GAMES WITH RATIONAL INATTENTION—COORDINATION WITH ENDOGENOUS INFORMATION

Ming Yang, Dept. of Economics, Princeton University

this version: May 10, 2010

Abstract

The equilibria of a coordination game with incomplete information depend on its information structure, i.e. the joint distribution of signals and fundamentals. Rather than exogenously assuming an information structure like most models in the literature, we allow the players to acquire information according to their own interests. The information structure then emerges as a part of the equilibrium rather than results in it. This setup avoids the arbitrariness in choosing information structure. The players' information acquisition behavior is modeled by rational inattention, a theory stating that human beings have limited capacity for information processing/acquisition and can optimally use it subject to such capacity constraint. Rational inattention has a solid foundation built upon Shannon's information theory. It frees the model from the behavioral details of human beings' information acquisition and thus is flexible enough to provide a general framework for the analysis of endogenous information acquisition. We show that MLRP (Monotonic Likelihood Ratio Property) holds if ratio of strategic complementarity to the marginal cost of information acquisition is no greater than unit and may not hold otherwise. Our model also generates some results distinct from most global games with exogenous information. For example, we show that lowering cost of information acquisition leads to multiplicity, which is opposite to the implication of a well known result that increasing the accuracy of private information facilitates uniqueness. We show that all the distinctions come from the difference between the flexible information structure of our model and the rigidity imposed on the previous ones.

*I am grateful to Stephen Morris for his guidance and comments. I also thank Marco Battaglini, Christopher Sims and Jorgen Weibull for their comments. I would like to thank Chi Li, Xin Wan and Yi Wang for the helpful discussion with them.

1 INTRODUCTION

Coordination is broadly oberserved in human activities and is often modeled by coordination games. Generally, in a coordination game with complete information, multiple equilibria emerge and undermine the prediction power of the model. As introduced by Carlsson and van Damme (1993), the approach of global games appears as a natural refinement to get rid of such multiplicity. A global game is an incomplete information game where the players are not sure about the payoff structure but can make inference according to their privately observed signals. This uncertainty about the payoff structure and the other players' believes weakens the common knowledge and thus facilitates the uniqueness. Regaining the prediction power, global game models are widely used in the study of currency attacks (Morris and Shin (1998)), debt pricing (Morris and Shin (2004)) and bank runs (Goldstein and Pauzner (2005)), etc. A common feature of these models is the exogeneity of their information structure, which determines the equilibrium outcomes. While bringing back the prediction power, however, this approach also brings in the arbrtrariness in choosing information structure, which makes the equilibrium outcomes manipulable. Moreover, the information structure is usually assumed to satisfy the Monotonic Likelihood Ratio Property (MLRP thereafter). Then a series of questions arise: where does MLRP come from? Is it always valid? If not, under what condition is it reasonable? These questions cannot be answered within a model having exogenous information structure since MLRP itself is a part of the information structure. To avoid this arbitrariness and justify MLRP, we study a global game model with endogenous information structure.

Two players coordinate in investing a risky project. The project's fundamental (profitability) is distributed according to a common prior. Each player's utility from investing depends on the uncertain fundamental as well as his opponents' actions. The players' actions are assumed to be strategic complements. To endogenize the information structure, we allow the players to actively acquire information according to their own interests rather than passively respond to the given signals. Specifically, a player can choose the joint distribution of his signals and the fundamental[1] and makes decision upon the realization of the signals. It is not surprising that the players would like to establish a ono-to-one mapping between the fundamental and their signals if there is no other constraints. This is equivalent to allow the players to observe the exact value of the fundamental and reduces to the games with complete information. This setup not only makes our problem a trivial one but is also unrealistic in many situations. Thus we need to add some constraints on the players' information acquisition behavior. Here we find rational inattention a proper modeling technique.

[1]It is equivalent to choose a conditional distribution of his signal on the fundamental since the marginal distribution of the fundamental is just the common prior.

The theory of rational inattention has two basic assumptions: i) the players' capacity for information processing/acquisition is limited; ii) facing this limit, the players can optimally use their capacity to acquire/process information according to their objectives. Here information is measured by the reduction of Shannon entropy of the fundamental. The first assumption says that there exists an upper bound for this entropy reduction. According to Shannon's information theory, this assumption is natural since nothing in the world is able to process infinte information within a given time period. In practice, this capacity upper bound may come from the limited computational power or the fixed number of analysts, etc.The second assumption may not be always realistic but serves as a benchmark. It can be viewed as the usual "rational man" assumption in this endogenous information acquisition context. Rational inattention imposes the least restrictions on the possible information structures and is hence consistent with our purpose of avoiding the arbitrariness. Having a solid fundation built upon information theory, rational inattention frees us from the behavioral details of information acquisition and thus has the potential to provide a general framework for analyzing endogenous information acquisition problems.

Rational inattention is first introduced by Sims (1998) to model the price stickiness. Sims (2003, 2006) further develop the theory. Mackowiak and Wiederholt (2009) examines its effects in a dynamic stochastic general equilibrium framework. While these macroeconomic models focus on the single person (the representative agent) decision problem, our model studies its effects in a strategic interactive environment. This is not merely another application of rational inattention as some new insight emerges from the combination of rational inattention and strategic games. The strategic complementarity between the players' actions induces the strategic complementarity between their information acquisition behavior, which leads to multiplicity. This result is similar to the main result of Hellwig and Veldkamp (2009)[2]. Moreover, if we set the parameter of strategic complementarity to zero, multiplicity disappears and our model reduces to a single person decision problem.

In our (capacity) constrained information acquisition problem, we fully characterize all the equilibria. We find that i) all equilibria are symmetric; ii) the equilibria can be divided into three types: pooling equilibria, perfect separating equilibria and partial separating equilibria. In the first type of equilibria the players pool in the same action and does not use their capacity to process information. In the second type of equilibria the players partition the possible states of the world into two positive-probability events corresponding to "invest" and "not invest", repectively. Their capacity is large enough and they know exactly which event happens. In

[2] Hellwig and Veldkamp (2009) also study the endogenous information acqusition in a strategic interactive environment, but they do not use rational inattention as a modeling technique.

the third type of equilibria the players' capacity constraint binds and he can not perfectly distinguish the above two events. The pooling equilibria and the perfect separating equilibria are two extreme cases where capacity constraint is slack, either because there is no need to acquire information or because the player's capacity is sufficiently large. In the main part of the paper, we provide necessary and sufficient conditions for the existence of these equilibria. We find that MLRP holds if the players' capacity is low enough but may not hold if otherwise.

When solving the constrained information acquisition problem, the Lagrangian multiplier for the capacity constraint is actually the shadow price of the capacity. Suppose $\mu = \mu(\kappa) > 0$ is the Lagrangian multiplier in an equilibrium and the corresponding capacity endowment is κ. If we remove the capacity constraint but allow the players to acquire information at a marginal cost $\mu(\kappa)$, we will end up with the same equilibrium[3] where each player acquires κ bits of information. We call this new problem the costly information acquisition problem, in which no capacity constraint is imposed but the players have to acquire information at some marginal cost. In practice, this marginal cost can be the expenditure of purchasing another computer or hiring one more analyst, etc. These two problems are closely related and their results are similar except that the latter one only has two types of equilibria: pooling equilibria and partial separating equilibria. In the main part of the paper, we first present the constrained information acquisition problem to provide a full characterization of the possible equilibria in this endogenous information acquisition environment and then use the costly information acquisition problem to further our analysis since it is easier to deal with mathematcially.

We provide a clear condition to justify MLRP. MLRP must hold if the ratio of the strategic complementarity to the marginal cost of information acquisition is less than unit and may not hold if otherwise. The strategic complementarity reflects the players' motive of coordination and the marginal cost measures the difficulty to do so. Thus we can define their ratio as the effective strategic complementarity of the game. Intuitively, MLRP is incentive compatible since it makes the players more likely to invest if it is more probable to be in a good state. However, if the effective strategic complementarity is too high, there may exist excessive coordination since a player now has enough incentive and ability to coordinate with his opponent's (irregular) non-MLRP strategy.

We also compare the implications of the current model to that of the global games literature. We show that large effective strategic complementarity leads to multiplicity (under some regularity conditions). In other words, if the cost of information acquisition is low enough, the players have multiple ways of coordinating in acquiring information. A well known result in global games literature is that increasing the precision of private information facilitates unique-

[3] There might be other equilibria.

ness. We extend the standard global game model by allowing the players to buy the precision of their private signals at some cost and show that they would like to buy infinitely precise signals if such cost approaches zero. Thus the traditional model says that lowering the cost of information acquisition favors the uniqueness rather than the multiplicity. Hence the current model and the previous models have totally opposite predictions about the effect of the cost of information acquisition.

Another well known result in the literature is that the uniqueness is guaranteed if the private signals are sufficiently accurate relative to the public signals (e.g. Morris and Shin (2004)). In other words, the effects on the uniqueness of increasing the precision of public signals can be offset through increasing the precision of private signals. For the extended standard global game model mentioned in last paragragh, we show that this effect of increasing the precision of public signals can be also offset by lowering the cost of private information acquisition. In our endogenous information model, however, under some regularity conditions there are always infinitely many equilibria when the effective strategic complementarity is large, regardless of the precision of public information, i.e. the effects of public information and private information acquisiiton are disentangled. The reason is that when the cost of information acquisition is samll, the players have enough freedom in allocating their attention, which in turn improves their coordination. We show that all these distinctions come from the difference between the flexible information structure of the current model and the rigidity imposed on the previous ones.

Our model also generates some results consistent with the previous global game models, e.g. we show that providing public information of high precision leads to multiplicity.

I proceed as following. Section II sets up the benchmark model and show some simple facts about the equilibrium information acquisition behavior. Section III studies the (capacity) constrained informaiton acqusition problem and fully characterizes the possible equilibria in the environment of endogenous information acquisition. Section IV addresses the costly information acquisition problem, further analyzes the effects of public information and cost of private information acquisition, and compare the implications of our model to previous global game models. We conclude in Section V with several directions for further research. Appendix A provides some basic knowledge of information theory and rational inattention. Appendix B is a collection of technical proofs.

2 THE BENCHMARK MODEL

2.1 The Basic Environment

Two players[4] play a coordination game with payoffs shown below

	I	N
I	(θ,θ)	$(\theta-r,0)$
N	$(0,\theta-r)$	$(0,0)$

(2.1)

where $\{I,N\}$ is the action set standing for $\{invest, not\ invest\}$, θ is the fundamental that represents the profitability of the project and $r>0$ measures the degree of strategic complementarity[5].

Let cummulative distribution function $P(\cdot)$ denote the common prior about the fundamental θ. $P(\cdot)$ can contain both absolutely continuous component and discrete components.[6] To avoid the trivial case where there is no uncertainty, we assume that supp$(P(\cdot))$ has strictly more than one point. This is called the non-triviality assumption thereafter.

Each player first independently collect information about the fundamental. Specifically, player $i\in\{1,2\}$ chooses a conditional density function $q^i\left(s^i|\theta\right)$ for his private signal s^i. To make it concrete, let $s^i\in S^i\in\mathbb{R}$.[7] These signals are private in the sense that they are independent when conditioned on θ. The conditional independence assumption models the players' independent information collecting behavior. The players then take actions after observing the realization of their private signals. Thus player i's strategy can be characterized by $\left(q^i(\cdot|\theta),\sigma^i(\cdot)\right)$, where $q^i:\mathbb{R}\rightarrow\Delta\left(S^i\right)$ determines the information being collected and $\sigma^i:S^i\rightarrow[0,1]$ is a mapping from possible realizations of i's private signal to the probability of choosing I.

The conditional density function $q^i(\cdot|\theta)$ describes player i's information collecting behavior. By choosing different functional forms for $q^i\left(s^i|\theta\right)$, player i can make his signal covariate with

[4] Here the "two-player" setup is not as restrictive as it seems at the first glance, since all our results also hold in the setup of a continuum of players if we slightly change the payoff for "invest" to $\theta-r\cdot(1-m)$, where m is the fraction of the players that invest.

[5] The more general case with $r(\theta)>0$ $a.s.$ can also be studied in our framework and many key results still hold. Here we assume constant strategic complementarity for the sake of simplicity.

[6] Note that the public signals affect the common prior, thus their effects on the equilibria can be studied through the comparative static analysis with respect to the common prior. The effects of the public signals are studied in the latter part of the paper.

[7] Here it is not essential to assume $S^i\in\mathbb{R}$. The realizations of the signal can belong to an arbitrary abstract space. All our results hold regardless of this assumption.

the fundamental in any way he would like. For example, if player i's welfare is sensitive to the fluctuation of the fundamental within some range $A \subset \text{supp}(P(\cdot))$, he would pay much attention to this event by making s^i highly correlated to $\theta \in A$. Generally, we use $I\left(s^i;\theta\right)$, the mutual information between the two random variables s^i and θ, to measure the amount of information about θ that is contained in s^i. $I\left(s^i;\theta\right)$ is uniquely determined by the functional form of $q^i\left(s^i|\theta\right)$ but not vice versa. A functional form of $q^i\left(s^i|\theta\right)$ defines a specific way of information collecting, or in other words, it determines what information about θ to be collected. Different forms of $q^i\left(s^i|\theta\right)$ may generate the same value for $I\left(s^i;\theta\right)$, i.e. the same amount of information may be collected from different aspects of θ.[8] Intuitively, if player i is allowed to choose the conditional distribution without any constraint, he would like to establish a one-to-one mapping between his signal and θ and thus obtain all the information of the fundamental. This makes our problem a trivial one since it is just a coordiantion game with complete information. Besides the triviality, this arbitrariness in choosing the functional form of $q^i\left(s^i|\theta\right)$ is also unrealistic. The mutual information between θ and s^i represents player i's capacity of information processing/collecting and thus must have an upper bound[9], i.e. $\exists\kappa > 0$ s.t. $I\left(s^i;\theta\right) \leq \kappa$. If this upper bound is strictly less then the uncertainty (i.e. Shannon entropy) of the fundamental, no one-to-one mapping can exist between θ and s^i. This is the case especially when the common prior $P(\cdot)$ has a continuous component, which leads to an infinite Shannon entropy of θ. Therefore, we impose the capacity constraint on the players' information acquisition behavior:

Assumption (A1): player $i \in \{1,2\}$ can choose any conditional density function $q^i\left(s^i|\theta\right) \in Q^i_\kappa$, where $\kappa > 0$ is a constant that measures the players' capacity of information processing, $Q^i \triangleq \left\{q : \mathbb{R} \to \Delta\left(S^i\right)\right\}$ and $Q^i_\kappa \triangleq \left\{q \in Q^i : I\left(s^i;\theta\right) \leq \kappa\right\}$.

Note that $\cup_{\kappa\in\mathbb{R}_+} Q^i_\kappa = Q^i$.

The objective of player $i \in \{1,2\}$ is to utilize his capacity to maximize his expected utility. When κ is large enough, player i may have multiple methods of information acquisition that lead to the same maximized expected utility. To avoid this trivial multiplicity and make our analysis brief and clear, we also assume that the players do not collect the information that is never used. We summarize the players' preferences in the assumption below:

Assumption (A2): i) the players prefer higher expected utility, where for any pair of strategies

[8] For a better understanding, please read Appendix A for some basics of information theory and rational inattention.

[9] As a general principle, nothing in the world can process infinite information within a given period.

$\left(\left(q^1\left(\cdot|\theta\right),\sigma^2\left(\cdot\right)\right),\left(q^2\left(\cdot|\theta\right),\sigma^2\left(\cdot\right)\right)\right)$, the expected utility of player $i\in\{1,2\}$ is

$$\begin{aligned} u_i &= u_i\left(q^1\left(\cdot|\theta\right),\sigma^2\left(\cdot\right),q^2\left(\cdot|\theta\right),\sigma^2\left(\cdot\right)\right) \\ &= \int_\theta\int_{s^i}\sigma^i\left(s^i\right)\cdot\int_{s^j}\left[\sigma^j\left(s^j\right)\cdot\theta+\left(1-\sigma^j\left(s^j\right)\right)\cdot\left(\theta-r\right)\right]\cdot q^j\left(s^j|\theta\right)\cdot ds^j\cdot q^i\left(s^i|\theta\right)\cdot ds^i\cdot dP\left(\theta\right) \\ &= \int_\theta\int_{s^i}\int_{s^j}\sigma^i\left(s^i\right)\cdot\left[\theta-r\cdot\left(1-\sigma^j\left(s^j\right)\right)\right]\cdot q^i\left(s^i|\theta\right)\cdot q^j\left(s^j|\theta\right)\cdot ds^j\cdot ds^i\cdot dP\left(\theta\right) \end{aligned}$$

ii) if two feasible strategies generate the same expected utility, the player prefers the one with less mutual information.

A constrained information acquisition problem $G\left(r,\kappa\right)$ can be stated as: two players with preference (A2) play the game with payoff matrix shown in (2.1). Player $i\in\{1,2\}$ chooses strategy $\left(q_i\left(s^i|\theta\right),\sigma^i\left(\cdot\right)\right)$ subject to the capacity constraint (A1), where $\sigma^i:S^i\to[0,1]$ is a mapping from possible realizations of i's private signal to the probability of choosing I. The equilibrium concept is Bayesian Nash equilibrium.

In principle, this problem seems hard to deal with, since the players' possible choices belong to a functional space and even S^1 and S^2, the sets of the possible realizations of the private signals are endogenous. However, some patterns emerge from the players' optimization behavior and next subsection provides a first simplification of our problem.

