

which is decreasing in $|\theta|$, the "price" of mistake. Therefore, the decision maker deliberately acquires information to balance the price of mistake and the cost of information.

Second, parameter μ measures the difficulty in acquiring information. Figure A1 shows how $m(\theta)$ varies with this parameter.

When $\mu = 0$, information acquisition incurs no cost and the decision maker's response is a step function. He never makes mistake. When μ becomes larger, he starts to com-
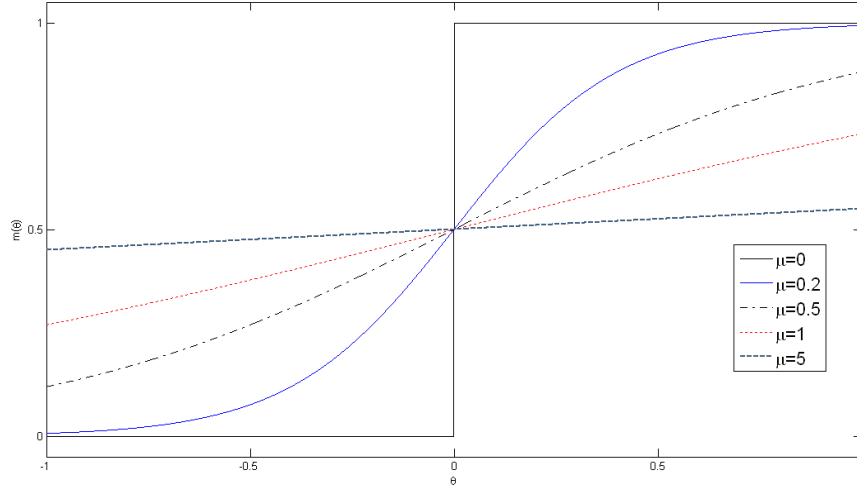
FIGURE A1. INFORMATION ACQUISITION UNDER VARIOUS INFORMATION COSTS

promise the accuracy of her decision to save information cost. Larger μ leads to flatter $m(\theta)$. Finally, when μ is extremely large, $m(\theta)$ is almost constant and the decision maker almost stops acquiring information.

Third, since the decision maker's action is highly sensitive to θ where slope $\left|\frac{dm(\theta)}{d\theta}\right|$ is large, $\left|\frac{dm(\theta)}{d\theta}\right|$ reflects her attentiveness around θ. Under this interpretation, Figure A1 reveals that the decision maker actively collects information for intermediate values of the fundamental but is rationally inattentive to values at the tails. This result coincides with our intuition. When θ is too high (low), the decision maker should take action 1 (action 0) anyway. Hence the information about θ on the tails are not so relevant to her payoff. When θ takes intermediate values, the decision maker's payoff gain from taking action 1 over action 0 depends crucially on the sign of θ. Therefore, the information about θ around zero is payoff-relevant and attracts most of her attention.

APPENDIX B

Proof of Proposition 3.
PROOF: According to (11),

$$\forall \theta \quad \in \quad \Theta,$$

$$(\text{B1}) \quad \theta - r \cdot (1 - m_1(\theta)) \quad = \quad \mu \cdot \left[\ln\left(\frac{m_2(\theta)}{1 - m_2(\theta)}\right) - \ln\left(\frac{p_{I2}}{1 - p_{I2}}\right) \right],$$

$$(\text{B2}) \quad \theta - r \cdot (1 - m_2(\theta)) \quad = \quad \mu \cdot \left[\ln\left(\frac{m_1(\theta)}{1 - m_1(\theta)}\right) - \ln\left(\frac{p_{I1}}{1 - p_{I1}}\right) \right].$$

(B1) and (B2) imply

$$\forall \theta \in \Theta,$$
$$\left[\ln \left(\frac{p_{I2}}{1 - p_{I2}} \right) - \ln \left(\frac{p_{I1}}{1 - p_{I1}} \right) \right]$$
$$\text{(B3)} \quad = \quad \left[\ln \left(\frac{m_2 (\theta)}{1 - m_2 (\theta)} \right) - \ln \left(\frac{m_1 (\theta)}{1 - m_1 (\theta)} \right) \right] + \frac{r}{\mu} (m_2 (\theta) - m_1 (\theta)).$$

If $p_{I2} = p_{I1}$, (B3) becomes

$$\forall \theta \in \Theta,$$
$$0 \quad = \quad \left[\ln \left(\frac{m_2 (\theta)}{1 - m_2 (\theta)} \right) - \ln \left(\frac{m_1 (\theta)}{1 - m_1 (\theta)} \right) \right] + \frac{r}{\mu} (m_2 (\theta) - m_1 (\theta)),$$

and we must have $m_2 (\theta) = m_1 (\theta) \, a.s.$ since $\frac{r}{\mu} > 0$. Now suppose $p_{I2} \neq p_{I1}$. Without loss of generality, let $p_{I2} > p_{I1}$. Denote $z = \ln \left(\frac{p_{I2}}{1 - p_{I2}} \right) - \ln \left(\frac{p_{I1}}{1 - p_{I1}} \right) > 0$. Then (B3) becomes

$$\forall \theta \in \Theta,$$
$$\text{(B4)} \quad 0 \quad < \quad z = \left[\ln \left(\frac{m_2 (\theta)}{1 - m_2 (\theta)} \right) - \ln \left(\frac{m_1 (\theta)}{1 - m_1 (\theta)} \right) \right] + \frac{r}{\mu} (m_2 (\theta) - m_1 (\theta)),$$

which suggests that $\Pr (m_2 (\theta) > m_1 (\theta)) = 1$. Let $\ln \left(\frac{m_2(\theta)}{1-m_2(\theta)} \right) = x (\theta)$ and $\ln \left(\frac{m_1(\theta)}{1-m_1(\theta)} \right) = y (\theta)$. (B4) implies

$$\forall \theta \in \Theta, x (\theta) < y (\theta) + z.$$

Note that $p_{Ii} = \int m_i (\theta) \cdot dP (\theta) = Em_i (\theta)$, $i \in \{1, 2\}$, $m_2 (\theta) = \frac{\exp(x(\theta))}{1+\exp(x(\theta))}$ and $m_1 (\theta) = \frac{\exp(y(\theta))}{1+\exp(y(\theta))}$, thus

$$
\begin{aligned}
z \quad &= \quad \ln \left(\frac{Em_2 (\theta)}{1 - Em_2 (\theta)} \right) - \ln \left(\frac{Em_1 (\theta)}{1 - Em_1 (\theta)} \right) \\
&= \quad \ln \left(\frac{E \left[\frac{\exp(x(\theta))}{1+\exp(x(\theta))} \right]}{E \left[\frac{1}{1+\exp(x(\theta))} \right]} \right) - \ln \left(\frac{E \left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))} \right]}{E \left[\frac{1}{1+\exp(y(\theta))} \right]} \right) \\
&< \quad \ln \left(\frac{E \left[\frac{\exp(y(\theta)+z)}{1+\exp(y(\theta)+z)} \right]}{E \left[\frac{1}{1+\exp(y(\theta)+z)} \right]} \right) - \ln \left(\frac{E \left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))} \right]}{E \left[\frac{1}{1+\exp(y(\theta))} \right]} \right).
\end{aligned}
$$