2.2 Some Simple Facts About The Equilibria Of The Constrained Information Acquisition Problem

Define $S^i=\cup_{\theta\in\mathrm{supp}(P(\cdot))}\left(\mathrm{supp}\left(q_i\left(\cdot|\theta\right)\right)\right)$, $i\in\{1,2\}$.

Let $S^i_I=\left\{s^i\in S^i:\sigma^i\left(s^i\right)=1\right\}$, $S^i_N=\left\{s^i\in S^i:\sigma^i\left(s^i\right)=0\right\}$ and $S^i_{ind}=\left\{s^i\in S^i:\sigma^i\left(s^i\right)\in(0,1)\right\}$. Then $\left(S^i_I,S^i_N,S^i_{ind}\right)$ is a partition of S^i and $\Pr\left(S^i_I\right)+\Pr\left(S^i_N\right)+\Pr\left(S^i_{ind}\right)=1$.

Lemma 01 *in the equilibrium of the constrained information acquisition problem,* $\#\left(S^i\right)=1$ *or* 2*, and* $\Pr\left(S^i_{ind}\right)=0$*,* $\forall i\in\{1,2\}$*.*

Proof. see Appendix B. ■

The intuition behind Lemma 01 is that player i has no incentive to distinguish different realizations within any set of S^i_I, S^i_N and S^i_{ind}, since this effort generates no utility but incurs cost of information collecting. It also suggests that the mixed strategies are not played in equilibria. Since player i is indifferent between I and N when event S^i_{ind} happens, he would pay no attention to distinguish it from other realizations. Thus there is no need for S^i_{ind} to exist and player i always plays the pure strategy upon receiving his signal.

Let $m_i(\theta) \triangleq \Pr(\text{player } i \text{ chooses } I|\theta)$ be the probability that player $i \in \{1,2\}$ invests when the true state of the fundamental is θ. Then $m_i(\cdot)$ is totally determined by player i's strategy $\left(q_i\left(s^i|\theta\right), \sigma^i(\cdot)\right)$, $i \in \{1,2\}$. On the other hand, Lemma 01 implies that player i's strategy $\left(q_i\left(s^i|\theta\right), \sigma^i(\cdot)\right)$ is also totally determined by $m_i(\cdot)$. Specifically, when $\Pr(m_i(\theta)=1)=1$ ($\Pr(m_i(\theta)=0)=1$), let $S^i=\left\{s_I^i\right\}$, $\forall\theta \in$supp$(P(\cdot))$, $q_i\left(s_I^i|\theta\right)=1$ and $\sigma^i\left(s_I^i\right)=I$ ($S^i=\left\{s_N^i\right\}$, $\forall\theta \in$supp$(P(\cdot))$, $q_i\left(s_N^i|\theta\right)=1$ and $\sigma^i\left(s_N^i\right)=N$); otherwise, let $S^i=\left\{s_I^i, s_N^i\right\}$, $\forall\theta \in$supp$(P(\cdot))$, $q_i\left(s_I^i|\theta\right)=m_i(\theta)$, $q_i\left(s_N^i|\theta\right)=1-m_i(\theta)$, $\sigma^i\left(s_I^i\right)=I$ and $\sigma^i\left(s_N^i\right)=N$. In other words, $m_i(\cdot)$ is a sufficient statistic of player i's equilibrium strategy and we can directly focus on the pair $(m_1(\cdot), m_2(\cdot))$ when studying the equilibria. Let $a^i=I$ or N be a generic action of player i, then $m_i(\cdot)$ determines the conditional distribution of a^i as well as the mutual information $I\left(a^i;\theta\right)$ between player i's action and the fundamental. According to Lemma1, the capacity constraint becomes $I\left(a^i;\theta\right) \leq \kappa$ and player i's objective is to maximize his expected utility through choosing $m_i(\cdot)$ subject to this capacity constraint.

In principle, the players can choose any function $m(\cdot)$ from some abstract functional space Ω. To make our analysis precise and rigorous, however, we need to require Ω to satisfy some regularity conditions.

Let $P_1(\cdot)$ and $P_2(\cdot)$ denote the absolutely continuous and the discrete components of the common prior $P(\cdot)$, respectively. Let $L(\mathbb{R},P(\cdot)) = \left\{f \mid \int |f(\theta)| \cdot dP(\theta) < \infty\right\}$ and define a distance $\rho(\cdot,\cdot)$ on $L(\mathbb{R},P(\cdot))$ as $\rho(m_1,m_2)=\int |m_1(\theta)-m_2(\theta)| \cdot dP(\theta)$, $\forall m_1, m_2 \in L(\mathbb{R},P(\cdot))$. $\forall m \in L(\mathbb{R},P(\cdot))$, $\forall\delta>0$, define the vibration function of m as

$$w_{m,\delta}(\theta) = \inf_{\{F\subset\mathbb{R} \mid \Pr(F)=0\}} \sup_{\theta_1,\theta_2\in[\theta-\delta/2,\theta+\delta/2]\cap\text{supp}(P_1(\cdot))\setminus F} |m(\theta_2)-m(\theta_1)|$$

. Then let $\Omega \subset L(\mathbb{R},P(\cdot))$ be a closed set of uniformly bounded functions and satisfy the following property:

Property A: if $P_1(\cdot)$ is the absolutely continuous component of the common prior $P(\cdot)$, then $\forall\epsilon>0$, $\exists\delta(\epsilon)>0$, s.t. $\forall m \in \Omega$, $\int w_{m,\delta(\epsilon)}(\theta) \cdot dP_1(\theta) < \epsilon$.

Property A is actually a "uniformly integrable" condition, it makes the possible strategy space Ω a convext and compact functional space. This assumption is not as restrictive as it looks like, at least in the sense that all the equilibria solved out later satisfy Property A. It actually excludes those strategies that vibrate so wildly with non-zero probability. Since these ill-conditioned strategies are impossible to belong to any equilibrium, our assumption here is appropriate.

3 THE GENERAL APPROACH TO THE CONSTRAINED INFORMATION ACQUISITION PROBLEM

We first formalize the constrained information acquisition problem and establish the existence of the equilibria. Then we fully characterize the three types of possible equilibria and provide the necessary and sufficient conditions for each type.

Given a pair of strategies $(m_1(\cdot), m_2(\cdot))$, $\forall i, j \in \{1,2\}, i \neq j$, player i's expected utility is

$$u_i(m_i(\cdot), m_j(\cdot)) = \int m_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta) \qquad (3.1)$$

. Player i's objective is to maximize his expected utility given by (3.1) subject to the capacity constraint $I(a^i; \theta) \leq \kappa$. Note that $I(a^i; \theta)$, the mutual information between player i's action and the fundamental is actually a functional of $\widetilde{m}_i(\cdot)$. To make it clear, we use $I(\widetilde{m}_i(\cdot))$ instead of $I(a^i; \theta)$ thereafter.

An equilibrium of the constrained information acquisition problem is a pair $(m_1(\cdot), m_2(\cdot))$ solving the the following problem:

$$i, j \in \{1,2\}, \, i \neq j, \;\; m_i(\cdot) \in \arg\max_{\widetilde{m}_i(\cdot)} u_i(\widetilde{m}_i(\cdot), m_j(\cdot)) = \int \widetilde{m}_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta) \qquad (3.2)$$

$$\begin{aligned} \text{s.t. } \widetilde{m}_i(\theta) &\in \Omega \\ \widetilde{m}_i(\theta) &\in [0,1] \quad (3.3) \\ \text{and} \quad \kappa &\geq I(\widetilde{m}_i(\cdot)) \\ &= \int [\widetilde{m}_i(\theta) \ln \widetilde{m}_i(\theta) + (1 - \widetilde{m}_i(\theta)) \ln (1 - \widetilde{m}_i(\theta))] \, dP(\theta) \\ &\quad - \widetilde{p}_{Ii} \ln \widetilde{p}_{Ii} - (1 - \widetilde{p}_{Ii}) \ln (1 - \widetilde{p}_{Ii}) \qquad (3.4) \end{aligned}$$

where $\widetilde{p}_{Ii} = \Pr(\text{player } i \text{ chooses } I) = \int \widetilde{m}_i(\theta) \cdot dP(\theta)$.

Proposition 1 *the Nash equlibrium of the constrained information acquisition problem exists.*

Proof. let $\Omega_\kappa = \{\widetilde{m}_i(\cdot) \in \Omega : \forall \theta \in \operatorname{supp}(P(\cdot)), \, \widetilde{m}_i(\theta) \in [0,1] \text{ and } I(\widetilde{m}_i(\cdot)) \leq \kappa\}$. Since $I(\widetilde{m}_i(\cdot))$ is a continuous functional of $\widetilde{m}_i(\cdot)$ and is always non-negative (a property of mutual information), Ω_κ is a closed subset of Ω. Note that $L^1(\mathbb{R})$ is Hausdorff and so is its subset Ω, thus Ω_κ is a closed subset of a compact Hausdorff space. This implies that Ω_κ is also compact. On the other hand, as a property of mutual information, $I(\widetilde{m}_i(\cdot))$ is a convext functional of $\widetilde{m}_i(\cdot)$, therefore Ω_κ is convext and compact. Player i's expected utility $u_i(m_i(\cdot), m_j(\cdot))$ is a continuous and linear functional with respect to his strategy $m_i(\cdot)$, thus it is also quasi-concave. According to Nash equilibrium theorem, the Nash equilibrium exists. ■

Since Ω_κ is convex and compact and player i's expected utility $u_i(\widetilde{m}_i(\cdot), m_j(\cdot)) = \int \widetilde{m}_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta)$ is a continuous and quasi-concave functional with respect to $\widetilde{m}_i(\cdot)$, we can solve the equilibrium by Lagrangian method. Let $\lambda_i(\theta) \geq 0$, $\eta_i(\theta) \geq 0$ and $\mu_i \geq 0$ be the multipliers for $\widetilde{m}_i(\theta) \leq 1$, $\widetilde{m}_i(\theta) \geq 0$ and the capacity constraint $\kappa \geq I(\widetilde{m}_i(\cdot))$, repectively. Player i's Lagrangian is

$$\begin{aligned} L_i = & \int \widetilde{m}_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta) + \mu_i \cdot \kappa + \mu_i \cdot [\widetilde{p}_{Ii} \cdot \ln \widetilde{p}_{Ii} - (1 - \widetilde{p}_{Ii}) \cdot \ln(1 - \widetilde{p}_{Ii})] \\ & - \mu_i \cdot \int [\widetilde{m}_i(\theta) \ln \widetilde{m}_i(\theta) + (1 - \widetilde{m}_i(\theta)) \ln(1 - \widetilde{m}_i(\theta))] \cdot dP(\theta) \\ & + \int [\lambda_i(\theta)[1 - \widetilde{m}_i(\theta)] + \eta_i(\theta) \widetilde{m}_i(\theta)] \cdot dP(\theta) \end{aligned} \tag{3.5}$$

where

$$\widetilde{p}_{Ii} = \Pr(\text{player } i \text{ chooses } I) = \int \widetilde{m}_i(\theta) \cdot dP(\theta) \qquad i \in \{1, 2\} \quad (3.6)$$

The first order condition implies

$$\begin{aligned} \forall \theta & \in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m_j(\theta)) & = \mu_i \cdot \left[\ln\left(\frac{m_i(\theta)}{1 - m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right) \right] + \lambda_i(\theta) - \eta_i(\theta) \quad i, j \in \{1, 2\}, i \neq j \end{aligned} \tag{3.7}$$

Then an equilibrium of the constrained information acquisition problem is a pair $(m_1(\cdot), m_2(\cdot)) \in \Omega_\kappa \times \Omega_\kappa$ satisfying (3.7).

Generally, three types of strategies are possible in this problem.

The first type is the pooling strategy where $\Pr(m(\theta) = 1) = 1$ or $\Pr(m(\theta) = 0) = 1$.

The second type is called the perfect separating strategy where $\Pr(m(\theta) = 1) + \Pr(m(\theta) = 0) = 1$ and $\Pr(m(\theta) = 1) \in (0, 1)$.

The third type is called the partial separating strategy where $\Pr(m(\theta) \in (0, 1)) > 0$.

With the pooling strategy, the player pools in the same action and does not use his copacity to process information.

With the perfect separating strategy, the player partitions supp$(P(\cdot))$ into two positive-probability events $\{\theta \in \operatorname{supp}(P(\cdot)) : m(\theta) = 1\}$ and $\{\theta \in \operatorname{supp}(P(\cdot)) : m(\theta) = 0\}$ corresponding to "invest" and "not invest", repectively. He knows exactly which event happens and his capacity constraint does not bind. With the partial separating strategy the player's capacity constraint binds and he can not perfectly distinguish the above two events. The pooling strategy and the perfect separating strategy are two extreme cases where capacity constraint is slack, either because there is no need to acquire information or because the player's capacity is large enough. A natural question here is whether an equilibrium can consist of different types of strategies. Our answer is no and we can prove an even stronger result.

Proposition 2 *All the equilibria of the constrained information acquisition problem are symmetric, i.e. if a pair* $(m_1(\cdot), m_2(\cdot))$ *is an equilibrium, then* $\Pr(m_1(\theta) = m_2(\theta)) = 1$.

Proof. see Appendix B. ■

This proposition allows us to use a single function $m(\cdot)$ to represent the equilibrium thereafter.

The symmetry of the equilibria comes from the symmetry of the payoff matrix (2.1) and the strategic complementarity $r > 0$. Each player attempts to match up to his opponent's strategy due to the coordinating motive. If there is strategic substitutability instead of strategic complementarity, the equilibria will not be symmetric even though the payoff matrix is symmetric. In the rest of this subsection, we characterize each of the three types of equilibria.

Lemma 02 *the constrained information acquisition problem has an equilibrium with at least one player pooling in* I *(*N*) iff* $\Pr(\theta \geq 0) = 1$ *(*$\Pr(\theta - r \leq 0) = 1$*).*

Proof. see Appendix B. ■

Proposition 3 *the constrained information acquisition problem has a pooling equilibrium* (I, I) *(*(N, N)*) iff* $\Pr(\theta \geq 0) = 1$ *(*$\Pr(\theta - r \leq 0) = 1$*).*

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument.

(Sufficiency) If $\Pr(\theta \geq 0) = 1$, both players pooling in I is obvious an equilibrium.

(Necessity) the necessity is a direct implication of Lemma 02. ■

Proposition 3 establishes the sufficient and necessary condition for the existence of the pooling equilibria.

In principle, the players have two motives in collecting information. One is to reduce the uncertainty of the fundamental and the other is to coordinate with their opponents. When his opponent is pooling, the player has no coordinating incentive. Moreover, if $\Pr(\theta \geq 0) = 1$ ($\Pr(\theta - r \leq 0) = 1$), I (N)always dominates N (I). The players have no incentive to collect information and thus must pool in the same action. However, the condition in Proposition 3 does not exclude other types of equilibria. For example, there might also exist perfect separating equilibria when some other conditions are satisfied, as shown in Proposition 4.

The pooling equilibrium is an extreme case where both players pay zero attention to the fundamental. The other extreme is the perfect separating equilibrium where both players' pay non-zero attention but their capacity constraints are slack. In this case, player i partitions supp($P(\cdot)$) into two positive-probability events and choose different actions upon the occurence of different events. Recall that player i's perfect separating strategy can be totally characterized by the event that he invests: $S_I^i \triangleq \{\theta \in \text{supp}(P(\cdot)) : m_i(\theta) = 1\}$ (note that $\Pr(m_i(\theta) = 1) + \Pr(m_i(\theta) = 0) = 1$). Let $A_+ = \{\theta \in \text{supp}(P(\cdot)) : \theta - r > 0\}$, $A_- = \{\theta \in \text{supp}(P(\cdot)) : \theta < 0\}$ and $A_0 =$supp$(P(\cdot)) \setminus (A_+ \cup A_-)$, then we have the following proposition:

Proposition 4 *i) the constrained information acquisition problem has a perfect separating equilibrium iff the following three condtions are satisfied: a)* $\Pr(A_+) < 1$, $\Pr(A_-) < 1$*; b)* $\Pr(A_+) = \Pr(A_-) = 0$ *implies* $\exists B \subset A_0$*, s.t.* $\Pr(B) \in (0,1)$ *and* $\kappa \geq H(\Pr(B))$*; c)* $\kappa \geq \min\{H(\Pr(A_+)), H(\Pr(A_-))\}$*; where* $\forall p \in [0,1]$*,* $H(p) \triangleq -p \cdot \ln p - (1-p) \cdot \ln(1-p)$ *is the bivariate Shannon entropy function*[10]*;*

ii) when the conditions in i) are satisfied, all the perfect separating equilibrium can be characterized by $\{S_I \subset \text{supp}(P(\cdot)) : A_+ \subset S_I \subset A_+ \cup A_0 \text{ and } H(\Pr(S_I)) \leq \kappa\}$ *with both players playing* I *when* S_I *happens and playing* N *othewise.*

Proof. By Proposition 2, all the equilibria are symmetric, thus we use S_I, the event upon which both players invest, to represent an arbitrary perfect separating equilibrium.

i) (Sufficiency) Suppose $\min\{H(\Pr(A_+)), H(\Pr(A_-))\} > 0$. In the case of $H(\Pr(A_+)) \geq H(\Pr(A_-)) > 0$, let $S_I = A_+ \cup A_0$ and both players play the strategy characterized by S_I, i.e. $\forall \theta \in S_I$, $m(\theta) = 1$ and $\forall \theta \in$supp$(P(\cdot)) \setminus S_I$, $m(\theta) = 0$. Since $I(m(\cdot)) = H(\Pr(S_I)) = H(\Pr(A_+ \cup A_0)) = H(\Pr(A_-)) \leq \kappa$, this strategy is feasible. It is obvious that S_I is the best response to itself. Also note that $\Pr(A_-) \in (0,1)$ since $H(\Pr(A_-)) > 0$ and $\Pr(A_-) < 1$ as assumed. Thus $\Pr(S_I) \in (0,1)$ and we have constructed a perfect separating equilibrium. In the case of $0 < H(\Pr(A_+)) < H(\Pr(A_-))$, let $S_I = A_+$, then (S_I, S_I) is an equilibrium by the same argument.