Take the exponential of both sides of the above inequality, we have

$$\exp(z) < \frac{E\left[\frac{\exp(y(\theta)+z)}{1+\exp(y(\theta)+z)}\right] \cdot E\left[\frac{1}{1+\exp(y(\theta))}\right]}{E\left[\frac{1}{1+\exp(y(\theta)+z)}\right] \cdot E\left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))}\right]},$$

i.e.,

$$E\left[\frac{\exp(y(\theta))}{1+\exp(y(\theta)+z)}\right] \cdot E\left[\frac{1}{1+\exp(y(\theta))}\right] > E\left[\frac{1}{1+\exp(y(\theta)+z)}\right] \cdot E\left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))}\right],$$

i.e.,

$$\int \frac{\exp(y(\theta_1))\,dP(\theta_1)}{1+\exp(y(\theta_1)+z)} \cdot \int \frac{dP(\theta_2)}{1+\exp(y(\theta_2))} + \int \frac{\exp(y(\theta_2))\,dP(\theta_2)}{1+\exp(y(\theta_2)+z)} \cdot \int \frac{dP(\theta_1)}{1+\exp(y(\theta_1))}$$
$$> \int \frac{dP(\theta_1)}{1+\exp(y(\theta_1)+z)} \cdot \int \frac{\exp(y(\theta_2))\,dP(\theta_2)}{1+\exp(y(\theta_2))} + \int \frac{dP(\theta_2)}{1+\exp(y(\theta_2)+z)} \cdot \int \frac{\exp(y(\theta_1))\,dP(\theta_1)}{1+\exp(y(\theta_1))},$$

i.e.,
(B5)
$$\int \frac{A+B-C-D}{\left[1+\exp(y(\theta_1)+z)\right]\left[1+\exp(y(\theta_2))\right]\left[1+\exp(y(\theta_2)+z)\right]\left[1+\exp(y(\theta_1))\right]}\,dP(\theta_1)\,dP(\theta_2) > 0,$$

where

$$A = \exp(y(\theta_1))\left[1+\exp(y(\theta_2)+z)\right]\left[1+\exp(y(\theta_1))\right],$$
$$B = \exp(y(\theta_2))\left[1+\exp(y(\theta_1)+z)\right]\left[1+\exp(y(\theta_2))\right],$$
$$C = \exp(y(\theta_2))\left[1+\exp(y(\theta_2)+z)\right]\left[1+\exp(y(\theta_1))\right]$$

and

$$D = \exp(y(\theta_1))\left[1+\exp(y(\theta_1)+z)\right]\left[1+\exp(y(\theta_2))\right].$$

Let $y(\theta_1) = u$ and $y(\theta_2) = v$, then the numerator in the integral becomes

$$A+B-C-D = \left[e^u - e^v\right]^2\left[1-e^z\right] < 0,$$

where the last inequality follows the fact that $z > 0$. Therefore, the left hand side of (B5) is strictly negative, which is a contradiction. Therefore, $\Pr(m_1(\theta) = m_2(\theta)) = 1$.

LEMMA 3: *Game $G(r, \mu)$ has an equilibrium with at least one player always investing if and only if $\mathbf{E}\exp\left(-\mu^{-1}\theta\right) \leq 1$, where expectations are formed according to common prior P.*

PROOF: (Sufficiency.) If $m_j(\theta) = 1$ for all $\theta \in \Theta$, player i's payoff gain from investing over not investing becomes

$$\Delta u_i(\theta) = \theta.$$

Then according to case a) in ii) of Proposition 1,

$$\mathbf{E} \exp\left(-\mu^{-1}\theta\right) \leq 1$$

implies $m_i(\theta) = 1$ for all $\theta \in \Theta$, which confirms that $m_j(\theta) = 1$ for all $\theta \in \Theta$ is player j's optimal strategy. Therefore, we have an equilibrium with both players always investing.

(Necessity.) Suppose $m_j(\theta) = 1$ for all $\theta \in \Theta$, but $\mathbf{E} \exp\left(-\mu^{-1}\theta\right) > 1$. Player i's payoff gain from investing over not investing is

$$\Delta u_i(\theta) = \theta .$$

According to case b) and c) in ii) of Proposition 1, $\mathbf{E} \exp\left(-\mu^{-1}\theta\right) > 1$ implies $m_i(\theta) < 1$ for all $\theta \in \Theta$. Then player j's payoff gain from investing over not investing becomes

$$\begin{aligned}
\Delta u_j(\theta) &= \theta - r \cdot [1 - m_i(\theta)] \\
&< \theta \text{ for all } \theta \in \Theta,
\end{aligned}$$

which implies

$$\begin{aligned}
&\mathbf{E} \exp\left(-\mu^{-1}\Delta u_j(\theta)\right) \\
> \; &\mathbf{E} \exp\left(-\mu^{-1}\theta\right) \\
> \; &1 .
\end{aligned}$$

Hence according to case b) and c) in ii) of Proposition 1, we find $m_j(\theta) < 1$ for all $\theta \in \Theta$, which is a contradiction.

LEMMA 4: *In an equilibrium of game $G(r, \mu)$ with one player always investing, the other player must also always invest.*

PROOF: By the necessity part of Lemma 3, we know that $\mathbf{E} \exp\left(-\mu^{-1}\theta\right) \leq 1$. Then the sufficiency part of Lemma 3 has already proved that the other player must also always invest.

PROPOSITION 9: *Game $G(r, \mu)$ has an equilibrium with both players always investing if and only if $\mathbf{E} \exp\left(-\mu^{-1}\theta\right) \leq 1$.*

PROOF: This proposition is a direct implication of Lemma 3 and 4.

LEMMA 5: *Game $G(r, \mu)$ has an equilibrium with at least one player always not investing if and only if $\mathbf{E} \exp\left(\mu^{-1}\theta\right) \leq e^{\mu^{-1}r}$, where expectations are formed according to common prior P.*

PROOF: (Sufficiency.) If $m_j(\theta) = 0$ for all $\theta \in \Theta$, player i's payoff gain from investing over not investing becomes
$$\Delta u_i(\theta) = \theta - r .$$

Then according to case b) in ii) of Proposition 1,

$$\mathbf{E} \exp\left(\mu^{-1}\left(\theta - r\right)\right) \leq 1$$

implies $m_i\left(\theta\right) = 0$ for all $\theta \in \Theta$, which confirms that $m_j\left(\theta\right) = 0$ for all $\theta \in \Theta$ is player j's optimal strategy. Therefore, we have an equilibrium with both players always not investing.

(Necessity.) Suppose $m_j\left(\theta\right) = 0$ for all $\theta \in \Theta$, but $\mathbf{E} \exp\left(\mu^{-1}\theta\right) > e^{\mu^{-1}r}$, i.e., $\mathbf{E} \exp\left(\mu^{-1}\left(\theta - r\right)\right) > 1$. Player i's payoff gain from investing over not investing is

$$\Delta u_i\left(\theta\right) = \theta - r .$$

According to case a) and c) in ii) of Proposition 1, $\mathbf{E} \exp\left(\mu^{-1}\left(\theta - r\right)\right) > 1$ implies $m_i\left(\theta\right) > 0$ for all $\theta \in \Theta$. Then player j's payoff gain from investing over not investing becomes

$$\begin{aligned}
\Delta u_j\left(\theta\right) &= \theta - r \cdot \left[1 - m_i\left(\theta\right)\right] \\
&> \theta - r \text{ for all } \theta \in \Theta,
\end{aligned}$$

which implies

$$\begin{aligned}
&\mathbf{E} \exp\left(\mu^{-1}\Delta u_j\left(\theta\right)\right) \\
> \ &\mathbf{E} \exp\left(\mu^{-1}\left(\theta - r\right)\right) \\
> \ &1 .
\end{aligned}$$

Hence according to case a) and c) in ii) of Proposition 1, we find $m_j\left(\theta\right) > 0$ for all $\theta \in \Theta$, which is a contradiction.

LEMMA 6: *In an equilibrium of game $G\left(r, \mu\right)$ with one player always not investing, the other player must always not invest either.*

PROOF: By the necessity part of Lemma 5, we know that $\mathbf{E} \exp\left(\mu^{-1}\theta\right) \leq e^{\mu^{-1}r}$. Then the sufficiency part of Lemma 5 has already proved that the other player must always not invest either.

PROPOSITION 10: *Game $G\left(r, \mu\right)$ has an equilibrium with both players always not investing if and only if $\mathbf{E} \exp\left(\mu^{-1}\theta\right) \leq e^{\mu^{-1}r}$.*

PROOF: This proposition is a direct implication of Lemma 5 and 6.

APPENDIX C

Proof of Lemma 2.
PROOF: Let $m \in \mathbf{M}\left(r, \mu\right)$ be an arbitrary shape. Let $\theta_0\left(p_I\right)$ be defined by (14) and

$$m\left(\theta, p_I\right) = m\left(\theta - \theta_0\left(p_I\right)\right) .$$

By definition, $m(\theta, p_I)$ satisfies

(C1) $\theta - r \cdot (1 - m(\theta, p_I)) = \mu \cdot \left[\ln\left(\dfrac{m(\theta, p_I)}{1 - m(\theta, p_I)} \right) - \ln\left(\dfrac{p_I}{1 - p_I} \right) \right]$ for all $\theta \in \Theta$.

Here $p_I \in (0, 1)$ is treated as an index and $m(\theta, p_I)$ is an equilibrium if and only if

(C2) $$p_I = \int_\Theta m(\theta, p_I)\, dP(\theta) \ .$$

Therefore, our objective is to show the existence of $p_I \in (0, 1)$ satisfying (C2).

Step 1. We show

$$\int_\Theta m(\theta, p_I)\, dP(\theta) < p_I$$

for p_I sufficiently close to 1.

By (C1),

$$\ln\left(\frac{m(\theta, p_I)}{1 - m(\theta, p_I)} \right) - \ln\left(\frac{p_I}{1 - p_I} \right) < \mu^{-1}\theta \text{ for all } \theta \in \Theta,$$

i.e.,

$$m(\theta, p_I) < \frac{\frac{p_I}{1 - p_I}}{e^{-\mu^{-1}\theta} + \frac{p_I}{1 - p_I}} \text{ for all } \theta \in \Theta.$$

Hence it suffices to show

$$\int_\Theta \frac{\frac{p_I}{1 - p_I}}{e^{-\mu^{-1}\theta} + \frac{p_I}{1 - p_I}}\, dP(\theta) \le p_I \ .$$

Let

$$w = \frac{1}{1 - p_I}$$

and

$$v(\theta) = e^{-\mu^{-1}\theta} - 1,$$

then it suffices to show

(C3) $$\int_\Theta \frac{w - 1}{v(\theta) + w}\, dP(\theta) \le \frac{w - 1}{w} \ .$$

Since $w > 1$ by definition, (C3) becomes

(C4) $$\int_\Theta \frac{1}{1 + v(\theta)/w}\, dP(\theta) \le 1 \ .$$

By assumption,

$$\int_{\Theta} e^{-\mu^{-1}\theta} dP(\theta) > 1,$$

i.e.,

(C5)
$$\int_{\Theta} \upsilon(\theta) dP(\theta) > 0.$$

Hence there exists $N > 0$ s.t.

$$\int_{\Theta \cap [-N,+\infty)} \upsilon(\theta) dP(\theta) > 0.$$

Let

$$B = \max\left(e^{\mu^{-1}N} - 1, 1\right),$$

then

$$|\upsilon(\theta)| \le B$$

for all $\theta \in [-N, +\infty)$. Since

$$\frac{1}{1+x} = 1 - x + x^2 + o\left(x^2\right)$$

for x close enough to zero, there exists $\overline{w} > 0$ s.t.

$$\frac{1}{1 + \upsilon(\theta)/w} < 1 - \frac{\upsilon(\theta)}{w} + \frac{2B^2}{w^2}$$

for all $\theta \in [-N, +\infty)$ and $w > \overline{w}$. Choose

(C6)
$$w > \max\left(\overline{w}, \frac{2B^2}{\int_{\Theta \cap [-N,+\infty)} \upsilon(\theta) dP(\theta)}\right),$$

then

$$\int_{\Theta \cap [-N,+\infty)} \frac{1}{1 + \upsilon(\theta)/w} dP(\theta)$$

$$< \int_{\Theta \cap [-N,+\infty)} \left[1 - \frac{\upsilon(\theta)}{w} + \frac{2B^2}{w^2}\right] dP(\theta)$$

$$= \Pr(\theta \ge -N) + \frac{2B^2}{w^2} \cdot \Pr(\theta \ge -N) - w^{-1} \int_{\Theta \cap [-N,+\infty)} \upsilon(\theta) dP(\theta)$$

$$\le \Pr(\theta \ge -N) + \frac{2B^2}{w^2} - w^{-1} \int_{\Theta \cap [-N,+\infty)} \upsilon(\theta) dP(\theta)$$

(C7)
$$< \Pr(\theta \ge -N),$$

where the last inequality follows (C6). Hence,

$$\int_\Theta \frac{1}{1 + \upsilon(\theta)/w} dP(\theta)$$

$$= \int_{\Theta \cap [-N,+\infty)} \frac{1}{1 + \upsilon(\theta)/w} dP(\theta) + \int_{\Theta \cap (-\infty,-N)} \frac{1}{1 + \upsilon(\theta)/w} dP(\theta)$$

$$\leq \int_{\Theta \cap [-N,+\infty)} \frac{1}{1 + \upsilon(\theta)/w} dP(\theta) + \int_{\Theta \cap (-\infty,-N)} 1 \cdot dP(\theta)$$

$$< \Pr(\theta \geq -N) + \Pr(\theta < -N)$$

$$= 1,$$

where the first inequality holds since $\upsilon(\theta)$ for all $\theta \in (-\infty, -N)$ and the last inequality comes from (C7). Therefore, (C4) holds and if we let

$$\overline{p}_I = \frac{w-1}{w},$$

we have

$$\int_\Theta m(\theta, \overline{p}_I) dP(\theta) < \overline{p}_I.$$

Step 2. We show

$$\int_\Theta m(\theta, p_I) dP(\theta) > p_I$$

for p_I sufficiently close to 0.