Now consider the case $\min\{H(\Pr(A_+)), H(\Pr(A_-))\} = 0$. If $H(\Pr(A_-)) = 0$, we know that $\Pr(A_-) = 0$ since $\Pr(A_-) < 1$. Thus $\Pr(A_+) < 1$ implies $\Pr(A_0) > 0$. When $\Pr(A_+) > 0$, $S_I = A_+$ is a perfect separating equilibrium. Otherwise, by condition b) $\Pr(A_+) = 0$ implies $\exists B \subset A_0$, s.t. $\Pr(B) \in (0,1)$ and $\kappa \geq H(\Pr(B))$. Then $S_I = B$ is a perfect separating

[10] Note that $H(0) \triangleq \lim_{p \to 0} H(p) = 0$, $H(1) \triangleq \lim_{p \to 1} H(p) = 0$.

equilibrium. The same argument also works if $H(\Pr(A_+)) = 0$.

(Necessity) Let $S_I \subset \text{supp}(P(\cdot))$ characterize the perfect separating equilibrium.

If $\Pr(A_+) = 1$ or $\Pr(A_-) = 1$, only the pooling equilibrium can exist. Then we must have $\Pr(A_+) < 1$ and $\Pr(A_-) < 1$, i.e. condition a) holds.

Suppose $\Pr(A_+) = \Pr(A_-) = 0$, and $\forall B \subset A_0$, $\Pr(B) \in \{0, 1\}$ or $\kappa < H(\Pr(B))$, then $S_I \subset A_0$ and $\Pr(S_I) \in (0, 1)$ implies $\kappa < H(\Pr(S_I))$, i.e. the capacity constraint is violated and $S_I \subset \text{supp}(P(\cdot))$ cannot be a perfect separating equilibrium. This contradiction shows the necessity of condition b).

Suppose $S_I \subset \text{supp}(P(\cdot))$ is a perfect separating equilibrium, thus $A_+ \subset S_I \subset A_+ \cup A_0$. This implies $\Pr(A_+) \leq \Pr(S_I) \leq \Pr(A_+ \cup A_0)$. Since the Shannon entropy function is concave, we have $\kappa \geq I(m(\cdot)) = H(\Pr(S_I)) \geq \min\{H(\Pr(A_+)), H(\Pr(A_+ \cup A_0))\} = \min\{H(\Pr(A_+)), H(1 - \Pr(A_+ \cup A_0))\} = \min\{H(\Pr(A_+)), H(\Pr(A_-))\}$, where the first inequality follows the capacity constraint, the first equality follows the definition of mutual information, the second inequality comes from the cacavity of $H(\cdot)$ and the last equlity holds since $H(p) = H(1 - p)$. Therefore, condition c) is also necessary.

ii) the proof is omitted here, since it can be directly derived from the above proof. ■

From Proposition 4, we see multiple equlibria may emerge if $\Pr(A_0) > 0$ and

$\kappa \geq \min\{H(\Pr(A_+)), H(\Pr(A_-))\}$. Note that $\Pr(A_0) = \Pr(r \geq \theta \geq 0)$ is the probability of the event that coordination is important. For given common prior, the larger the strategic complementarity, the more probable that coordination matters the players welfare. Thus multiplicity results from high strategic complementarity and high capacity.

The capacity constraint is slack for both the pooling equilibrium and the perfect separating equilibrium, thus the players have no need to take care of their attention allocation. The most interesting case is the partial separating equilibrium, the intermediate case between the pooling and the perfect separating equilibria. As shown in the last case of the proof of Lemma 06 in Appendix B, however, there might exist partial separating equilibrium with slack capacity constraints. This case is very rare since it exists under very restrictive conditions (see the remark of Lemma 06). In the main part of the paper, we only study the partial separating equilibria with binding capacity constraints.

Proposition 5 *the constrained information acquisition problem has a partial separating equilibrium if* $\Pr(\theta < 0) > 0$, $\Pr(\theta - r > 0) > 0$ *and* $\kappa < \min\{H(\Pr(A_+)), H(\Pr(A_-))\}$.

Proof. by Proposition 3 and 4, if $\Pr(\theta < 0) > 0$, $\Pr(\theta - r > 0) > 0$ and $\kappa < \min\{H(\Pr(A_+)), H(\Pr(A_-))\}$, there is no pooling equilibrium and perfect separating equilibrium. Then Proposition 1 implies the existence of a partial separating equilibrium. ■

The following corollaries characterize the partial separating equilibria.

Corollary 5.1 *i) let* $m_i(\theta) = m(\theta)$*,* $i \in \{1,2\}$ *be a partial separating equilibrium with binding capacity constraints, then* $\Pr(m(\theta) \in (0,1)) = 1$*; ii) a partial separating equilibrium with binding capapcity constraints is characterized by the following equation:*

$$\forall \theta \in supp(P(\cdot)),\ \theta - r \cdot [1 - m(\theta)] = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] \tag{3.8}$$

where $p_I = \int m(\theta) \cdot dP(\theta)$ *and* $\mu > 0$ *is the Lagrangian multiplier for the capacity constraint.*

Proof. i) recall that any equilibrium of the constrained information acquisition problem must satisfy (3.7)

$$\begin{aligned} \forall \theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m_j(\theta)) &= \mu_i \cdot \left[\ln\left(\frac{m_i(\theta)}{1-m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right] + \lambda_i(\theta) - \eta_i(\theta) \quad i,j \in \{1,2\}, i \neq j \end{aligned} \tag{3.7}$$

Proposition 2 says that the equilibrium is symmetric, thus we have

$$\begin{aligned} \forall \theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m(\theta)) &= \mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] + \lambda(\theta) - \eta(\theta) \end{aligned} \tag{3.9}$$

Note that the Lagrangian multiplier μ must be strictly positive, otherwise the capacity constraint does not bind. Suppose $\Pr(m(\theta) = 0 \text{ or } m(\theta) = 1) > 0$. Since $m(\cdot)$ is a partial separating equilibrium, $\Pr(m(\theta) \in (0,1)) > 0$ and thus $p_I = \int m(\theta) \cdot dP(\theta) \in (0,1)$. Then for any θ, s.t. $m(\theta) = 0$ (or $m(\theta) = 1$), $\mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] = -\infty$ (or ∞) and (3.9) does not hold. Therefore, we must have $\Pr(m(\theta) = 0 \text{ or } m(\theta) = 1) = 0$, i.e. $\Pr(m(\theta) \in (0,1)) = 1$.

ii) $\Pr(m(\theta) \in (0,1)) = 1$ implies $\lambda(\theta) = \eta(\theta) = 0$ and then (3.9) becomes

$$\forall \theta \in \operatorname{supp}(P(\cdot)),\ \theta - r \cdot (1 - m(\theta)) = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] \tag{3.8}$$

■

Remarks: (3.8) is intuitive. Given player j's strategy $m_j(\theta) = m(\theta)$, the left hand side of (3.8) is player i's marginal benefit of increasing $m_i(\theta)$. $\mu > 0$ is the shadow price of an extra bit of information and $\left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right]$ is the derivative of the mutual information with respect to $m_i(\theta)$, thus the right hand side of (3.8) is player i's marginal cost of increasing $m_i(\theta)$. (3.8) says that the marginal cost must equal the marginal benefit. Also note that $\ln\left(\frac{p_I}{1-p_I}\right)$ is the unconditional (or say, average) log-likelihood ratio of "I" relative to "N" , while $\ln\left(\frac{m(\theta)}{1-m(\theta)}\right)$ is player i's log-likelihood ratio of choosing "I" over "N" conditional on θ. Then (3.8) says that if the marginal benefit $\theta - r(\theta)(1-m(\theta))$ is positive (negative) for some θ, player i should choose "I" with a probability higher (lower) than the average level. If there is no capacity constraint, i.e. $\mu = 0$, he would choose $m(\theta) = 1$. Then we return to the pooling or perfect separating equilibirum.

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[11] in the $\theta \sim m$ plane about the point $(\theta_0, 1/2)$, where $\theta_0 = r/2 - \mu\cdot\ln\left(\frac{p_I}{1-p_I}\right)$. The single equation (3.8) (let's ignore $p_I = \int m(\theta)\cdot dP(\theta)$ at this point) actually suggests that θ_0, or say, p_I determines the position of $m(\cdot)$ while r and $\mu = \mu(\kappa, P(\cdot))$ determine its possible shapes, as shown in the corollary below.

Corollary 5.2 *let* $\widetilde{r} = \frac{r}{4\cdot\mu}$, *i) if* $\widetilde{r} \le 1$, *then* $\forall\theta \in supp(P(\cdot))$, *there is a unique value* m *such that* (3.8) *holds; ii) if* $\widetilde{r} > 1$, *let*

$$\begin{aligned}
m_1 &= \left(1 + (1-\widetilde{r}^{-1})^{1/2}\right)/2 \\
\theta_1 &= \mu\cdot\left[\ln\left(\frac{m_1}{1-m_1}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] + r\cdot[1-m_1] \\
&= \mu\cdot\left[\ln\left(\frac{1+(1-\widetilde{r}^{-1})^{1/2}}{1-(1-\widetilde{r}^{-1})^{1/2}}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] + \frac{r}{2}\cdot\left(1-(1-\widetilde{r}^{-1})^{1/2}\right) \\
m_2 &= \left(1 - (1-\widetilde{r}^{-1})^{1/2}\right)/2 \\
\theta_2 &= \mu\cdot\left[\ln\left(\frac{m_2}{1-m_2}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] + r\cdot[1-m_2] \\
&= \mu\cdot\left[\ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] + \frac{r}{2}\cdot\left(1+(1-\widetilde{r}^{-1})^{1/2}\right)
\end{aligned}$$

and define $\overline{m} : (\theta_1, +\infty) \cap supp(P(\cdot)) \to (m_1, 1)$ *by* $\theta - r\cdot[1-\overline{m}(\theta)] = \mu\cdot\left[\ln\left(\frac{\overline{m}(\theta)}{1-\overline{m}(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right]$, $\underline{m} : (-\infty, \theta_2) \cap supp(P(\cdot)) \to (0, m_2)$ *by* $\theta - r\cdot[1-\underline{m}(\theta)] = \mu\cdot\left[\ln\left(\frac{\underline{m}(\theta)}{1-\underline{m}(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right]$ *and* $\widetilde{m} : [\theta_1, \theta_2] \cap supp(P(\cdot)) \to [m_2, m_1]$ *by* $\theta - r\cdot[1-\widetilde{m}(\theta)] = \mu\cdot\left[\ln\left(\frac{\widetilde{m}(\theta)}{1-\widetilde{m}(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right]$, *then* (3.8) *implies that* $\forall\theta \in supp(P(\cdot)) \cap [\theta_1, \theta_2]$, $m(\theta) \in \{\underline{m}(\theta), \overline{m}(\theta), \widetilde{m}(\theta)\}$, $\forall\theta \in (-\infty, \theta_1) \cap supp(P(\cdot))$, $m(\theta) = \underline{m}(\theta)$, *and* $\forall\theta \in (\theta_2, +\infty) \cap supp(P(\cdot))$, $m(\theta) = \overline{m}(\theta)$.

Proof. i) take derivative of both sides of (3.8) with respect to m, we find that

$$\frac{d\theta}{dm} = \frac{\mu}{m(1-m)} - r$$

when $\widetilde{r} = \frac{r}{4\cdot\mu} \le 1$, $\frac{d\theta}{dm} > 0$ for all $m \in (0, 1/2) \cup (1/2, 1)$, thus there exists a one-to-one mapping between m and θ.

[11] This symmetry comes from the constant strategic complementarity.

ii)if $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, $\exists$ $m_1 = \left(1 + (1 - \widetilde{r}^{-1})^{1/2}\right)/2$ and $m_2 = \left(1 - (1 - \widetilde{r}^{-1})^{1/2}\right)/2$, such that $\frac{d\theta}{dm} = 0$ at these two points. Then the result follows the fact that $\frac{d^2\theta}{dm^2}|_{m=m_1} > 0$ and $\frac{d^2\theta}{dm^2}|_{m=m_2} < 0$. ■

The shape of $m(\cdot)$ evolves as $\widetilde{r}$ increases, as shown in the figures below:

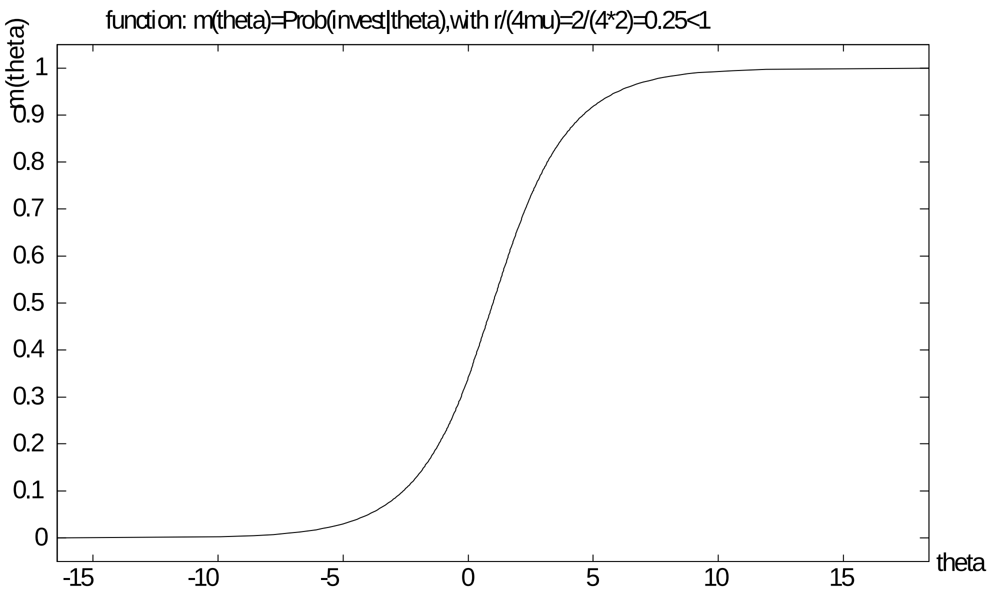

figure 01 (effective strategic complementarity $\widetilde{r} < 1$)

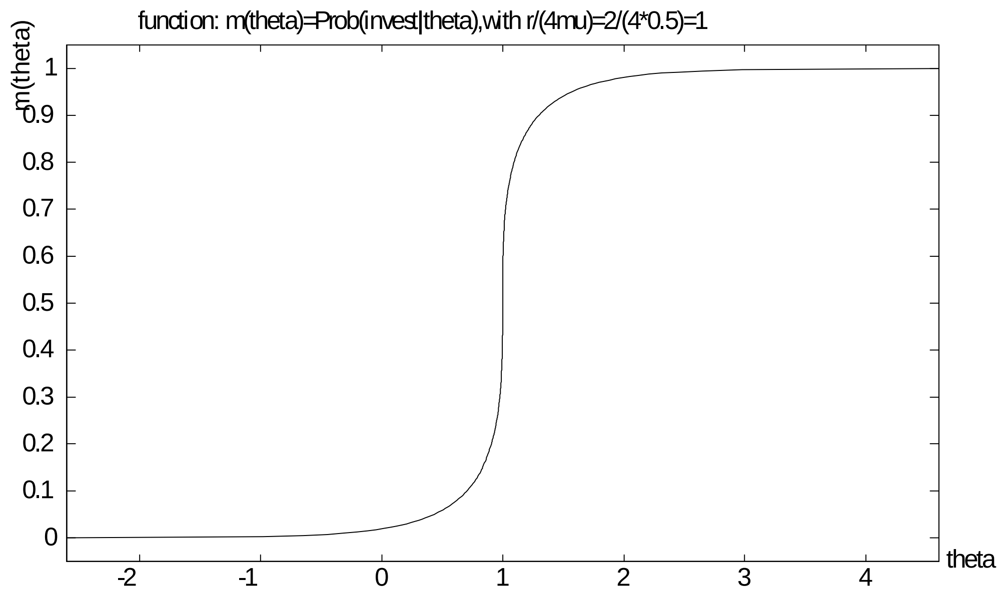

figure 02 (effective strategic complementarity $\widetilde{r} = 1$)