By (C1),

$$\ln\left(\frac{m(\theta, p_I)}{1 - m(\theta, p_I)}\right) - \ln\left(\frac{p_I}{1 - p_I}\right) > \mu^{-1}(\theta - r) \text{ for all } \theta \in \Theta,$$

i.e.,

$$1 - m(\theta, p_I) < \frac{1}{1 + e^{\mu^{-1}(\theta-r)\frac{p_I}{1-p_I}}} \text{ for all } \theta \in \Theta.$$

Hence it suffices to show

$$\int_\Theta \frac{1}{1 + e^{\mu^{-1}(\theta-r)\frac{p_I}{1-p_I}}} dP(\theta) \leq 1 - p_I.$$

Let

$$w = \frac{1}{p_I}$$

and

$$\upsilon(\theta) = e^{\mu^{-1}(\theta-r)} - 1,$$

then it suffices to show

(C8)
$$\int_{\Theta} \frac{w-1}{v(\theta)+w} dP(\theta) \leq \frac{w-1}{w} .$$

By assumption,
$$\int_{\Theta} e^{\mu^{-1}(\theta-r)} dP(\theta) > 1,$$

i.e.,

(C9)
$$\int_{\Theta} v(\theta) dP(\theta) > 0 .$$

Note that (C8) and (C9) are the same as (C4) and (C5), thus (C8) can be proved by the same argument in Step 1. Therefore, we can find a $\underline{p}_I \in (0, 1)$ s.t.

$$\int_{\Theta} m\left(\theta, \underline{p}_I\right) dP(\theta) > \underline{p}_I .$$

Step 3. Since common prior P is absolutely continuous with respect to Lebesgue measure over $\mathbb{R}$,

$$\int_{\Theta} m(\theta, p_I) dP(\theta) - p_I$$

is a continuous function of $p_I \in (0, 1)$. Hence Step 1 and Step 2 imply the existence of $p_I^* \in (0, 1)$ s.t.

$$\int_{\Theta} m\left(\theta, p_I^*\right) dP(\theta) = p_I^* .$$

According to (14), let

$$\theta_0^* = r/2 - \mu \cdot \ln\left(\frac{p_I^*}{1-p_I^*}\right) ,$$

then $m\left(\theta - \theta_0^*\right)$ is an equilibrium with shape m. This concludes the proof.

LEMMA 7: *Let P be any probability measure over $\mathbb{R}$. A set of functions $M \subset L^1(\mathbb{R}, P)$ is relatively compact if M is uniformly bounded and equicontinuous.*

PROOF: Let $B > 0$ be the uniform bound and $\{m_n\}_{n=1}^{\infty} \subset M$ be a sequence of functions. Let

$$A_T = \left\{-T, -T + \frac{1}{T}, -T + \frac{2}{T}, \cdots, T - \frac{2}{T}, T - \frac{1}{T}, \right\} ,$$

then $\cup_{T=1}^{\infty} A_T$ is dense in $\mathbb{R}$. Since $\cup_{T=1}^{\infty} A_T$ is countable, we can list its elements as $\{\theta_1, \theta_2, \theta_3, \cdots\}$. Note that the numerical sequence $\{m_n(\theta_1)\}_{n=1}^{\infty}$ is bounded, so by Bolzano-Weierstrass theorem it has a convergent subsequence, which we will write using double subscripts: $\{m_{1,n}(\theta_1)\}_{n=1}^{\infty}$. Now the numerical sequence $\{m_{1,n}(\theta_2)\}_{n=1}^{\infty}$ is also bounded, so it has a convergent subsequence $\{m_{2,n}(\theta_2)\}_{n=1}^{\infty}$. Note that the sequence of

functions $\{m_{2,n}\}_{n=1}^{\infty}$ converges at both θ_1 and θ_2 since it is a subsequence of $\{m_{1,n}\}_{n=1}^{\infty}$. Proceeding in this fashion we obtain a countable collection of subsequences of our original sequence:

$$
\begin{array}{cccc}
m_{1,1} & m_{1,2} & m_{1,3} & \cdots \\
m_{2,1} & m_{2,2} & m_{2,3} & \cdots \\
m_{3,1} & m_{3,2} & m_{3,3} & \cdots \\
\cdot & \cdot & \cdot & \cdots \\
\cdot & \cdot & \cdot & \cdots \\
\cdot & \cdot & \cdot & \cdots
\end{array} \; ,
$$

where the sequence in the n-th row converges at the points $\theta_1, \theta_2, \cdots, \theta_n$ and each row is a subsequence of the one above it. Hence the diagonal sequence $\{m_{n,n}\}_{n=1}^{\infty}$ is a subsequence of the original sequence $\{m_n\}_{n=1}^{\infty}$ that converges at each point of $\cup_{T=1}^{\infty} A_T$. Now we show that $\{m_{n,n}\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^1(\mathbb{R}, P)$.

For any $\varepsilon > 0$, there exists T_0 such that

(C10) $$\Pr\left([-T_0, T_0]\right) \geq 1 - \frac{\varepsilon}{5 \cdot B} \; ,$$

where B is the uniform bound such that $|m(\theta)| < B$ for all $\theta \in \Theta$ and $m \in M$. Since M is equicontinuous, there exists $T_1 > T_0$ such that $\forall m \in M, \forall \theta_1, \theta_2 \in \Theta$,

$$|\theta_1 - \theta_2| < \frac{1}{T_1}$$

implies

$$|m(\theta_1) - m(\theta_2)| < \frac{\varepsilon}{5} \; .$$

As A_{T_1} is finite and $\{m_{n,n}\}_{n=1}^{\infty}$ converges at every point of A_{T_1}, there exists $n_0 \in \mathbb{N}$ such that

$$\left| m_{n,n}(\theta) - m_{n',n'}(\theta) \right| < \frac{\varepsilon}{5}$$

for all $n, n' > n_0$ and all $\theta \in A_{T_1}$. For any $y \in [-T_1, T_1]$, there exists $\theta \in A_{T_1}$ such that

$$|y - \theta| < \frac{1}{T_1} \; ,$$

thus we have

$$\left| m_{n,n}(y) - m_{n,n}(\theta) \right| < \frac{\varepsilon}{5}$$

and

$$\left| m_{n',n'}(y) - m_{n',n'}(\theta) \right| < \frac{\varepsilon}{5} \; .$$

Hence for any $y \in [-T_1, T_1]$ and $n, n' > n_0$,

$$
\begin{aligned}
&\left| m_{n',n'}(y) - m_{n,n}(y) \right| \\
\leq\ & \left| m_{n',n'}(y) - m_{n',n'}(\theta) \right| + \left| m_{n',n'}(\theta) - m_{n,n}(\theta) \right| + \left| m_{n,n}(\theta) - m_{n,n}(y) \right| \\
<\ & \frac{3 \cdot \varepsilon}{5} \ .
\end{aligned}
$$

(C11)

Then

$$
\begin{aligned}
&\left\| m_{n',n'} - m_{n,n} \right\|_{L^1(\mathbb{R}, P)} \\
=\ & \int_\Theta \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y) \\
=\ & \int_{[-T_1, T_1]} \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y) + \int_{\Theta \setminus [-T_1, T_1]} \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y) \\
<\ & \frac{3 \cdot \varepsilon}{5} \cdot \Pr([-T_1, T_1]) + \int_{\Theta \setminus [-T_1, T_1]} 2 \cdot B \cdot dP(y) \\
\leq\ & \frac{3 \cdot \varepsilon}{5} \cdot 1 + \frac{\varepsilon}{5 \cdot B} \cdot 2 \cdot B \\
=\ & \varepsilon \ ,
\end{aligned}
$$

where the first inequality follows (C11) and the second inequality comes from (C10).

Therefore, $\left\{ m_{n,n} \right\}_{n=1}^\infty$ is a Cauchy subsequence of $\{m_n\}_{n=1}^\infty$ in $L^1(\mathbb{R}, P)$ and M is relatively compact in $L^1(\mathbb{R}, P)$. This concludes the proof.

Proof of Proposition 5.

PROOF: We prove by contradiction. Suppose the proposition does not hold, then $\exists i \in \{1, 2\}$, $\overline{\beta}_i > 0$ and a sequence $\{c_n\}_{n=1}^\infty$ s.t. $\lim_{n \to \infty} c_n = 0$ and $\forall n$, $\beta_i(c_n) \leq \overline{\beta}_i$.

We write g_β for the density function over signals induced by precision β, i.e.,

$$
g_\beta(x) = \int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot p(\theta) \cdot d\theta,
$$

and write $l_\beta(\cdot|x)$ for the induced posterior density over θ:

$$
l_\beta(\theta|x) = \frac{\beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot p(\theta)}{g_\beta(x)}.
$$

A sufficient statistic for a player j's conjecture over i's play is the probability he attaches to player i investing as a function of θ, which is a function $m_i : \Theta \to [0, 1]$. Let

$$
S \triangleq \{s \text{ Lebesgue measurable} : \forall x \in \mathbb{R}, s(x) \in [0, 1]\}
$$

and

$$M_{\overline{\beta}_i} \triangleq \left\{ m \in \Omega : \exists \beta \in \left[0, \overline{\beta}_i\right] \text{ and } s \in S, \text{ s.t. } m\left(\theta\right) = \int_x \beta^{1/2} \cdot f\left(\beta^{1/2}\left(x - \theta\right)\right) \cdot s\left(x\right) \cdot dx \text{ for all } \theta \in \Theta \right\},$$

where $\Omega = \left\{ m \in L^1\left(\Theta, P\right) : \forall \theta \in \Theta, m\left(\theta\right) \in [0, 1] \right\}$. $M_{\overline{\beta}_i}$ contains all player j's possible conjectures of player i's play when $\beta_i \in \left[0, \overline{\beta}_i\right]$.

Step 1: We prove that $M_{\overline{\beta}_i}$ is relatively compact in Ω, i.e., its closure $\overline{M_{\overline{\beta}_i}}$ is compact.

Let f' denote the derivative of f. f' could be a generalized function. Since f' is Lebesgue integrable over $\mathbb{R}$,

$$\int_y \max\left(f'\left(y\right), 0\right) dy < \infty$$

and

$$\int_y \max\left(-f'\left(y\right), 0\right) dy < \infty$$

hold by definition. $\forall m \in M_{\overline{\beta}_i}, \forall \theta \in \Theta$,

$$
\begin{aligned}
\left|\frac{dm\left(\theta\right)}{d\theta}\right| &= \left|\int_x \beta \cdot f'\left(\beta^{1/2}\left(x - \theta\right)\right) \cdot s\left(x\right) \cdot dx\right| \\
&= \left|\int_y \beta^{1/2} \cdot f'\left(y\right) \cdot s\left(\beta^{-1/2} \cdot y + \theta\right) \cdot dy\right| \\
&\leq \beta^{1/2} \cdot \max\left[\int_y \max\left(f'\left(y\right), 0\right) dy, \int_y \max\left(-f'\left(y\right), 0\right) dy\right].
\end{aligned}
$$

Hence for any $\varepsilon > 0$,

$$|\theta_1 - \theta_2| < \frac{\varepsilon}{\beta^{1/2} \cdot \max\left[\int_y \max\left(f'\left(y\right), 0\right) dy, \int_y \max\left(-f'\left(y\right), 0\right) dy\right]}$$

implies

$$|m\left(\theta_1\right) - m\left(\theta_2\right)| < \varepsilon$$

for all $m \in M_{\overline{\beta}_i}$, i.e., $M_{\overline{\beta}_i}$ is equicontinuous. By definition, $\forall m \in M_{\overline{\beta}_i}, \forall \theta \in \Theta$, $|m\left(\theta\right)| \leq 1$, i.e., $M_{\overline{\beta}_i}$ is uniformly bounded. Therefore, according to Lemma 7, $M_{\overline{\beta}_i}$ is relatively compact in Ω.

If player j chooses $\left(\beta_j, s_j\right)$ against conjecture m_i, her expected utility is

$$V_j\left(\beta_j, s_j, m_i\right) = \int_{x_j} s_j\left(x_j\right) \cdot \left[\int_\theta \left(\theta - r \cdot \left(1 - m_i\left(\theta\right)\right)\right) \cdot l_{\beta_j}\left(\theta | x_j\right) \cdot d\theta\right] \cdot g_{\beta_j}\left(x_j\right) \cdot dx_j.$$

With an optimal choice of s_j this gives

$$
\begin{aligned}
& V_j^* \left(\beta_j, m_i \right) \\
= \ & \int_{x_j} \max \left\{ 0, \int_\theta (\theta - r \cdot (1 - m_i(\theta))) \cdot l_{\beta_j} \left(\theta | x_j \right) \cdot d\theta \right\} \cdot g_{\beta_j} \left(x_j \right) \cdot dx_j \\
= \ & \int_{x_j} \max \left\{ 0, \int_\theta (\theta - r \cdot (1 - m_i(\theta))) \cdot \beta_j^{1/2} f \left(\beta_j^{1/2} \left(x_j - \theta \right) \right) p(\theta) \, d\theta \right\} \cdot dx_j \ .
\end{aligned}
$$

(C12)

Note that $\lim_{\beta \to \infty} \beta^{1/2} \cdot f \left(\beta^{1/2} (x - \theta) \right) = \delta(x - \theta)$, where $\delta(\cdot)$ is the Dirac delta function. Then (C12) implies

$$
\begin{aligned}
V_j^{**} (m_i) \ & \triangleq \ \lim_{\beta_j \to \infty} V_j^* \left(\beta_j, m_i \right) \\
& = \ \int_{x_j} \max \left\{ 0, \left[x_j - r \cdot \left(1 - m_i \left(x_j \right) \right) \right] \cdot p\left(x_j \right) \right\} \cdot dx_j \\
& = \ \int_\theta \max \left\{ 0, \left[\theta - r \cdot (1 - m_i(\theta)) \right] \cdot p(\theta) \right\} \cdot d\theta \ .
\end{aligned}
$$

(C13)

$V_j^{**}(m_i)$ is player j's ex ante expected utility against conjecture m_i if he can always observe the exact realization of the fundamental.

Step 2: We show that $\forall m_i \in \Omega$, $\forall \beta_j > 0$, $V_j^{**}(m_i) > V_j^*\left(\beta_j, m_i \right)$.