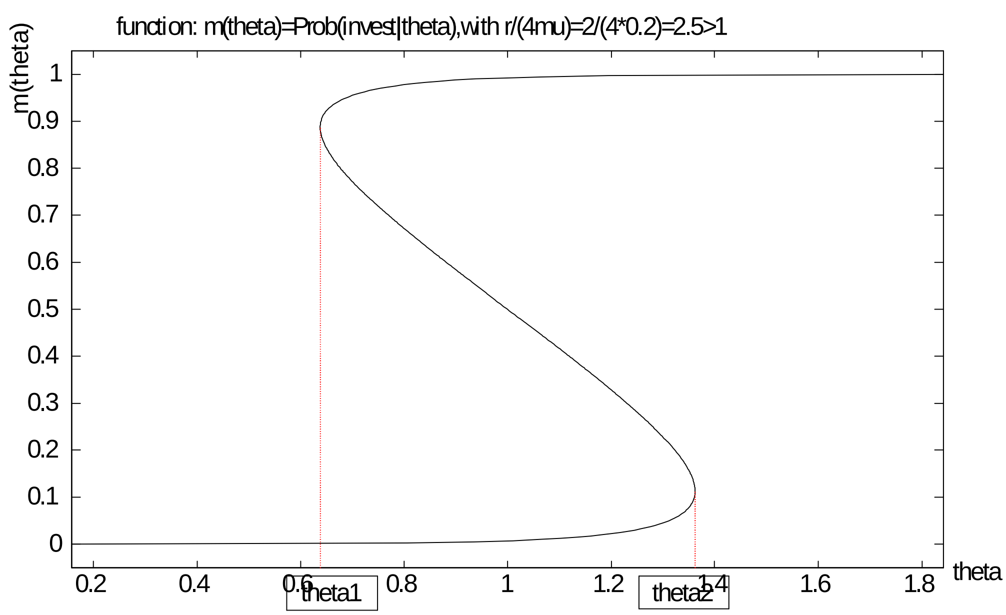

figure 03 (effective strategic complementarity $\widetilde{r} > 1$)

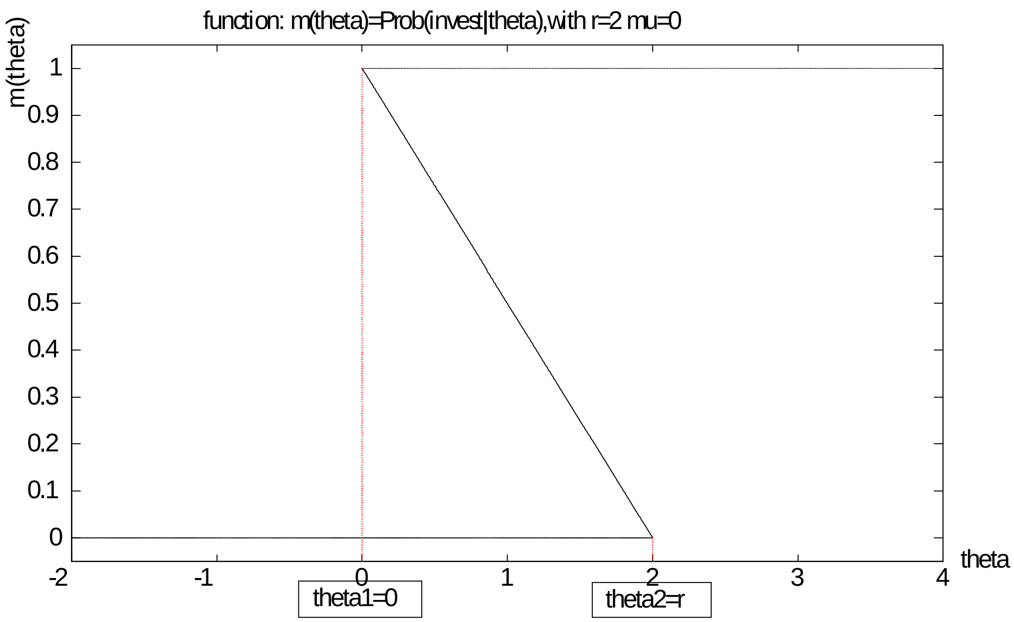

figure 04 (shadow price $\mu = 0$, it reduces to perfect separating equilibrium)

Several messages of the model:

1) generally, the equilibrium $m(\cdot)$ has an "S"-like shape. This represents the players' desire to play a switching strategy, i.e. invest if the fundamental is high and do not invest if it is low. They can definitely play such a strategy when the capacity constraint is slack and thus return to the perfect separating equilibrium. When the capacity constraint binds, however, the players can only approximate the switching strategy but cannot attain it. As shown in the above figures, they almost certainly invest (not invest) as the fundamental becomes large (negatively large) but are reluctant to do so for the intermediate values.

2) these figures clearly show how multiple equilibria may emerge[12] as $\widetilde{r} = \frac{r}{4 \cdot \mu}$ becomes larger than unit. When κ, the players' capacity of information processing increases, the shadow price of information processing, $\mu = \mu(\kappa, P(\cdot))$, becomes lower and facilitates the emergence of multiplicity. This is intuitive. More restrictive capacity constraint imposes more randomness in the players' behavior. In the limit $\kappa \to 0$, they are just the noisy traders. The players always desire more coordination but lessening capacity reduces their ability to do so. Moreover, it also harms the players' incentive to coordinate. Knowing that his opponent has samller capacity and thus noisier behavior, the player will no longer have as much incentive to coordinate as before. In other words, increasing capacity facilitates coordination and thus the emergence of multiple equilibria. We call $\widetilde{r} = \frac{r}{4 \cdot \mu}$ the effective strategic complementarity thereafter. It shows

[12] We will prove this multiplicity in the costly information acquisition problem.

that not only does the physical strategic complementarity r affect the equilibrium, but also the shadow price, or say, the capacity of information processing plays a role in this coordination game.

3) Monotonic Likelihood Ratio Property (MLRP) is a key assumption in most of the global games literature. As a part of the the information structure, MLRP cannot be justified within the models with exogenous information structure. Our model provides a way to examine MLRP. By the Lemma1, $\Pr(s_I|\theta) = m(\theta)$, thus MLRP may not hold for the effective strategic complementarity larger than one, as shown in figure 03. The reason is that when the motive of coordination is very large (i.e. r is large) or acquiring information is very easy (i.e. μ is small or κ is large), a player has enough incentive and ability to coordinate with his oppenent who is playing the wierd inverse-MLRP strategy. MLRP always holds when $\widetilde{r} = \frac{r}{4 \cdot \mu} \leq 1$.

4) As capacity becomes large enough, $\mu = \mu \to 0$ and the equilibrium approches the shape shown in figure 04, which actually approximates the perfect separating equilibria.

Through analyzing the constraint information acquisition problem, we provide a fully characterization of the possible equilibria. However, μ depends on $(\kappa, P(\cdot))$ in a complex way, which restricts our further analysis. In Section 4, we turn to the costly information acquisition problem where $\mu > 0$ is an exogenously given marginal cost of information acquisition and each player decides the amount of capacity to purchase according to his own interest.

4 THE COSTLY INFORMATION ACQUISITION PROBLEM

The basic setup of the costly information acquisition problem is almost the same as before, except that there is no capacity constraint and the players are able to acquire information at some marginal cost $\mu > 0$. In practice, μ can be the cost of buying a new computer or hiring another analyst, etc.

Note that the second part of the preference assumption (A2) is automatically satisfied in the costly information acquisition problem.

Formally, an equilibrium of the costly information acquisition problem is a pair $(m_1(\cdot), m_2(\cdot))$ solving the the following problem:

$$\begin{aligned} i, j \quad \in \quad & \{1, 2\}, \, i \neq j, \;\; m_i(\cdot) \in \arg\max_{\widetilde{m}_i(\cdot)} u_i(\widetilde{m}_i(\cdot), m_j(\cdot)) = \int \widetilde{m}_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta) \\ & -\mu \cdot I(\widetilde{m}_i(\cdot)) \qquad (4.1) \end{aligned}$$

$$\text{s.t.} \quad \widetilde{m}_i(\cdot) \in \Omega$$
$$\forall\theta \in \text{supp}(P(\cdot)),\ m(\theta) \in [0,1] \qquad (4.2)$$

where $I(\widetilde{m}_i(\cdot)) = \int [\widetilde{m}_i(\theta)\ln\widetilde{m}_i(\theta) + (1-\widetilde{m}_i(\theta))\ln(1-\widetilde{m}_i(\theta))]\,dP(\theta) - \widetilde{p}_{Ii}\ln\widetilde{p}_{Ii} - (1-\widetilde{p}_{Ii})\ln(1-\widetilde{p}_{Ii})$ is the amount of information acquired and $\widetilde{p}_{Ii} = \Pr(\text{player } i \text{ chooses } I) = \int \widetilde{m}_i(\theta)\cdot dP(\theta)$.

In this section, first we characterize all the possible equilibria of the costly information acquisition problem; second we study the effects of private information acquisition through the comparative static analysis with respect to the marginal cost μ; third we study the effects of public information through the comparative analysis with respect to the common prior $P(\cdot)$. We also compare our results to that of the previous models and find significant difference.

4.1 Characterizing The Equilibria

Under the same conditions in the previous section, we show that the equilibirum of this new problem exists.

Proposition 6 *the equlibrium of the costly information acquisition problem exists.*

Proof. let $\widetilde{\Omega} = \{\widetilde{m}_i(\cdot) \in \Omega : \forall\theta \in \text{supp}(P(\cdot)), \widetilde{m}_i(\theta) \in [0,1]\}$, then $\widetilde{\Omega}$ is compact since it is a closed subset of a compact space Ω. Because $I(\widetilde{m}_i(\cdot))$ is a continuous and convex functional of $\widetilde{m}_i(\cdot)$, $u_i(\widetilde{m}_i(\cdot), m_j(\cdot))$ is continuous and concave in $\widetilde{m}_i(\cdot)$. According to Nash equilibrium theorem, the Nash equilibrium exists. ■

The main difference between the two problems is:

Lemma 03 *the perfect separating strategies can not be played in an equilibrium of the costly information acquisition problem.*

Proof. Player i's Lagrangian is

$$\begin{aligned} L_i = & \int \widetilde{m}_i(\theta)\cdot[\theta - r\cdot(1-m_j(\theta))]\cdot dP(\theta) + \mu\cdot[\widetilde{p}_{Ii}\cdot\ln\widetilde{p}_{Ii} - (1-\widetilde{p}_{Ii})\cdot\ln(1-\widetilde{p}_{Ii})] \\ & -\mu\cdot\int[\widetilde{m}_i(\theta)\ln\widetilde{m}_i(\theta) + (1-\widetilde{m}_i(\theta))\ln(1-\widetilde{m}_i(\theta))]\cdot dP(\theta) \\ & + \int[\lambda_i(\theta)[1-\widetilde{m}_i(\theta)] + \eta_i(\theta)\widetilde{m}_i(\theta)]\cdot dP(\theta) \end{aligned} \qquad (4.3)$$

where

$$\widetilde{p}_{Ii} = \Pr(\text{player } i \text{ chooses } I) = \int \widetilde{m}_i(\theta)\cdot dP(\theta) \quad i \in \{1,2\} \qquad (4.4)$$

The first order condition implies

$$\begin{aligned}\forall\theta &\in \text{supp}\left(P\left(\cdot\right)\right), \\ \theta - r\cdot\left(1-m_j\left(\theta\right)\right) &= \mu\cdot\left[\ln\left(\frac{m_i\left(\theta\right)}{1-m_i\left(\theta\right)}\right)-\ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right]+\lambda_i\left(\theta\right)-\eta_i\left(\theta\right)\quad i,j\in\{1,2\},i\neq j \quad (4.5)\end{aligned}$$

Suppose player i plays a perfect separating strategy, i.e. $\Pr\left(m_i\left(\theta\right)\in(0,1)\right)=0$ and $\Pr\left(m_i\left(\theta\right)=1\right)\in(0,1)$. Then $p_{Ii}=\Pr\left(m_i\left(\theta\right)=1\right)\in(0,1)$. Since the cost of acquiring information is $\mu>0$, $\forall\theta\in$supp$(P\left(\cdot\right))$ s.t. $m_i\left(\theta\right)=1$ implies $\mu\cdot\left[\ln\left(\frac{m_i(\theta)}{1-m_i(\theta)}\right)-\ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right]=\infty$, thus (4.5) does not hold with probability $\Pr\left(m_i\left(\theta\right)=1\right)>0$ and $m_i\left(\cdot\right)$ cannot be an equilibrium strategy. ■

Compared to the Lagrangian multiplier in the constrained information acquisition problem, here the cost of acquiring information is exogenously given and always strictly positive, which makes the margianl cost of acquiring information too high (or too low) to support a perfect separating equilibirum. According to Lemma 03, only the pooling strategies and the partial separating strategies can be played in the equilibria. A natural question here is whether an equilibrium can consist of different types of strategies. Our answer is no and we can prove an even stronger result.

Proposition 7 *All the equilibria of the costly information acquisition problem are symmetric, i.e. if a pair* $\left(m_1\left(\cdot\right),m_2\left(\cdot\right)\right)$ *is an equilibrium, then* $\Pr\left(m_1\left(\theta\right)=m_2\left(\theta\right)\right)=1$.

Proof. see Appendix B. ■

This proposition allows us to use a single function $m\left(\cdot\right)$ to represent the equilibrium thereafter.

Proposition 8 *the costly information acquisition problem has a pooling equilibrium* (I,I) *(*(N,N)*) iff* $\Pr\left(\theta\geq 0\right)=1$ *(*$\Pr\left(\theta-r\leq 0\right)=1$*).*

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument.

(Sufficiency) If $\Pr\left(\theta\geq 0\right)=1$, both players pooling in I is obvious an equilibrium.

(Necessity) the necessity is a direct implication of Lemma 03. ■

Proposition 8 establishes the sufficient and necessary condition for the existence of the pooling equilibria. The following proposition chracterizes the partial equilibria.

Proposition 9 *the costly information acquisition problem has a partial separating equilibrium if* $\Pr\left(\theta<0\right)>0$ *and* $\Pr\left(\theta-r>0\right)>0$.

Proof. by Proposition 8, if $\Pr(\theta < 0) > 0$, $\Pr(\theta - r > 0) > 0$, there is no pooling equilibrium. Then Proposition 6 and Lemma 03 imply the existence of a partial separating equilibrium. ■

The following corollaries characterize the partial separating equilibria.

Corollary 9.3 *i) let* $m_i(\theta) = m(\theta)$, $i \in \{1,2\}$ *be a partial separating equilibrium, then* $\Pr(m(\theta) \in (0,1)) = 1$*; ii) a partial separating equilibrium is characterized by the following equation:*

$$\forall\theta \in supp\left(P(\cdot)\right), \ \theta - r\cdot[1 - m(\theta)] = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right)\right] \qquad (4.6)$$

where $p_I = \int m(\theta)\cdot dP(\theta)$ *and* $\mu > 0$ *is the cost of acquiring information.*

Proof. i) is proved in Lemma 09 in Appendix B. ii) directly follows i) and (4.5). ■

The intuition behind (4.6) is the same as that of (3.8).

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[13] in the $\theta \sim m$ plane about the point $(\theta_0, 1/2)$, where

$$\theta_0 = r/2 - \mu\cdot\ln\left(\frac{p_I}{1-p_I}\right) \qquad (4.7)$$

. Combining (4.6) and (4.7) leads to

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

Thus we can index $m(\cdot)$ by θ_0, i.e. $m(\cdot) = m(\cdot,\theta_0)$ and

$$\theta - \theta_0 = \mu\cdot\ln\left(\frac{m(\theta,\theta_0)}{1-m(\theta,\theta_0)}\right) + r\cdot\left(\frac{1}{2} - m(\theta,\theta_0)\right) \qquad (4.8)$$

θ_0 determines the position of $m(\cdot,\theta_0)$ but has no effect on its shape. The shape of $m(\cdot,\theta_0)$ is determined by the two key parameters, r and μ.

Corollary 9.4 *let* $\widetilde{r} = \frac{r}{4\cdot\mu}$*, i) if* $\widetilde{r} \leq 1$*, then* $\forall\theta \in$*supp*$(P(\cdot))$*, there is a unique* m *such that*

[13] This symmetry comes from the constant strategic complementarity.

(4.6) *holds; ii) if* $\widetilde{r}>1$, *let*

$$\begin{aligned}
m_1 &= \left(1+(1-\widetilde{r}^{-1})^{1/2}\right)/2\\
\theta_1 &= \mu\cdot\left[\ln\left(\frac{m_1}{1-m_1}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]+r\cdot[1-m_1]\\
&= \mu\cdot\left[\ln\left(\frac{1+(1-\widetilde{r}^{-1})^{1/2}}{1-(1-\widetilde{r}^{-1})^{1/2}}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]+\frac{r}{2}\cdot\left(1-(1-\widetilde{r}^{-1})^{1/2}\right)\\
&= \theta_0-\ \mu\cdot\ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right)-\ \frac{r}{2}\cdot(1-\widetilde{r}^{-1})^{1/2} \qquad (4.9)\\
m_2 &= \left(1-(1-\widetilde{r}^{-1})^{1/2}\right)/2\\
\theta_2 &= \mu\cdot\left[\ln\left(\frac{m_2}{1-m_2}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]+r\cdot[1-m_2]\\
&= \mu\cdot\left[\ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]+\frac{r}{2}\cdot\left(1+(1-\widetilde{r}^{-1})^{1/2}\right)\\
&= \theta_0+\mu\cdot\ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right)+\frac{r}{2}\cdot(1-\widetilde{r}^{-1})^{1/2} \qquad (4.10)
\end{aligned}$$

and define $\overline{m}:(\theta_1,+\infty)\cap supp(P(\cdot))\rightarrow(m_1,1)$ *by* $\theta-r\cdot[1-\overline{m}(\theta)]=\mu\cdot\left[\ln\left(\frac{\overline{m}(\theta)}{1-\overline{m}(\theta)}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]$, $\underline{m}:(-\infty,\theta_2)\cap supp(P(\cdot))\rightarrow(0,m_2)$ *by* $\theta-r\cdot[1-\underline{m}(\theta)]=\mu\cdot\left[\ln\left(\frac{\underline{m}(\theta)}{1-\underline{m}(\theta)}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]$ *and* $\widetilde{m}:[\theta_1,\theta_2]\cap supp(P(\cdot))\rightarrow[m_2,m_1]$ *by* $\theta-r\cdot[1-\widetilde{m}(\theta)]=\mu\cdot\left[\ln\left(\frac{\widetilde{m}(\theta)}{1-\widetilde{m}(\theta)}\right)-\ln\left(\frac{p_I}{1-p_I}\right)\right]$, *then* (4.6) *implies that* $\forall\theta\in supp(P(\cdot))\cap[\theta_1,\theta_2]$, $m(\theta)\in\{\underline{m}(\theta),\overline{m}(\theta),\widetilde{m}(\theta)\}$, $\forall\theta\in(-\infty,\theta_1)\cap supp(P(\cdot))$, $m(\theta)=\underline{m}(\theta)$, *and* $\forall\theta\in(\theta_2,+\infty)\cap supp(P(\cdot))$, $m(\theta)=\overline{m}(\theta)$.

Proof. the proof is the same as Corollary 5.2 and is omitted here. ■

Remarks: the equilibrium is determined by r, μ and the common prior $P(\cdot)$. On the one hand, r and μ, the parameters characterizing the players' private information acquisition behavior, determine the possible shapes of $m(\theta,\theta_0)$. According to Corollary 9.4, there are infinitely many possible shapes of $m(\cdot,\theta_0)$ to satisfy (4.8) when $\widetilde{r}=\frac{r}{4\cdot\mu}>1$. The following figures show several typical shapes:

a) the strategy is inverse-MLRP. For $\theta\in(\theta_1,\theta_2)$, we choose the decreasing component, otherwise choose the unique $m(\theta)$ satisfying (4.8). This strategy is shown below:

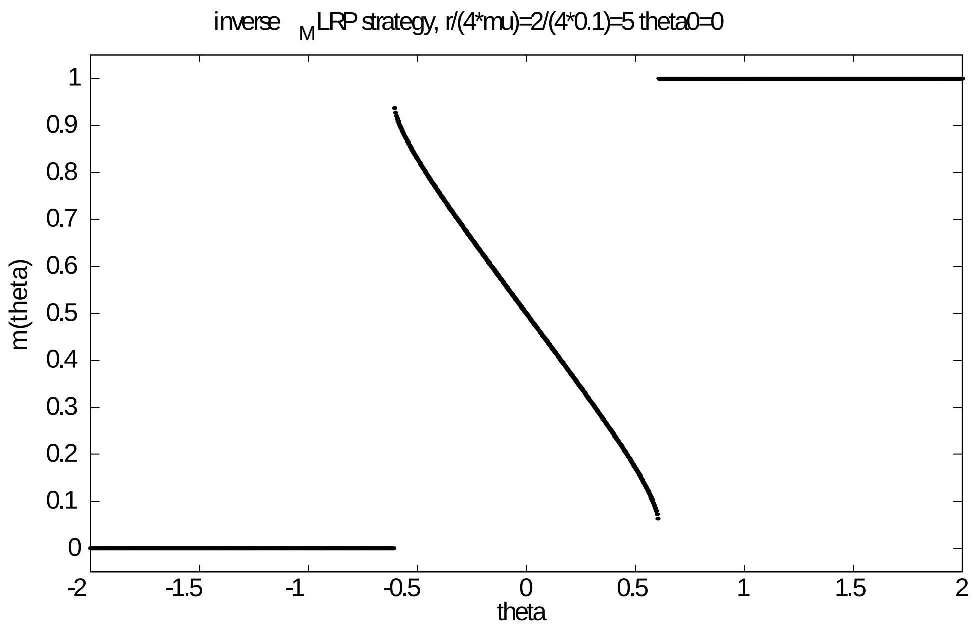

figure 05: $m(\cdot,\theta_0)$ is of shape a)

b) the strategy is MLRP. For $\theta \in (\theta_1, \infty)$, choose the upper part of the increasing component, otherwise choose the unique $m(\theta)$ satisfying (4.8). This strategy is shown below:

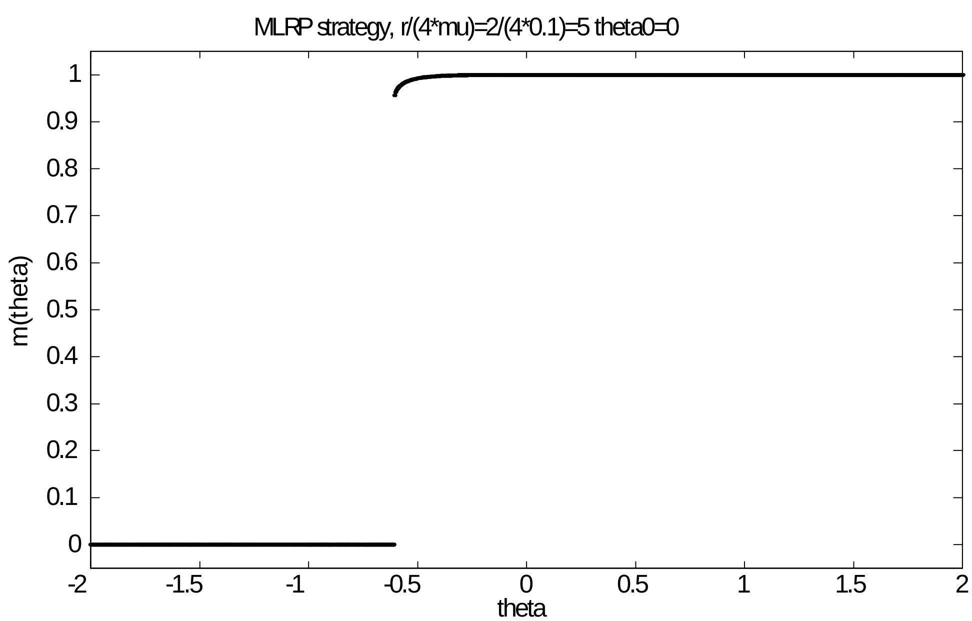

figure 06: $m(\cdot,\theta_0)$ is of shape b)

c) the strategy is MLRP. For $\theta \in (\theta_0, \infty)$, choose the upper part of the increasing component, otherwise choose the lower part of the increasing component. This strategy is shown below:

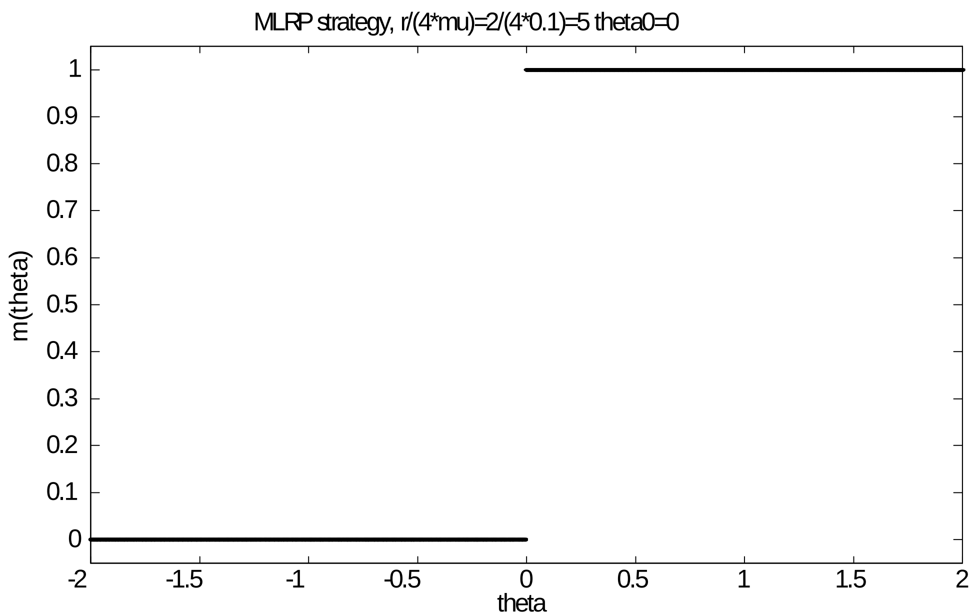

figure 07: $m(\cdot, \theta_0)$ is of shape c)

d) the strategy is MLRP. For $\theta \in (-\infty, \theta_2)$, choose the lower part of the increasing component, otherwise choose the unique $m(\theta)$ satisfying (4.8). This strategy is shown below:

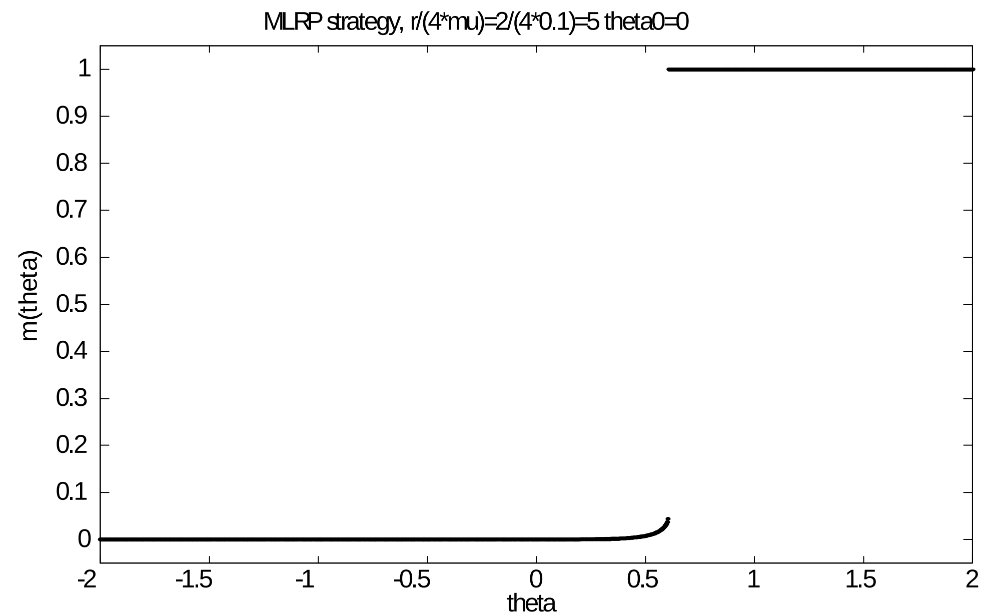

figure 08: $m(\cdot, \theta_0)$ is of shape d)

Define a set of functions

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

M is actually the set of all possible shapes of the partial separating equilibrium strategies. Note that $\#M = \left\{ \begin{array}{ll} 1 & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} \leq 1 \\ \infty & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} > 1 \end{array} \right\}$. When r and μ are given, a partial separating equilibrium $m(\theta - \theta_0)$ is determined by its shape $m \in M(r,\mu)$ and its position θ_0. Recall that the unconditional probability of investing is $p_I = \int m(\theta - \theta_0) \cdot dP(\theta)$ and (4.7) implies that the common prior $P(\cdot)$ determines θ_0 through the following equation:

$$\theta_0 = r/2 - \mu \cdot \ln\left(\frac{\int m(\theta-\theta_0)\cdot dP(\theta)}{1-\int m(\theta-\theta_0)\cdot dP(\theta)}\right) \qquad (4.11)$$

Since the public information is summarized in the common prior $P(\cdot)$, the above argument shows that it affects the equilibrium only through changing its position $\theta_0 = r/2 - \mu \cdot \ln\left(\frac{p_I}{1-p_I}\right)$ but leaves its shape unaffected. All in all, to find an equilibrum with any given shape $m \in M(r,\mu)$ is equivalent to find a fixed point θ_0 of the following mapping:

$$g(\theta_0, m) \triangleq r/2 - \mu \cdot \ln\left(\frac{\int m(\theta-\theta_0)\cdot dP(\theta)}{1-\int m(\theta-\theta_0)\cdot dP(\theta)}\right) \qquad (4.12)$$

Here $g(\theta_0, \cdot)$ is a functional of the possible shape $m \in M(r,\mu)$.

Although (4.12) is derived from the conditions of partial separating equilibria, it can cover the case of pooling equilibria as well if we let $\theta_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$. When (4.12) has a fixed point

$\theta_0 = -\infty$ ($\theta_0 = \infty$), it actually means an equilibrium pooling in I (N), since $\forall\theta \in$supp$(P(\cdot))$, $\lim_{\theta_0 \to -\infty} m(\theta - \theta_0) = 1$ ($\lim_{\theta_0 \to +\infty} m(\theta - \theta_0) = 0$) for any possible shape.

As $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$ allows for multiple possible shapes, a natural question is whether this multiplicity of possible shapes leads to multiple equilibria of the game. We answer this question in next subsection.

4.2 The Comparison Between Private Information Acquisition With Endogenous And With Exogenous Information Structures

Proposition 10 *If $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, the common prior $P(\cdot)$ is an absolutely continuous distribution with full support (i.e. the density $p(\theta)$ exists and supp$(P(\cdot)) = \mathbb{R}$), then the costly information acquisition problem has infinitely many equilibria.*

Proof. see Appendix B. ■

This proposition says that multiplicity emerges if the marginal cost of information acquisition is small relative to the strategic complementarity (i.e. $\mu < r/4$). In other words, if acquiring/processing information becomes easier (e.g. the computational power becomes cheaper, the wage of the analysts becomes smaller relative to the total profit of the firms, etc.), the players have higher capability as well as incentive to coordinate and thus more equilibria can be supported.

Another interesting observation is that even though there is no exogenous discreteness (i.e. the common prior $P(\cdot)$ only has an absolutely continuous componnent) there is a jump (maybe jumps) within any equilibrium $m(\cdot)$ due to the condition $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$. Here the existence of jumps is an essential feature of all the equilibria and does not come from changing from one equilibrium to another as did in many previous models trying to capture the sudden changes in financial systems. Thus the intrinsic discontinuity in this model have the potential to explain financial crisis within a single equilibrium.

To compare our result to that of the previous models, consider a semi-endogenous information structure model where the players are allowed to increase the accuracy of their private signals at some cost but cannot change any other aspect of the information structure. Specifically, let two players play the game with payoff matrix (2.1). The common prior about the fundamental θ is $P(\theta)$. Player $i \in \{1, 2\}$ takes action $a^i \in \{I, N\}$ after observing his private signal $x^i = \theta + \beta_i^{-1/2} \cdot \varepsilon^i$, where ε^i is distributed according to a smooth density function $f(\cdot)$, $E\varepsilon^i = 0$ and $Var(\varepsilon^i) < \infty$. Here β_i measures the precision of player i's information. The cost of acquiring information of precision β is $c \cdot h(\beta)$, where $c > 0$ is an exogenous parameter and $h(\cdot)$ is a continuous increasing function with $h(0) = 0$.

Each player's strategy involves simultaneously chooses a precision $\beta_i \in \mathbb{R}_+$ and an action rule $s^i : \mathbb{R} \to \{0,1\}$, where $s^i\left(x^i\right) = 1$ ($s^i\left(x^i\right) = 0$) means that player i chooses I (N) when observing x^i.

We write $G(c)$ for the game with cost parameter c.

Proposition 11 *If the common prior $P(\cdot)$ is an absolutely continuous distribution with density $p(\cdot)$, $\Pr(\theta > r) > 0$, $\Pr(\theta < 0) > 0$ and the noise ε^i has a full support, i.e. supp($f(\cdot)$) $= \mathbb{R}$, then $\forall \beta > 0$, $\exists \overline{c} > 0$, s.t. every rationalizable strategy in $G(c)$ for $c < \overline{c}$, has each player acquiring information of precision at least β.*

Proof. see Appendix B. ■

This proposition says that the players would like to acquire information of arbitrarily large precision if the cost of doing so is arbitrarily small. A well known result in the literature of global games is that in the models with exogenous information structure, unique equilibrium is guaranteed if the private signals are sufficiently accurate relative to the accuracy of public signals (e.g. Morris and Shin (2004)). Proposition 11 allows us to establish the standard global game result in the semi-endogenous information structure model.

Corollary 11.5 *If the common prior $P(\cdot)$ is an absolutely continuous distribution with full support (i.e. the densitiy $p(\cdot)$ exists and supp($P(\cdot)$) $= \mathbb{R}$), then $\forall \delta > 0$, $\exists \overline{c} > 0$, s.t. $\forall$ strategy $s : \mathbb{R} \to \{0,1\}$ surviving iterated deletion of strictly dominated strategies in the game $G(c)$ for $c < \overline{c}$ satisfies: $s(x) = 0$ if $x \leq r/2 - \delta$ and $s(x) = 1$ if $x \geq r/2 + \delta$.*

Proof. The proof is a direct application of Proposition 2.2 in Morris and Shin (2003) and Proposition 11. According to Proposition 2.2 of Morris and Shin (2003), $\forall \delta > 0$, $\exists \overline{\beta} > 0$, s.t. the above statement holds for all $\beta > \overline{\beta}$. Then Proposition 11 says there exists $\overline{c} > 0$ such that the players acquire information of precision at least $\overline{\beta}$. ■

Corollary 11.5 says that when the information cost approaches zero all the equilibria become approximately the unique switching strategy $s(x) = \left\{ \begin{array}{ll} 0 & \text{if } x \leq r/2 \\ 1 & \text{if } x > r/2 \end{array} \right\}$. This result is consistent with the standard global game arguments in that lowering the information cost induces more accurate private signals, reduces the common knowledge and thus facilitates the uniqueness of the equilibirum. In our model with endogenous information structure, however, Proposition 10 says that lowering information cost enhances common knowledge and facilitates multiplicity. This sharp comparison comes from the fact that too much rigidity exists in the previous global game models, where the private signal is the fundamental plus some noise. This noise is ususally assumed to be independent with the fundamental, which implicitly imposes a strong restriction on the players' information acquisition. In the language of rational inattention, this is equivalent to require the players to pay equal attention to every aspects of the fundamental. This restriction on the information structure underestimates the players' active information acquisition behavior and might not be always realistic.