Note that our assumptions

$$
\mathbf{E} \exp \left(-\mu^{-1} \theta \right) > 1
$$

and

$$
\mathbf{E} \exp \left(\mu^{-1} \theta \right) > e^{\mu^{-1} r}
$$

imply that

$$
\Pr \left(\theta < 0 \right) > 0
$$

and

$$
\Pr \left(\theta > r \right) > 0 \, ,
$$

respectively. Hence we have

$$
\begin{aligned}
& \Pr \left(\theta - r \cdot (1 - m_i(\theta)) > 0 \right) \\
\geq \ & \Pr \left(\theta - r > 0 \right) \\
> \ & 0
\end{aligned}
$$

and

$$
\begin{aligned}
& \Pr\left(\theta - r \cdot (1 - m_i(\theta)) < 0\right) \\
\geq \quad & \Pr\left(\theta < 0\right) \\
> \quad & 0\,.
\end{aligned}
$$

Since function $\max\{0, \cdot\}$ is convex, Jensen's inequality implies that

$$
\begin{aligned}
& \max\left\{0, \int_\theta [\theta - r \cdot (1 - m_i(\theta))] \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) \cdot p(\theta)\,d\theta\right\} \\
\leq \quad & \int_\theta \max\left\{0, [\theta - r \cdot (1 - m_i(\theta))] \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right)\right\} \cdot p(\theta)\,d\theta \\
(\text{C14}) \quad = \quad & \int_\theta \max\{0, [\theta - r \cdot (1 - m_i(\theta))]\} \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) \cdot p(\theta)\,d\theta\,.
\end{aligned}
$$

Since

$$
\begin{aligned}
& \Pr\left([\theta - r \cdot (1 - m_i(\theta))] \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) > 0\right) \\
= \quad & \Pr\left(\theta - r \cdot (1 - m_i(\theta)) > 0\right) \\
> \quad & 0
\end{aligned}
$$

and

$$
\begin{aligned}
& \Pr\left([\theta - r \cdot (1 - m_i(\theta))] \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) < 0\right) \\
= \quad & \Pr\left(\theta - r \cdot (1 - m_i(\theta)) < 0\right) \\
> \quad & 0
\end{aligned}
$$

for all $x_j \in \mathbb{R}$, (C14) holds strictly. Then, (C12) implies

$$
\begin{aligned}
V_j^*(\beta_j, m_i) \quad = \quad & \int_{x_j} \max\left\{0, \int_\theta [\theta - r \cdot (1 - m_i(\theta))] \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) \cdot p(\theta) \cdot d\theta\right\} \cdot dx_j \\
< \quad & \int_{x_j} \int_\theta \max\{0, [\theta - r \cdot (1 - m_i(\theta))]\} \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) \cdot p(\theta) \cdot d\theta \cdot dx_j \\
= \quad & \int_\theta \max\{0, (\theta - r \cdot (1 - m_i(\theta))) \cdot p(\theta)\} \cdot \int_{x_j} \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}(x_j - \theta)\right) \cdot dx_j \cdot d\theta \\
= \quad & \int_\theta \max\{0, (\theta - r \cdot (1 - m_i(\theta))) \cdot p(\theta)\} \cdot 1 \cdot d\theta \\
= \quad & V_j^{**}(m_i)\,,
\end{aligned}
$$

where the last equality follows (C13). Therefore,

$$
(\text{C15}) \qquad \forall m_i \in \Omega,\ \forall \beta_j > 0,\ V_j^{**}(m_i) > V_j^*(\beta_j, m_i)\,.
$$

Step 3: We prove $\lim_{n\to\infty} \beta_j(c_n) = \infty$.

If this is not true, there exists a $\overline{\beta}_j > 0$ and a subsequence $\{c_{n_k}\}_{k=1}^{\infty} \subset \{c_n\}_{n=1}^{\infty}$ s.t. $\lim_{k\to\infty} \beta_j(c_{n_k}) = \overline{\beta}_j$.

We first show $\forall \beta_j > 0, \exists \beta'_j > 0$ and $\delta(\beta_j, \beta'_j) > 0$, s.t. $\forall m_i \in M_{\overline{\beta}_i}$,

$$V_j^*(\beta'_j, m_i) - V_j^*(\beta_j, m_i) > \delta(\beta_j, \beta'_j) \ .$$

Otherwise, $\exists \beta_j > 0, \forall \beta'_j > 0, \forall l \in \mathbb{N}, \exists m^l_{\beta_j, \beta'_j} \in M_{\overline{\beta}_i}$, s.t.

$$V_j^*\left(\beta'_j, m^l_{\beta_j, \beta'_j}\right) - V_j^*\left(\overline{\beta}_j, m^l_{\beta_j, \beta'_j}\right) \leq 1/l \ .$$

Hence $\forall \beta'_j > 0$, there exists a $m_{\beta_j, \beta'_j} \in \Omega$ and a subsequence $\left\{l_{k, \beta_j, \beta'_j}\right\}_{k=1}^{\infty}$ s.t.

$$\lim_{k\to\infty} m^{l_{k, \beta_j, \beta'_j}}_{\beta'_j} = m_{\beta_j, \beta'_j}$$

and

$$V_j^*\left(\beta'_j, m_{\beta_j, \beta'_j}\right) - V_j^*\left(\beta_j, m_{\beta_j, \beta'_j}\right) \leq 0,$$

since $M_{\overline{\beta}_i}$ is relatively compact and $V_j^*(\beta, m)$ is a continuous functional of m for all $\beta > 0$. However, (C15) implies that

$$V_j^*\left(\beta'_j, m_{\beta_j, \beta'_j}\right) - V_j^*\left(\beta_j, m_{\beta_j, \beta'_j}\right) > 0$$

for β'_j large enough, which is a contradiction.

Note that $V_j^*(\beta, m)$ is continuous in β, hence $\exists \beta'_j > \overline{\beta}_j$ and $K \in \mathbb{N}$ s.t. $\forall k > K, \forall m_i \in M_{\overline{\beta}_i}$,

$$V_j^*(\beta'_j, m_i) - V_j^*(\beta_j(c_{n_k}), m_i) > \delta(\overline{\beta}_j, \beta'_j)/2 \ .$$

Since $\lim_{n\to\infty} c_{n_k} = 0$, we can choose k large enough such that

$$c_{n_k} < \frac{\delta(\overline{\beta}_j, \beta'_j)}{2 \cdot [h(\beta'_j) - h(\beta_j(c_{n_k}))]} \ .$$

Hence we have $\forall m_i \in M_{\overline{\beta}_i}$,

$$V_j^*(\beta'_j, m_i) - c_{n_k} \cdot h(\beta'_j) > V_j^*(\beta_j(c_{n_k}), m_i) - c_{n_k} \cdot h(\beta_j(c_{n_k})),$$

which contradicts the assumption that $\beta_j(c_{n_k})$ is player j's equilibrium response in $G(c_{n_k})$. Therefore we prove $\lim_{n\to\infty} \beta_j(c_n) = \infty$.

Step 4: Finally we derive a contradiction to complete the proof. Since $\forall n$, $\beta_i\left(c_n\right) \in \left[0, \overline{\beta}_i\right]$, there exists a $\beta_i^* \in \left[0, \overline{\beta}_i\right]$ and a subsequence $\left\{c_{n_k}\right\}_{k=1}^{\infty} \subset \left\{c_n\right\}_{n=1}^{\infty}$ s.t.

$$\lim_{k \to \infty} \beta_i\left(c_{n_k}\right) = \beta_i^* \, .$$

Let $m_i\left(\cdot, \beta_i\left(c\right)\right)$ characterize player i's equilibrium strategy in $G\left(c\right)$. $\forall k \in \mathbb{N}$, $\forall \theta \in \Theta$, we have

$$
m_i\left(\theta, \beta_i\left(c_{n_k}\right)\right)
$$
$$
= \int_{x_i} \left[\begin{array}{c} \left[\beta_i\left(c_{n_k}\right)\right]^{1/2} \cdot f\left(\left[\beta_i\left(c_{n_k}\right)\right]^{1/2}\left(x_i - \theta\right)\right) \\ \cdot 1_{\left\{\int_{\theta'}\left(\theta' - r \cdot\left(1 - m_j\left(\theta', \beta_j\left(c_{n_k}\right)\right)\right)\right) \cdot \left[\beta_i\left(c_{n_k}\right)\right]^{1/2} f\left(\left[\beta_i\left(c_{n_k}\right)\right]^{1/2}\left(x_i - \theta'\right)\right) p\left(\theta'\right) d\theta' > 0\right\}} \end{array} \right] \cdot dx_i
$$

and

$$
m_j\left(\theta, \beta_j\left(c_{n_k}\right)\right)
$$
$$
= \int_{x_j} \left[\begin{array}{c} \left[\beta_j\left(c_{n_k}\right)\right]^{1/2} \cdot f\left(\left[\beta_j\left(c_{n_k}\right)\right]^{1/2}\left(x_j - \theta\right)\right) \\ \cdot 1_{\left\{\int_{\theta'}\left(\theta' - r \cdot\left(1 - m_i\left(\theta', \beta_i\left(c_{n_k}\right)\right)\right)\right) \cdot \left[\beta_j\left(c_{n_k}\right)\right]^{1/2} f\left(\left[\beta_j\left(c_{n_k}\right)\right]^{1/2}\left(x_j - \theta'\right)\right) p\left(\theta'\right) d\theta' > 0\right\}} \end{array} \right] \cdot dx_j .
$$

Since Ω is a complete functional space and $m_i\left(\cdot, \beta_i\right)$ is continuous in β_i, $i \in \{1, 2\}$, there exists $\left(m_i^*\left(\cdot\right), m_j^*\left(\cdot\right)\right) \in \Omega \times \Omega$ such that $\lim_{k \to \infty}\left(m_i\left(\cdot, \beta_i\left(c_{n_k}\right)\right), m_j\left(\cdot, \beta_j\left(c_{n_k}\right)\right)\right) = \left(m_i^*\left(\cdot\right), m_j^*\left(\cdot\right)\right)$. Especially, as a result of Step 3, $\forall \theta \in \Theta$,

$$m_j^*\left(\theta\right) = 1_{\left\{\theta - r \cdot\left(1 - m_i^*\left(\theta\right)\right) > 0\right\}}$$

and

$$m_i^*\left(\theta\right) = \int_{x_i} \beta_i^{*1/2} \cdot f\left(\beta_i^{*1/2}\left(x_i - \theta\right)\right) \cdot 1_{\left\{\int_{\theta'}\left(\theta' - r \cdot\left(1 - m_j^*\left(\theta\right)\right)\right) \cdot \beta_i^{*1/2} f\left(\beta_i^{*1/2}\left(x_i - \theta'\right)\right) p\left(\theta'\right) d\theta' > 0\right\}} \cdot dx_i .$$

Choose a sequence $\left\{\beta_{i, n_k}'\right\}_{k=1}^{\infty}$ such that

$$\lim_{k \to \infty} \beta_{i, n_k}' = \infty$$

and

$$\lim_{k \to \infty} c_{n_k} \cdot h\left(\beta_{i, n_k}'\right) = 0 \, .$$

Then

$$
\lim_{k \to \infty} \left\{ \begin{array}{c} \left[V_i^*\left(\beta_i\left(c_{n_k}\right), m_j\left(\cdot, \beta_j\left(c_{n_k}\right)\right)\right) - c_{n_k} \cdot h\left(\beta_i\left(c_{n_k}\right)\right)\right] \\ - \left[V_i^*\left(\beta_{i, n_k}', m_j\left(\cdot, \beta_j\left(c_{n_k}\right)\right)\right) - c_{n_k} \cdot h\left(\beta_{i, n_k}'\right)\right] \end{array} \right\}
$$
$$
= \left[V_i^*\left(\beta_i^*, m_j^*\right) - 0\right] - \left[V_i^{**}\left(m_j^*\right) - 0\right] < 0,
$$

where the last inequality follows (C15). Therefore, for $k \in \mathbb{N}$ large enough,

$$V_i^* \left(\beta_i \left(c_{n_k} \right), m_j \left(\cdot, \beta_j \left(c_{n_k} \right) \right) \right) - c_{n_k} \cdot h \left(\beta_i \left(c_{n_k} \right) \right) < V_i^* \left(\beta'_{i,n_k}, m_j \left(\cdot, \beta_j \left(c_{n_k} \right) \right) \right) - c_{n_k} \cdot h \left(\beta'_{i,n_k} \right),$$

which contradicts to the assumption that $\beta_i \left(c_{n_k} \right)$ is player i's equilibrium response in $G \left(c_{n_k} \right)$.

This concludes the proof.

APPENDIX D

Proof of Proposition 7.

PROOF: According to Proposition 9 and 10 in Appendix B, both "always invest" (i.e., $m(\theta) = 1$ for all $\theta \in \Theta$) and "never invest" (i.e., $m(\theta) = 0$ for all $\theta \in \Theta$) are equilibria. By the way, there may also exist "intermediate" equilibria with information acquisition. This concludes the proof.

Proof of Proposition 8.

PROOF: We consider a special class of equilibria with the following form

$$m(\theta) = \left\{ \begin{array}{l} m_1 \text{ if } \theta \leq m_\lambda \\ m_0 \text{ if } \theta > m_\lambda \end{array} \right. ,$$

where $m_0, m_1 \in (0, 1)$, $m_0 < m_1$, $m_\lambda = \lambda \cdot m_0 + (1 - \lambda) \cdot m_1$ and $\lambda \in [0, 1]$.

Given $p_I \in (0, 1)$, (20) and (21) imply that

$$m_1(p_I) = \frac{\frac{p_I}{1-p_I}}{\frac{p_I}{1-p_I} + e^{-\mu^{-1}}}$$

and

$$m_0(p_I) = \frac{\frac{p_I}{1-p_I}}{\frac{p_I}{1-p_I} + e^{\mu^{-1}}} .$$

Let

$$g(p_I) = \frac{1}{1 + 2A} \left[m_1 \cdot (m_\lambda + A) + m_0 \cdot (1 + A - m_\lambda) \right].$$

If

$$g(p_I) = p_I ,$$

according to (20), (21) and (22), $m_1(p_I)$, $m_0(p_I)$,

$$S_I = \left[-A, \lambda \cdot m_0(p_I) + (1 - \lambda) \cdot m_1(p_I) \right]$$

and

$$S_N = (\lambda \cdot m_0(p_I) + (1 - \lambda) \cdot m_1(p_I), 1 + A]$$

constitute an equilibrium. Now we prove $g(p_I)$ has an fixed point in $(0, 1)$ for any $\lambda \in [0, 1]$.

Step 1. We show

$$g\left(p_I\right) < p_I$$

for p_I close to 1.

Since $m_\lambda \leq 1$,

$$
\begin{aligned}
g\left(p_I\right) &= \frac{1}{1+2A}\left[m_1\left(p_I\right)\cdot\left(m_\lambda + A\right) + m_0\left(p_I\right)\cdot\left(1 + A - m_\lambda\right)\right] \\
&\leq \frac{1+A}{1+2A}\cdot m_1\left(p_I\right) + \frac{A}{1+2A}\cdot m_0\left(p_I\right) .
\end{aligned}
$$

Hence it suffices to show

(D1)
$$\frac{1+A}{1+2A}\cdot m_1\left(p_I\right) + \frac{A}{1+2A}\cdot m_0\left(p_I\right) < p_I$$

for p_I close to 1. Let

$$w = \frac{1}{1-p_I} ,$$

$$v_0 = e^{-\mu^{-1}} - 1$$

and

$$v_1 = e^{\mu^{-1}} - 1 ,$$

then (D1) can be rewritten as

(D2)
$$\frac{1+A}{1+2A}\cdot\frac{w-1}{w+v_0} + \frac{A}{1+2A}\cdot\frac{w-1}{w+v_1} < \frac{w-1}{w} .$$

Since $p_I \in (0, 1)$, $w - 1 > 0$ and (D2) is equivalent to

(D3)
$$\frac{1+A}{1+2A}\cdot\frac{1}{1+v_0/w} + \frac{A}{1+2A}\cdot\frac{1}{1+v_1/w} < 1 .$$

It suffices to show (D3) for w large enough.