It is well known that the unique equilibrium (when $c \to 0$ and thus $\beta \to \infty$) in the global games with exogenous information structure is inefficient. Both players would have enjoyed higher payoffs if they could commite to a strategy $\widetilde{s}(x) = \left\{ \begin{array}{l} 0 \text{ if } x \leq 0 \\ 1 \text{ if } x > 0 \end{array} \right\}$. In our model with endogenous information structure, however, $\forall \widehat{\theta} \in [0, r]$, $m(\theta) = \left\{ \begin{array}{l} 0 \text{ if } \theta \leq \widehat{\theta} \\ 1 \text{ if } \theta > \widehat{\theta} \end{array} \right\}$ is an equilibrium when marginal cost $\mu \to 0$, as shown in figure 04 in Section 3. Thus both the most efficient strategy $\widehat{\theta} = 0$ and the most inefficient strategy $\widehat{\theta} = r$ can be supported as equilibria. In other words, our model shows that rather than converging to the unique but inefficient equilibrium, lowering the marginal cost of information acquisition provides an opportunity to achieve a better equilibrium at the risk of being trapped in an even worse one.

The results of these two-player games still hold for the case with a continuum of players if we slightly change the payoff to "investing" to $\theta - r \cdot (1 - m)$, where m is the fraction of the players that invest. Here m is the aggregate variable of this economy and is a function of the fundamental θ. In the semi-endogenous information structure model, $m(\theta)$ is always continuous in θ provided that the noise ε^i is a continuous random variable. In our model with endogenous information structure, however, there exists some intrinsic discontinuity of m with respect to θ in all equilibria when the conditions in Proposition 10 are satisfied. Here the key condition resulting in such discontinuity is $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$. The message is that the players are able to tell if the fundamental is larger than a cutoff or not when the marginal cost of information acquisition μ is small, and they would like to do so when the strategic complementarity r is large. As before, this distinction also comes from the difference between the flexible information structure of the current model and the rigidity imposed on the previous ones.

4.3 Some Comparative Static Analysis With Respect To The Public Information

In this subsection, we study the effects of the public information through the comparative static analysis with respect to the common prior $P(\cdot)$, since the public information is summarized in it.

Proposition 12 *$\forall r > 0$, $\forall \mu > 0$ and $\forall \epsilon \in (0, r/2)$, $\exists \delta > 0$ s.t. the costly information acquisition problem has multiple equilibria for all absolutely continuous common prior $P(\cdot)$ (i.e. its density $p(\cdot)$ exists) satisfying* $\Pr(\theta \in [\epsilon, r - \epsilon]) > 1 - \delta$.

Proof. see Appendix B. ■

This proposition says that multiple equilibria emerge if the public information makes the players confident enough about the event $\{\theta \in (0, r)\}$. This result is intuitive since in the limit case $\Pr(\theta \in [0, r]) = 1$, we have two pooling equilibria and some intermediate equilibria (partial separating equilibria). This is a strong proposition in the sense that the criterion $\Pr(\theta \in [\epsilon, r - \epsilon]) > 1 - \delta$ is uniform for all absolutely continuous common priors. To make this result comparable to the standard global game results, we establish the following corollary:

Corollary 12.6 *Let $p(\theta)$ be the probability density function of a common prior with finite expectation $E_p(\theta)$ and variance $Var_p(\theta)$, then $\forall r > 0$, $\forall \mu > 0$, $\forall y \in (0, r)$, $\exists \overline{\beta} > 0$ s.t. $\forall \beta > \overline{\beta}$, the costly information acquisition problem with the common prior (density function) $\beta^{1/2} \cdot p\left(\beta^{1/2} \cdot (\theta - y)\right)$ has multiple equilibria.*

Proof. Since $y \in (0, r)$, we can choose $\epsilon > 0$ small enough such that $y \in (\epsilon, r - \epsilon)$. Choose the corresponding $\delta > 0$ as suggested in Proposition 12. Let $E_{p,\beta}(\theta)$ and $Var_{p,\beta}(\theta)$ be the expectation and variance corresponding to the density $\beta^{1/2} \cdot p\left(\beta^{1/2} \cdot (\theta - y)\right)$, respectively. Let $\Pr(\cdot, \beta)$ denote the probability measure induced by the density $\beta^{1/2} \cdot p\left(\beta^{1/2} \cdot (\theta - y)\right)$ and $d = \frac{1}{2} \cdot \min\{y - \epsilon, r - \epsilon - y\}$. Since

$$\begin{aligned}
& \lim_{\beta \to \infty} E_{p,\beta}(\theta) \\
= & \lim_{\beta \to \infty} \int_{-\infty}^{\infty} \theta \cdot \beta^{1/2} \cdot p\left(\beta^{1/2} \cdot (\theta - y)\right) \cdot d\theta \\
= & \lim_{\beta \to \infty} \int_{-\infty}^{\infty} \left(\beta^{-1/2} \cdot \widetilde{\theta} + y\right) \cdot p\left(\widetilde{\theta}\right) \cdot d\widetilde{\theta} \\
= & \lim_{\beta \to \infty} \left[y + \beta^{-1/2} \cdot E_p(\theta)\right] \\
= & \; y \in (\epsilon, r - \epsilon)
\end{aligned}$$

there exists $\beta' > 0$, s.t. $\forall \beta > \beta'$, $E_{p,\beta}(\theta) \in (y - d, y + d)$. Also note that $\lim_{\beta \to \infty} Var_{p,\beta}(\theta) = \lim_{\beta \to \infty} \beta^{-1} \cdot Var_p(\theta) = 0$, thus $\exists \overline{\beta} > \beta'$ s.t. $\forall \beta > \overline{\beta}$, $Var_{p,\beta}(\theta) < 4 \cdot d^2 \cdot (1 - \delta)$. Finally we have $\forall\beta > \overline{\beta}$,

$$\begin{aligned} & \left[E_{p,\beta}(\theta) - \frac{1}{2}\cdot(1-\delta)^{-1/2}\cdot\left[Var_{p,\beta}(\theta)\right]^{1/2}, E_{p,\beta}(\theta) + \frac{1}{2}\cdot(1-\delta)^{-1/2}\cdot\left[Var_{p,\beta}(\theta)\right]^{1/2}\right] \\ \subset\ & \left[E_{p,\beta}(\theta) - d, E_{p,\beta}(\theta) + d\right] \\ \subset\ & \left[y - 2\cdot d, y + 2\cdot d\right] \\ \subset\ & \left[y - \min\{y-\epsilon, r-\epsilon-y\}, y + \min\{y-\epsilon, r-\epsilon-y\}\right] \\ \subset\ & [\epsilon, r-\epsilon] \end{aligned}$$

thus

$$\begin{aligned} & \Pr\left(\{\theta \in [\epsilon, r-\epsilon]\}, \beta\right) \\ \geq\ & \Pr\left(\left\{\theta \in \left[E_{p,\beta}(\theta) - \frac{1}{2}\cdot(1-\delta)^{-1/2}\cdot\left[Var_{p,\beta}(\theta)\right]^{1/2}, E_{p,\beta}(\theta) + \frac{1}{2}\cdot(1-\delta)^{-1/2}\cdot\left[Var_{p,\beta}(\theta)\right]^{1/2}\right]\right\}, \beta\right) \\ \geq\ & 4\cdot(1-\delta) > (1-\delta) \end{aligned}$$

where the second inequality comes from Chebyshev's inequality. Then the result directly follows Proposition 12. ■

Here β measures the precision of public information. Corollary 12.6 says that providing public information of high precision facilitates the multiplicity. This is consistent with the well known result in the global games literature. Another well known result is that the uniqueness is guaranteed if the private signals are sufficiently accurate relative to the public signals (e.g. Morris and Shin (2004)). In other words, the effects on the uniqueness of increasing the precision of public signals can be offset by increasing the precision of private signals. Based on Corollary 11.5, in the semi-endogenous information structure model the effect of increasing the precision of public signals can be offset by lowering the cost of private information acquisition. In our endogenous information model, however, if the density of the common prior has a full support, Proposition 10 states that there are always infinitely many equilibria when $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$, regardless of the precision of public information, i.e. the effects of public information and private information acquisiiton are disentangled. The reason is that when the cost of information acquisition is samll, the players have enough freedom in allocating their attention, which in turn improves their coordination. Also this freedom has nothing to do with the public information. The entangled effects in the previous models actually comes from the rigidity imposed on the information structure.

In the rest of this section, we numerically solve the model for various common priors to see the effects of public information. In order to exclude the possible indeterminancy caused by the multiple equilibrium choices of private information acquisition, we conduct the numerical analysis under the the condition $\widetilde{r} = \frac{r}{4\cdot\mu} \leq 1$. Thus there is a unique shape m satisfying

$\theta = \mu \cdot \ln\left(\frac{m(\theta)}{1-m(\theta)}\right) + r \cdot \left(\frac{1}{2} - m(\theta)\right)$ and an equilibrium $m(\theta - \theta_0)$ is determined by its position paramter θ_0. In other words, whether there is multiplicity only depends on the public information. Specifically, for $r = 2.8$ and $\mu = 0.8$, we calculate all the equilibrium position θ_0's for normal common priors with mean and standard deviation ranging over $[-1, 3.5]$ and $[0.02, 2.5]$, respectively. These ranges are large enough to cover all the possible situations, as shown in the figures below. To make it applicable for the numerical analysis, we use the trancated normal distributions and thus pooling equilibria may also exist[14].

For the trancated normal common priors, generally speaking, there may be one, three or five equilibrium θ_0's. The stable and unstable equilibria always appear alternately and the number of the stable equilibria is always one more than that of the unstable equilibria. Here an equilibrium θ_0 is stable if it has a neighborhood such that any deviation within it approaches θ_0 in the long run.

We plot the equilibrium θ_0 as a function of ("mean", "standard deviation") of the common prior. However, it is hard to plot the multiple equilibria in one graph. To preserve as much information as possible, we plot the graph according to the following rule: for any pair ("mean", "standard deviation"), if there is a unique equilibrium θ_0, just plot it and record it as globally stable; if there are three equilibria $\theta_0^1 < \theta_0^2 < \theta_0^3$, plot θ_0^2 and record it as unstable; if there are five equilibria $\theta_0^1 < \theta_0^2 < \theta_0^3 < \theta_0^4 < \theta_0^5$, plot θ_0^3 and record it as locally stable with neighborhood $\left(\theta_0^2, \theta_0^4\right)$. Note that in the last case, θ_0^2 and θ_0^4 are unstable equilibria and $\theta_0^4 - \theta_0^2$ is a stability coefficient for the locally stable equilibrium θ_0^3. The larger is this coefficient, the more stable is the corresponding equilibrium. As a natural extension, the stability coefficient is zero for the unstable equilibria and is infinity for the globally stable equilibria. Besides plotting θ_0 against each pair of ("mean", "standard deviation"), we also use a colored plane to indicate the stability coefficient of the corresponding θ_0. Higher coefficient is represented by the color close to the red end of the color bar and lower coefficient is represented by the color close to the blue end. The stability coefficient of the globally stable equilibria is ∞, which cannot be directly shown in the graph. Here we use a large number to represent this ∞. Recall that $\theta_0 = \infty$ ($\theta_0 = -\infty$) means a pooling in N (I) equilibrium and a parital separating equilibrium with θ_0 large (negatively large) enough is closed to the pooling in N (I) equilibrium. These θ_0's are too large (negatively large) to plot and we trancate them to the same large number representing the pooling equilibria.

The graph is shown by the following figures from four directions as if it is rotated counterclockwise.

[14] Note that normal distributions have full support and thus there cannot exist any pooling equilibrium according to Proposition 8.

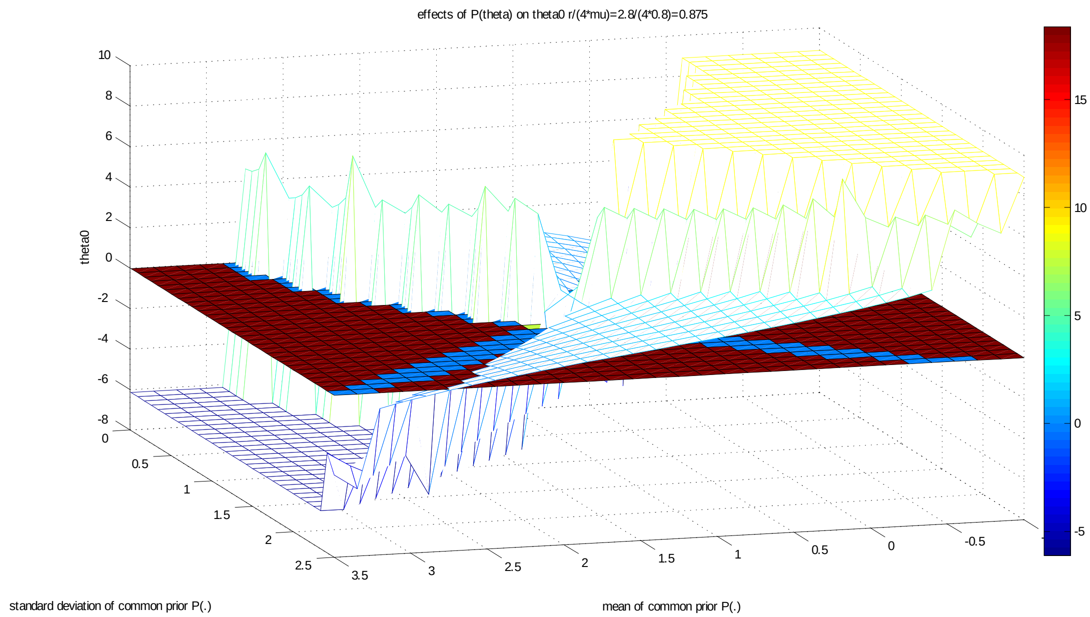

figure 09

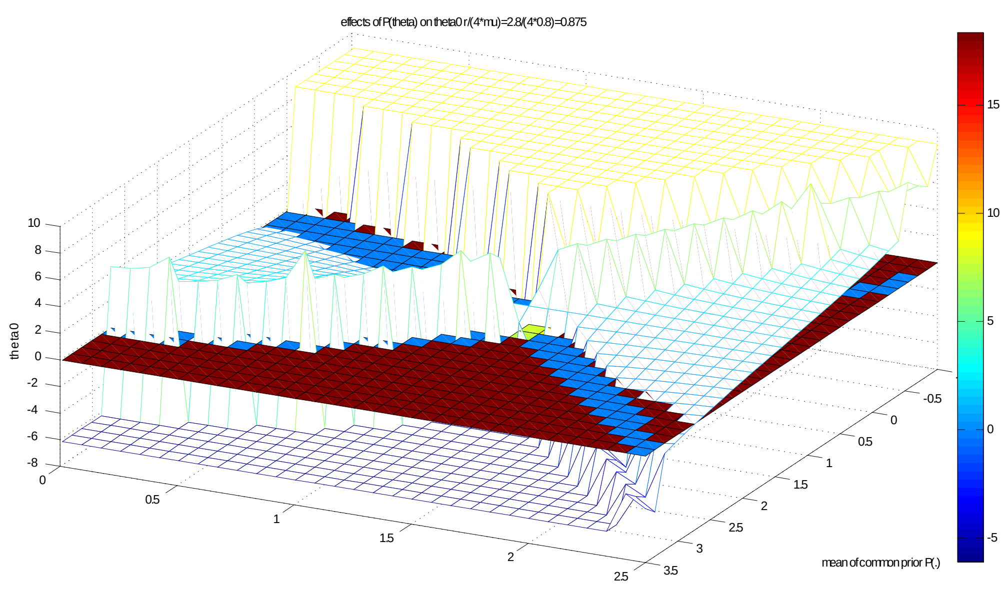

figure 10

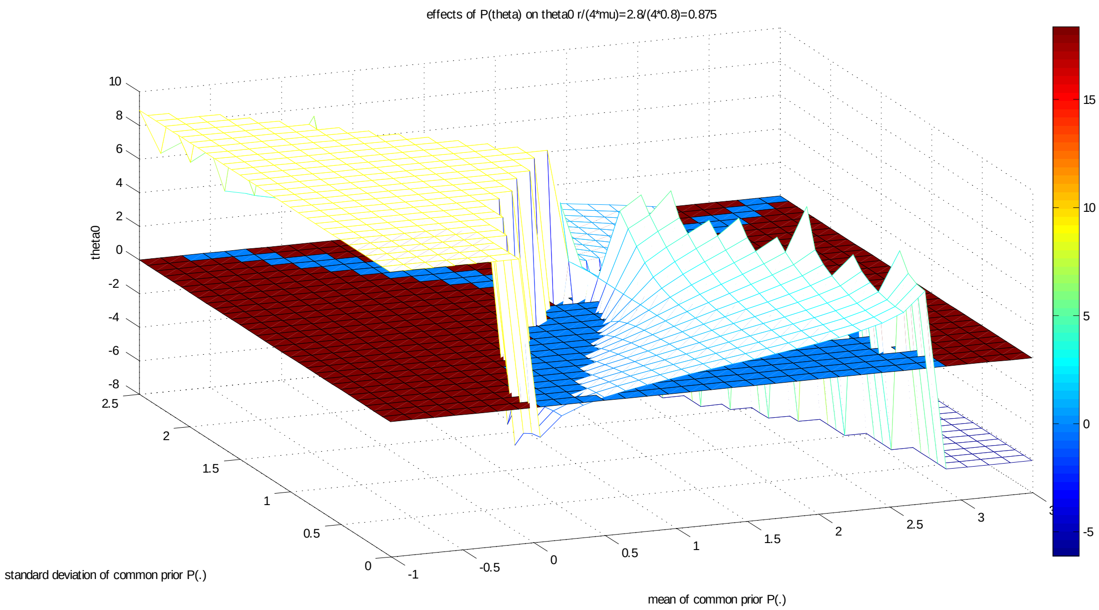

figure 11

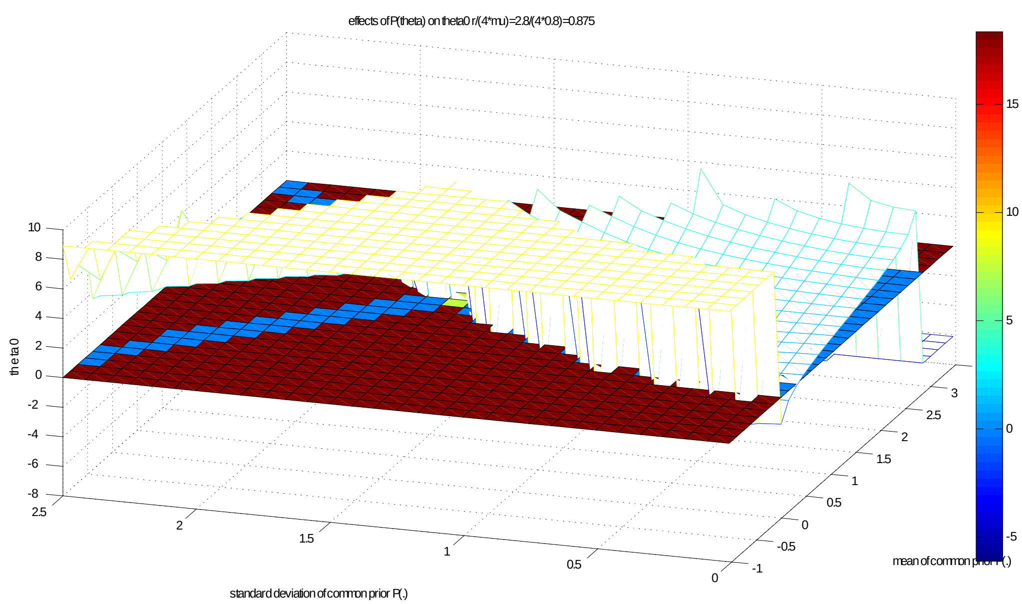

figure 12

The plane $\theta_0 = 0$ indicates the stability of equilibria under different compositions of "mean" and "standard deviation" of the common prior. The dark-red area corresponds to the globally stable equilibria (which are also the unique equilibria for the corresponding "mean" and "standard deviation"); the blue area and the green area correspond to the unstable equilibria and the locally stable equilibria, respectively. We see that the whole plane is almost occupied by the stable and unstable area, and the locally stable equilibria only exist for intermediate values of "mean" and "standard deviation". Both the stable area and the unstable area are disconnected, which implies that moving from one type[15] of stable equilibrium to another type must experience some jumps. In the area of stable equilibria, θ_0 decreases in the mean of the common prior, i.e. a better public signal induces a higher probability of investing. This is intuitive. The opposite happens in the area of unstable equilibria. It means nothing since the unstable equilibria cannot exist in practice. This graph provides some guidance for the strategic disclosure of public information, which can be discussed in future research.

5 CONCLUSION

We endogenize the information structure of a global game model to avoid the arbitrariness in choosing the information structure. Since this model generates distinct results, it might be interesting to apply it to the usual problems like currency attacks, debt pricing and bank run to see if there is any new implications. To attack these more applied topics, we need to allow r, the strategic complementarity, to vary with the fundamental, i.e. $r = r(\theta)$. Fortunately, most results of this paper can be easily extended to cope with this new setup. Dynamic endogenous information acquisition models might be another interesting direction, especially the propaganda of a new standard or techonology. Our analysis of stability of the equilibria in Subsection 4.3 is a preliminary attempt to address this problem, but a full understanding calls for a real dynamic model of endogenous information acquisition.

References

[1] Carlsson, H. and E. van Damme. (1993). Global Games and Equilibrium Selection. *Econometrica* 61 989–1018.

[2] Goldstein, I. and A. Pauzner. (2005). Demand-Deposit Contracts and the Probability of Bank Runs. *Journal of Finance* Vol.60, No.3 1293-1327.

[3] Mackowiak, B. and M. Wiederholt. (2009). Optimal Sticky Prices under Rational Inattention. *American Economic Review* Vol.99(3) 769-803.

[15] Here we divide the stable equilibria into three types: pooling in I, pooling in N and partial separating.

[4] Morris, S. and H. S. Shin. (1998). Unique Equilibrium in a Model of Self-Fulfilling Attacks. *American Economic Review* Vol.88 587-597.

[5] Morris, S. and H. S. Shin. (2003). Global Games: Theory and Applications, in Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society), edited by M. Dewatripont, L. Hansen and S. Turnovsky; Cambridge University Press (2003).

[6] Morris, S. and H. S. Shin. (2004). Coordination Risk and the Price of Debt. *European Economic Review* Vol.48 133-153.

[7] Hellwig, C. and L. Veldkamp. (2009). Knowing What Others Know: Coordination Motives in Information Acquisition. *Review of Economic Studies* Vol.76(1) 223-251, 01.

[8] Sims, Christopher A. (1998). Stickiness. *Carnegie-Rochester Series on Public Policy* Vol.49(1) 317-356.

[9] Sims, Christopher A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics* Vol.50, No.3(April) 665-690.

[10] Sims, Christopher A. (2005). An introduction to rational inattention. Lecture notes for the course "Bounded Rationality and Macroeconomics.

[11] Sims, Christopher A. (2006). Rational inattention: A research agenda. Working paper, Princeton University.

A Basics of Information Theory and Rational Inattention

In Shannon's information theory, information is defined as the reduction of uncertainty, while uncertainty is measured by entropy. For a discrete random vector $\overrightarrow{X}$ with probability weights $p(\overrightarrow{x})$, $\overrightarrow{x} \in \mathbf{X}$, its entropy is

$$H\left(\overrightarrow{X}\right) = -E_{\overrightarrow{x}}\left[\log p\left(\overrightarrow{x}\right)\right] = -\sum_{\overrightarrow{x} \in \mathbf{X}} p\left(\overrightarrow{x}\right) \cdot \log p\left(\overrightarrow{x}\right)$$

, where we define $p(\overrightarrow{x}) \log p(\overrightarrow{x}) = 0$ when $p(\overrightarrow{x}) = 0$. Shannon proves that any function measuring the uncertainty and satisfying three axioms must have this form. Thus it is a natural and objective measurement of uncertainty. The base of the logarithm is not essential, it just changes the unit of entropy. For example, when the base is 2 , the entropy of a discrete random variable with equal probability on two values is 1 bit.

When $\overrightarrow{X} = \left(\overrightarrow{X}_1, \overrightarrow{X}_2\right)$, we also call $H\left(\overrightarrow{X}\right) = H\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right)$ the joint entropy of $\overrightarrow{X}_1$ and $\overrightarrow{X}_2$. The conditional entropy of $\overrightarrow{X}_1$ given $\overrightarrow{X}_2$ is the expected conditional entropy over $\overrightarrow{X}_2$, which is defined as

$$\begin{aligned} H\left(\overrightarrow{X}_1|\overrightarrow{X}_2\right) &= E_{\overrightarrow{X}_2}\left[H\left(\overrightarrow{X}_1|\overrightarrow{x}_2\right)\right] \\ &= -\sum_{\overrightarrow{x}_2 \in \mathbf{X}_2} p\left(\overrightarrow{x}_2\right) \cdot \sum_{\overrightarrow{x}_1 \in \mathbf{X}_1} p\left(\overrightarrow{x}_1|\overrightarrow{x}_2\right) \cdot \log p\left(\overrightarrow{x}_1|\overrightarrow{x}_2\right) \\ &= -\sum_{\overrightarrow{x} \in \mathbf{X}} p\left(\overrightarrow{x}_1, \overrightarrow{x}_2\right) \cdot \log \frac{p\left(\overrightarrow{x}_1, \overrightarrow{x}_2\right)}{p\left(\overrightarrow{x}_2\right)} \end{aligned}$$

. Note that $H\left(\overrightarrow{X}_1|\overrightarrow{X}_2\right) = H\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right) - H\left(\overrightarrow{X}_2\right)$. $H\left(\overrightarrow{X}_1|\overrightarrow{X}_2\right)$ measures the remaining uncertainty of $\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right)$ when $\overrightarrow{X}_2$ is known.

The mutual information between two random vectors measures the amount of information that can be obtained about one random vector when the other one is known. Its definition is

$$I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right) = \sum_{\overrightarrow{x} \in \mathbf{X}} p\left(\overrightarrow{x}_1, \overrightarrow{x}_2\right) \log \frac{p\left(\overrightarrow{x}_1, \overrightarrow{x}_2\right)}{p\left(\overrightarrow{x}_1\right) \cdot p\left(\overrightarrow{x}_2\right)}$$

. Note that

$$\begin{aligned} I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right) &= I\left(\overrightarrow{X}_2; \overrightarrow{X}_1\right) \\ &= H\left(\overrightarrow{X}_2\right) - H\left(\overrightarrow{X}_2|\overrightarrow{X}_1\right) = H\left(\overrightarrow{X}_1\right) - H\left(\overrightarrow{X}_1|\overrightarrow{X}_2\right) \\ &= H\left(\overrightarrow{X}_1\right) + H\left(\overrightarrow{X}_2\right) - H\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right) \end{aligned}$$

. Mutual information is always non-negative. If the above joint probability $p\left(\overrightarrow{x}_1, \overrightarrow{x}_2\right)$ is replaced by the conditional joint probability $p\left(\overrightarrow{x}_1, \overrightarrow{x}_2|\overrightarrow{y}\right)$ on some random vector $\overrightarrow{Y}$, we get the conditional mutual information, which measures the mutual information between two random vectors when the third one is known. Conditional mutual information is also always non-negative.

Some properties of the mutual information: i) given the marginal probability $p\left(\overrightarrow{x}_1\right)$, $I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right)$ is a convex functional with respect to the conditional probability $p\left(\overrightarrow{x}_2|\overrightarrow{x}_1\right)$; ii) given the conditional probability $p\left(\overrightarrow{x}_2|\overrightarrow{x}_1\right)$, $I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right)$ is a concave functional with respect to the marginal probability $p\left(\overrightarrow{x}_1\right)$.

For a continuous random vector, its Shannon entropy is infinity, since it can take a continuum of possible values. In this case the differential entropy is defined as an extension of the Shannon entropy. Formally, the differential entropy of a continuous random vector $\overrightarrow{X}$ with probability density function $p\left(\overrightarrow{x}\right)$ is defined as

$$h\left(\overrightarrow{X}\right) = -\int_{\overrightarrow{X} \in \mathbb{R}^n} p\left(\overrightarrow{x}\right) \cdot \log p\left(\overrightarrow{x}\right) \cdot d\overrightarrow{x}$$

, where n is the dimension of $\overrightarrow{X}$. All the above properties of Shannon entropy still hold for differential entropy. Since the base of the logarithm is not essential, I just use the natural logarithm in the rest of this paper.

Let's look at an example of normal random vectors: let $\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right) \sim N\left(\left(\overrightarrow{\mu}_1, \overrightarrow{\mu}_2\right), \Sigma\right)$, where $\overrightarrow{\mu}_i \in \mathbb{R}^{n_i}$, $i = 1, 2$ and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Then the joint entropy of $\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right)$ is $h\left(\overrightarrow{X}_1, \overrightarrow{X}_2\right) = \frac{n}{2} \cdot [\ln(2\pi) + 1] + \frac{1}{2} \cdot \ln |\Sigma|$, where $n = n_1 + n_2$.

The conditional entropy is $h\left(\overrightarrow{X}_1 | \overrightarrow{X}_2\right) = \frac{n_1}{2} \cdot [\ln(2\pi) + 1] + \frac{1}{2} \cdot \ln |\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$, and the mutual information is $I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right) = -\frac{1}{2} \cdot \ln |I - \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$, where I is the identical matrix with dimension n_1 .

Channel refers to the medium to convey information. The channel between two random vectors $\overrightarrow{X}_1$ and $\overrightarrow{X}_2$ is described by the conditional distribution $p\left(\overrightarrow{x}_1 | \overrightarrow{x}_2\right)$ or $p\left(\overrightarrow{x}_2 | \overrightarrow{x}_1\right)$. Given the unconditional distributions of $\overrightarrow{X}_1$ and $\overrightarrow{X}_2$, different conditional distributions define different channels. Thus a specific channel determines which kind of information about one random vector is conveyed by the other. The mutual information $I\left(\overrightarrow{X}_1; \overrightarrow{X}_2\right)$ measures the amont of information transmitted through the channel between $\overrightarrow{X}_1$ and $\overrightarrow{X}_2$.

The basic idea of rational inattention is that the agent's actions can depend on observations of the state variables only through a channel with finite capacity. Thus he should appropriately allocate his capacity to collect the information most relevent to his objective and ignore others. Specifically, if $\overrightarrow{\theta}$ denotes the state random variables, then the agent can choose an appropriate channel to generate the signals $\overrightarrow{S}$ subject to the constraint that the mutual information $I\left(\overrightarrow{\theta}; \overrightarrow{S}\right)$ is upper bounded by his capacity. For more about rational inattention, please see Sims (2003 and 2005).

B TECHNICAL PROOFS

Proof of Lemma 01.

Proof. suppose player i collects information in a way described by $q_i\left(s^i|\theta\right)$, $s^i \in S^i$. Construct a new strategy with three possible realizations of the signal $\left\{s_I^i, s_N^i, s_{ind}^i\right\}$ such that $\forall\theta \in$supp$(P(\cdot))$, $q_i\left(s_I^i|\theta\right) = \int_{S_I^i} q_i\left(s^i|\theta\right) ds^i$, $q_i\left(s_N^i|\theta\right) = \int_{S_N^i} q_i\left(s^i|\theta\right) ds^i$ and $q_i\left(s_{ind}^i|\theta\right) =$

$\int_{S^i_{ind}} q_i\left(s^i|\theta\right) ds^i$. Player i chooses I when $s^i = s^i_I$, chooses N when $s^i = s^i_N$ and is indifferent when $s^i = s^i_{ind}$. Obviously this modification does not change i's expected utility. However, if the original S^i has more than three elements, this new strategy costs strictly less mutual information. According to the preference assumption (A2), $q_i\left(s^i|\theta\right)$ is suboptimal and cannot be an equilibrium strategy. Thus we proved that $\#\left(S^i\right) \leq 3$, $\forall i \in \{1,2\}$.

Suppose $\exists i \in \{1,2\}$, $\Pr\left(S^i_{ind}\right) > 0$, i.e. $\Pr\left(s^i_{ind}\right) > 0$, since we proved that in the equilibrium $S^i_{ind} = \left\{s^i_{ind}\right\}$.

If $\Pr\left(s^i_{ind}\right) \in (0,1)$, then $\Pr\left(s^i_I\right) > 0$ or $\Pr\left(s^i_N\right) > 0$. Without loss of generality, let $\Pr\left(s^i_I\right) > 0$. Replace the two realizations s^i_{ind} and s^i_I by a new one $\widetilde{s}^i_I$ such that $\forall\theta \in$supp$(P(\cdot))$, $q_i\left(\widetilde{s}^i_I|\theta\right) = q_i\left(s^i_I|\theta\right) + q_i\left(s^i_{ind}|\theta\right)$. Let player i choose I when receiving $\widetilde{s}^i_I$. Obviously this modification does not change i's expected utility. Since $\Pr\left(s^i_I\right) > 0$ and $\Pr\left(s^i_{ind}\right) > 0$, this new strategy costs strictly less mutual information. According to the preference assumption (A2), $q_i\left(s^i|\theta\right)$ is suboptimal and cannot be an equilibrium strategy.