Choosing w large enough such that

$$1 + v_0/w > 0$$

and multiplying both sides of (D3) with $\left(1 + v_0/w\right)\left(1 + v_1/w\right)$ lead to

$$1 + \frac{1+A}{1+2A}\cdot v_1/w + \frac{A}{1+2A}\cdot v_0/w < 1 + v_1/w + v_0/w + \frac{v_0 v_1}{w^2} ,$$

i.e.,

(D4)
$$-\frac{v_0 v_1}{w} < \frac{A}{1+2A}\cdot v_1 + \frac{1+A}{1+2A}\cdot v_0 .$$

Note that (18) implies

$$\frac{A}{1 + 2A} \cdot v_1 + \frac{1 + A}{1 + 2A} \cdot v_0 > 0 \, .$$

Hence, we can choose w large enough such that (D4) holds. Therefore, let

$$\overline{p}_I = \frac{w - 1}{w} \, ,$$

we have

$$g\left(\overline{p}_I\right) < \overline{p}_I \, .$$

Step 2. We show

$$g\left(p_I\right) > p_I$$

for p_I close to 0.

Since $m_\lambda \geq 0$,

$$\begin{aligned} g\left(p_I\right) & = \frac{1}{1 + 2A}\left[m_1\left(p_I\right) \cdot \left(m_\lambda + A\right) + m_0\left(p_I\right) \cdot \left(1 + A - m_\lambda\right)\right] \\ & \geq \frac{A}{1 + 2A} \cdot m_1\left(p_I\right) + \frac{1 + A}{1 + 2A} \cdot m_0\left(p_I\right) \, . \end{aligned}$$

Hence it suffices to show

$$\frac{A}{1 + 2A} \cdot m_1\left(p_I\right) + \frac{1 + A}{1 + 2A} \cdot m_0\left(p_I\right) > p_I \, ,$$

i.e.,

(D5) $$\frac{A}{1 + 2A} \cdot \left[1 - m_1\left(p_I\right)\right] + \frac{1 + A}{1 + 2A} \cdot \left[1 - m_0\left(p_I\right)\right] < 1 - p_I \, ,$$

for p_I close to 0. Let

$$w = \frac{1}{p_I} \, ,$$

$$v_0 = e^{-\mu^{-1}} - 1$$

and

$$v_1 = e^{\mu^{-1}} - 1 \, ,$$

then (D5) can be rewritten as

(D6) $$\frac{A}{1 + 2A} \cdot \frac{w - 1}{w + v_1} + \frac{1 + A}{1 + 2A} \cdot \frac{w - 1}{w + v_0} < \frac{w - 1}{w} \, .$$

It suffices to show (D6) for w large enough. Note that (D6) is the same as (D2), thus by

the same argument in Step 1, we can find a $\underline{p}_I$ close to 0 such that

$$g\left(\underline{p}_I\right) > \underline{p}_I \ .$$

Step 3. Since $g\left(p_I\right)$ is continuous in p_I, there must exist a $p_I^* \in (0, 1)$ such that

$$g\left(p_I^*\right) = p_I^* \ .$$

Hence we find an equilibrium for the given λ. Since we can find an equilibrium for any $\lambda \in [0, 1]$, there exist infinitely many equilibria. This concludes the proof.

Proof of Proposition 6.

PROOF: Since information acquisition is flexible, there exists $K > 0$ such that

$$\forall m, \widetilde{m} \in \Omega, \ |c\left(m\right) - c\left(\widetilde{m}\right)| \leq K \cdot \int |m\left(\theta\right) - \widetilde{m}\left(\theta\right)| \, dP\left(\theta\right) \ ,$$

where P is the common prior that is mutually continuous with respect to Lebesgue measure on $\mathbb{R}$.

Let $V_i\left(m_i, m_j\right)$ denote player i's expected payoff from playing strategy m_i when player j plays m_j, i.e.,

$$V_i\left(m_i, m_j\right) = \int \left[\theta - r \cdot \left(1 - m_j\left(\theta\right)\right)\right] \cdot m_i\left(\theta\right) dP\left(\theta\right) - \mu \cdot c\left(m_i\right) \ .$$

Note that any strict Nash equilibrium of the game with complete information (except the two extreme ones) involves both players playing a strategy parameterized by $x \in (0, r)$, i.e., $1_{\{\theta > x\}}$. (Symmetrically, we should also consider $1_{\{\theta \geq x\}}$, which is omitted here.)

Let $m_i \in \Omega$ and $\int \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta) > 0$. Then

$$V_i \left(1_{\{\theta > x\}}, 1_{\{\theta > x\}} \right) - V_i \left(m_i, 1_{\{\theta > x\}} \right)$$

$$= \int \left[\theta - r \cdot \left(1 - 1_{\{\theta > x\}} \right) \right] \cdot \left[1_{\{\theta > x\}} - m_i(\theta) \right] dP(\theta) - \mu \cdot \left[c \left(1_{\{\theta > x\}} \right) - c(m_i) \right]$$

$$= \int_{\theta \leq x} [r - \theta] \cdot m_i(\theta) \, dP(\theta) + \int_{\theta > x} \theta \cdot [1 - m_i(\theta)] \, dP(\theta)$$
$$- \mu \cdot \left[c \left(1_{\{\theta > x\}} \right) - c(m_i) \right]$$

$$\geq \int_{\theta \leq x} [r - \theta] \cdot m_i(\theta) \, dP(\theta) + \int_{\theta > x} \theta \cdot [1 - m_i(\theta)] \, dP(\theta) - \mu \cdot K \cdot \int \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta)$$

$$= \int_{\theta \leq x} [r - \theta] \cdot \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta) + \int_{\theta > x} \theta \cdot \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta)$$

$$- \mu \cdot K \cdot \int \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta)$$

$$= \int [t(\theta) - \mu \cdot K] \cdot \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta) \, ,$$

Where

$$t(\theta) \triangleq \begin{cases} r - \theta & \text{if } \theta \leq x \\ \theta & \text{if } \theta > x \end{cases}.$$

Note that

$$t(\theta) \geq \min(x, r - x) > 0$$

for all $\theta \in \Theta$. Let $\overline{\mu} < [\min(x, r - x)] / K$, we have

$$V_i \left(1_{\{\theta > x\}}, 1_{\{\theta > x\}} \right) - V_i \left(m_i, 1_{\{\theta > x\}} \right) > 0$$

for all $m_i \in \Omega$ s.t. $\int \left| 1_{\{\theta > x\}} - m_i(\theta) \right| dP(\theta) > 0$ and all $\mu \in [0, \overline{\mu}]$. Therefore, $1_{\{\theta > x\}}$ is a strict Nash equilibrium of the game with information acquisition.