Now we know that $\Pr\left(s^i_{ind}\right) = 1$, i.e. player j's strategy makes player i always indifferent between I and N. Thus, $\exists m_i \in [0,1]$, s.t. player i always invests with probability m_i and does not collect any information. Let $m_i(\theta) \triangleq \Pr(\text{player } i \text{ chooses } I|\theta)$ be the probability that player $i \in \{1,2\}$ invests when the realization of the fundamental is θ. Then $m_i(\cdot)$ is totally determined by player i's strategy $\left(q_i\left(s^i|\theta\right), \sigma^i(\cdot)\right)$, $i \in \{1,2\}$. If $\Pr\left(\theta - r\cdot(1-m_j(\theta)) \neq 0\right) > 0$, e.g. $\Pr\left(\theta - r\cdot(1-m_j(\theta)) > 0\right) > 0$, then player i can always benefit from using some more capacity[16] to distinguish some non-zero probability event $A \subset \{\theta - r\cdot(1-m_j(\theta)) > 0\}$ and investing when this event happens. Thus we have $\Pr(m_i(\theta) = m_i) = 1$ and $\Pr\left(\theta - r\cdot(1-m_j(\theta)) = 0\right) = 1$. Let $F_+ = \{\theta \in \text{supp}(P(\cdot))\,|\theta - r\cdot(1-m_i) > 0\}$, $F_- = \{\theta \in \text{supp}(P(\cdot))\,|\theta - r\cdot(1-m_i) < 0\}$ and $B = \{\theta \in \text{supp}(P(\cdot))\,|\theta - r\cdot(1-m_i) = 0\}$. Construct a new strategy $\left(\widetilde{q}_j\left(s^j|\theta\right), \widetilde{\sigma}^j(\cdot)\right)$, $S^j = \left\{s^j_I, s^j_N\right\}$ for player j, s.t. $\widetilde{q}_j\left(s^j = s^j_I|\theta\right) = 1$ if $\theta \in F_+ \cup B$, $\widetilde{q}_j\left(s^j = s^j_N|\theta\right) = 1$ if $\theta \in F_-$, and $\widetilde{\sigma}^j\left(s^j_I\right) = I$, $\widetilde{\sigma}^j\left(s^j_N\right) = N$. Then this new strategy brings player j an expected utility no less than that gained with the original strategy and uses a capacity no more than that used in the original strategy. Moreover, the non-triviality assumption implies $\Pr(B) < 1$, thus $\Pr(F_+) + \Pr(F_-) > 0$. Without loss of generality, assume $\Pr(F_+) > 0$. $\forall\theta \in F_+$, $\theta - r\cdot(1-m_i) > 0 = \theta - r\cdot(1-m_j(\theta))$, i.e. $m_j(\theta) < m_i \leq 1$. Thus the new strategy generates at least $\int_{F_+}(1-m_j(\theta))\,dP(\theta) > 0$ amount of extra expected utility for player j. Therefore, $\left(\widetilde{q}_j\left(s^j|\theta\right), \widetilde{\sigma}^j(\cdot)\right)$ strictly dominates $\left(q_j\left(s^j|\theta\right), \sigma^j(\cdot)\right)$ and thus $\left(q_j\left(s^j|\theta\right), \sigma^j(\cdot)\right)$ cannot be player j's equilibrium strategy.

[16] This is feasible since player i's current strategy uses zero capacity and his capacity constraint is $\kappa > 0$.

Now we have proved $\Pr\left(S^i_{ind}\right) = 0$, $\forall i \in \{1,2\}$ by contradiction. Combined with the previous result that $\#\left(S^i\right) \leq 3$, $\forall i \in \{1,2\}$, it also implies that $\#\left(S^i\right) = 1$ or 2. ■

Proof of Lemma 02.

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument.

(Sufficiency) If $\Pr(\theta \geq 0) = 1$, both players pooling in I is obvious an equilibrium.

(Necessity) Suppose player i pools in I, i.e. $\forall\theta \in$supp$(P(\cdot))$, $m_i(\theta) = 1$, which implies $\lambda_i(\theta) \geq 0$, $\eta_i(\theta) = 0$, $p_{Ii} = m_i(\theta) = 1$ and thus the first term of the right hand side of (3.7) vanishes. Then (3.7) becomes

$$\forall\theta \in \text{supp}(P(\cdot)),\ \theta - r\cdot(1 - m_j(\theta)) = \lambda_i(\theta) \geq 0$$

i.e.

$$\forall\theta \in \text{supp}(P(\cdot)),\ \theta \geq r\cdot(1 - m_j(\theta)) \geq 0$$

i.e. $\Pr(\theta \geq 0) = 1$. ■

Lemma 04 *in an equilibrium of the constrained information acquisition problem with one player pooling in I (N), the other player must also pool in I (N).*

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument. Suppose player i pools in I, then Lemma 02 implies $\Pr(\theta \geq 0) = 1$. By the non-triviality assumption, $\Pr(\theta = 0) < 1$. Thus pooling in I always generates a non-negative utility for player j and any other strategy either reduces his expected utility or incurs non-zero attention. According to the preference assumption A2, player j must pool in I. ■

Lemma 05 *in an equilibrium of the constrained information acquisition problem with one player playing the perfect separating strategy, the other player must play the same strategy.*

Proof. Suppose player i plays a perfect separating strategy, i.e. player i partitions supp$(P(\cdot))$ into two positive-probability events and choose different actions upon the occurence of different events. This strategy can be totally characterized by the event that player i invests: $S^i_I \triangleq \{\theta \in \text{supp}(P(\cdot)) : m_i(\theta) = 1\}$ (note that $\Pr(m_i(\theta) = 1) + \Pr(m_i(\theta) = 0) = 1$). We first find out player j's best response $m_j(\cdot)$ and then verify its feasibility. As an equilibrium strategy, S^i_I satisfies $\{\theta \in \text{supp}(P(\cdot)) : \theta - r > 0\} \subset S^i_I \subset \{\theta \in \text{supp}(P(\cdot)) : \theta \geq 0\}$. Thus $\forall\theta \in S^i_I \backslash \{\theta = 0\}$, player j must choose $m_j(\theta) = 1$ to enjoy a positive conditional expected utility $\theta > 0$; $\forall\theta \in$supp$(P(\cdot)) \backslash S^i_I \backslash \{\theta = r\}$, player j must choose $m_j(\theta) = 0$ to avoid a negative conditional expected utility $\theta - r < 0$. If $\Pr(\theta = 0) = \Pr(\theta - r = 0) = 0$, then the above argument leads to $\Pr(m_i(\theta) = m_j(\theta)) = 1$. Now suppose $\Pr(\theta = 0) > 0$. If $0 \in S_I^i$ but $m_j(0) < 1$, then player i's conditional expected utility by choosing $m_i(0) = 1$ is $0 - r \cdot (1 - m_j(0)) < 0$, thus $\theta = 0$ should not belong toS_I^i, which is a contradition. Therefore $0 \in S_I^j$ if $0 \in S_I^i$. By the same argument we know that when $\Pr(\theta = r) > 0$, if S_I^i does not contain $\{\theta = r\}$, S_I^j does not either. Thus we showed that $\Pr(m_i(\theta) = m_j(\theta)) = 1$. Since both players have the same capacity, player j's strategy is feasible. ■

Lemma 06 *in an equilibrium of the constrained information acquisition problem with one player playing the partial separating strategy, the other player must play the same strategy.*

Proof. Suppose player i plays a partial separating strategy $m_i(\cdot)$, i.e. $\Pr(m_i(\theta) \in (0,1)) > 0$.

Recall that an equilibrium is characterized by

$$\begin{aligned} \forall\theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m_j(\theta)) &= \mu_i \cdot \left[\ln\left(\frac{m_i(\theta)}{1 - m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right)\right] + \lambda_i(\theta) - \eta_i(\theta) \quad i, j \in \{1,2\}, i \neq j \quad (3.7) \end{aligned}$$

where

$$p_{Ii} = \Pr(\text{player } i \text{ chooses } I) = \int m_i(\theta) \cdot dP(\theta) \qquad i \in \{1,2\} \quad (3.6)$$

We first show that $\mu_i = 0$ implies $\mu_j = 0$.

Suppose $\mu_i = 0$ but $\mu_j > 0$. $\mu_i = 0$ implies

$$\begin{aligned} \forall\theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m_j(\theta)) &= \lambda_i(\theta) - \eta_i(\theta) \qquad (B.1) \end{aligned}$$

If player j's equilibrium strategy is pooling or perfect separating, his capacity constraint must be slack and $\mu_j = 0$. Thus $\mu_j > 0$ implies that player j uses partial separating strategy, i.e. $\Pr(m_j(\theta) \in (0,1)) > 0$, which suggests that $p_{Ij} = \int m_j(\theta) \cdot dP(\theta) \in (0,1)$. Combind with the fact that $\mu_j > 0$, we must have $\Pr(m_j(\theta) \in (0,1)) = 1$. Otherwise the right handside of (3.7) for player j would be infinite with non-zero probability. Then player j's equilibrium strategy $m_j(\cdot)$ satisfies the following equation

$$\begin{aligned} \forall\theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r \cdot (1 - m_i(\theta)) &= \mu_j \cdot \left[\ln\left(\frac{m_j(\theta)}{1 - m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1 - p_{Ij}}\right)\right] \qquad (B.2) \end{aligned}$$

Let $A_+ = \{\theta \in \operatorname{supp}(P(\cdot)) \,|\, \theta - r \cdot (1 - m_j(\theta)) > 0\}$, $A_- = \{\theta \in \operatorname{supp}(P(\cdot)) \,|\, \theta - r \cdot (1 - m_j(\theta)) < 0\}$ and $A_0 = \operatorname{supp}(P(\cdot)) \setminus (A_+ \cup A_-) = \{\theta \in \operatorname{supp}(P(\cdot)) \,|\, \theta - r \cdot (1 - m_j(\theta)) = 0\}$. Then $(B.1)$ implies that

$$m_i(\theta) = \left\{ \begin{array}{ll} 1 & \text{if } \theta \in A_+ \\ \in [0,1] & \text{if } \theta \in A_0 \\ 0 & \text{if } \theta \in A_- \end{array} \right\}$$

Now construct a new strategy $\widetilde{m}_j(\theta)$ for player j, s.t. $\forall\theta \in$supp$(P(\cdot))$, $\widetilde{m}_j(\theta) = m_i(\theta)$. Note that $\widetilde{m}_j(\cdot)$ is feasible for player j since the two players have the same capacity. With this new strategy, $\forall\theta \in A_+$, player j's conditional expected utility becomes $\widetilde{m}_j(\theta)\cdot[\theta - r\cdot(1-m_i(\theta))] = 1\cdot[\theta - r\cdot(1-1)] = \theta > m_j(\theta)\cdot\theta = m_j(\theta)\cdot[\theta - r\cdot(1-m_i(\theta))]$ ="player j's conditional expected utility with his original strategy $m_j(\theta)$". The inequality is true since $\theta > r\cdot(1-m_j(\theta)) > 0$ and we proved $\Pr(m_j(\theta)\in(0,1)) = 1$. Similarly, we can show that $\forall\theta \in A_-$, player j's conditional expected utility with his new strategy strictly exceeds that generated by his original strategy. $\forall\theta \in A_0$, if $\theta - r\cdot(1-m_i(\theta)) > 0 = \theta - r\cdot(1-m_j(\theta))$, then $\widetilde{m}_j(\theta) = m_i(\theta) > m_j(\theta)$ and player j strictly benefits from change from $m_j(\theta)$ to $\widetilde{m}_j(\theta)$; if $\theta - r\cdot(1-m_i(\theta)) < 0 = \theta - r\cdot(1-m_j(\theta))$, then $\widetilde{m}_j(\theta) = m_i(\theta) < m_j(\theta)$ and player j also strictly benefits from change from $m_j(\theta)$ to $\widetilde{m}_j(\theta)$. Now we prove that $\widetilde{m}_j(\cdot)$ generates strictly larger ex ante expected utility than does $m_j(\cdot)$. If not, we must have $\Pr(A_0) = 1$ and $\Pr(m_i(\theta) = m_j(\theta)) = 1$. Recall that by definition, $\forall\theta \in A_0$, $\theta - r\cdot(1-m_j(\theta)) = 0$, thus $\theta - r\cdot(1-m_i(\theta)) = 0$ and $(B.2)$ implies $\mu_j\cdot\left[\ln\left(\frac{m_j(\theta)}{1-m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right] = 0$, i.e. $\forall\theta \in A_0$, $p_{Ij} = m_j(\theta) = 1-\theta/r$. Since $\Pr(A_0) = 1$, this implies #supp$(P(\cdot)) = 1$, which is a contradiction to the non-triviality assumption. Therefore, we proved that $\widetilde{m}_j(\cdot)$ is a feasible strategy strictly dominating $m_j(\cdot)$, which contradicts the fact that $m_j(\cdot)$ is player j's equilibrium strategy. This final contradiction shows that either $\mu_i = \mu_j = 0$ or $\mu_i > 0$ and $\mu_j > 0$.

i) the case of $\mu_i > 0$ and $\mu_j > 0$. Note that $\mu_i > 0$ and $\Pr(m_i(\theta)\in(0,1)) > 0$ implies that $\Pr(m_j(\theta)\in(0,1)) = 1$. Otherwise the right handside of (3.7) for player i would be infinite with non-zero probability. If player j's equilibrium strategy is pooling or perfect separating, his capacity constraint must be slack and $\mu_j = 0$. Thus $\mu_j > 0$ implies that player j uses partial separating strategy, i.e. $\Pr(m_j(\theta)\in(0,1)) > 0$, then the same argument leads to $\Pr(m_j(\theta)\in(0,1)) = 1$. Now (3.7) becomes

$$\begin{aligned} \forall\theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r\cdot(1-m_j(\theta)) &= \mu_i\cdot\left[\ln\left(\frac{m_i(\theta)}{1-m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right] \qquad (B.3) \\ \theta - r\cdot(1-m_i(\theta)) &= \mu_j\cdot\left[\ln\left(\frac{m_j(\theta)}{1-m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right] \qquad (B.4) \end{aligned}$$

Suppose $\Pr(m_i(\theta) \neq m_j(\theta)) > 0$. Let $\langle\cdot,\cdot\rangle$ denote the inner product on $L(\mathbb{R},P(\cdot))$, i.e.

$$\forall f_1, f_2 \in L(\mathbb{R},P(\cdot)),\ \langle f_1, f_2\rangle \triangleq \int f_1(\theta)\cdot f_2(\theta)\cdot dP(\theta)$$

First we show that $\langle m_i - m_j, \theta/r - (1-m_i)\rangle > 0$ or $\langle m_j - m_i, \theta/r - (1-m_j)\rangle > 0$. If not, we have

$$\begin{aligned} \langle m_i - m_j, \theta/r - (1-m_i)\rangle &\leq 0 \\ \langle m_i - m_j, \theta/r - (1-m_j)\rangle &\geq 0 \end{aligned}$$

Taking difference of these two inequalities leads to $\langle m_i - m_j, m_j - m_i\rangle \geq 0$, i.e. $\|m_i - m_j\|^2 \leq 0$, thus $\|m_i - m_j\|^2 = 0$, which is a contradiction to $\Pr(m_i(\theta) \neq m_j(\theta)) > 0$.

Without loss of generality, suppose $\langle m_i - m_j, \theta/r - (1-m_i)\rangle > 0$. According to $(B.4)$, we have $\left\langle m_i - m_j, \mu_j \cdot \left[\ln\left(\frac{m_j(\theta)}{1-m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right]\right\rangle > 0$, i.e. $\left\langle m_i - m_j, \left[\ln\left(\frac{m_j(\theta)}{1-m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right]\right\rangle > 0$, since $\mu_j > 0$. Let $\widetilde{m}(\theta) = t \cdot m_i(\theta) + (1-t) \cdot m_j(\theta)$, $t \in [0,1]$, $\forall\theta \in$supp$(P(\cdot))$. Then $\widetilde{m}(\cdot) \in \Omega$ since Ω is convex. If a player plays the strategy characterized by $\widetilde{m}(\cdot)$, the mutual information between his action and the fundamental is

$$I(\widetilde{m}(\cdot)) = \int [\widetilde{m}(\theta)\ln\widetilde{m}(\theta) + (1-\widetilde{m}(\theta))\ln(1-\widetilde{m}(\theta))]\, dP(\theta) - \widetilde{p}_I \ln \widetilde{p}_I - (1-\widetilde{p}_I)\ln(1-\widetilde{p}_I)$$

where $\widetilde{p}_I = \int \widetilde{m}(\theta) \cdot dP(\theta) = t \cdot p_{Ii} + (1-t) \cdot p_{Ij}$.

Note that

$$\frac{dI(\widetilde{m}(\cdot))}{dt}\,|_{t=0} = \left\langle m_i - m_j, \left[\ln\left(\frac{m_j(\theta)}{1-m_j(\theta)}\right) - \ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right]\right\rangle > 0$$

thus $\exists t \in (0,1)$ s.t. $I(\widetilde{m}(\cdot)) > I(m_j(\cdot)) = \kappa$. On the other hand, the mutual information $I(m(\cdot))$ is a convex functional of $m(\cdot)$, thus $I(\widetilde{m}(\cdot)) \leq \kappa$. This generates a contradiction.

Therefore, we must have $\Pr(m_i(\theta) = m_j(\theta)) = 1$.

ii) the case of $\mu_i = \mu_j = 0$. Now (3.7) becomes

$$\begin{aligned} \forall\theta &\in \operatorname{supp}(P(\cdot)), \\ \theta - r\cdot(1-m_j(\theta)) &= \lambda_i(\theta) - \eta_i(\theta) \qquad (B.1) \\ \theta - r\cdot(1-m_i(\theta)) &= \lambda_j(\theta) - \eta_j(\theta) \qquad (B.5) \end{aligned}$$

Let $A_+^j = \{\theta \in \operatorname{supp}(P(\cdot))\,|\theta - r\cdot(1-m_j(\theta)) > 0\}$, $A_-^j = \{\theta \in \operatorname{supp}(P(\cdot))\,|\theta - r\cdot(1-m_j(\theta)) < 0\}$ and A_0^j =supp$(P(\cdot)) \setminus \left(A_+^j \cup A_-^j\right) = \{\theta \in \operatorname{supp}(P(\cdot))\,|\theta - r\cdot(1-m_j(\theta)) = 0\}$, $j \in \{1,2\}$. Note that $\forall\theta \in A_+^j$, $(B.1)$ implies $m_i(\theta) = 1 \geq m_j(\theta)$ and $\theta - r\cdot(1-m_i(\theta)) \geq \theta - r\cdot(1-m_j(\theta)) > 0$, thus by definition $\theta \in A_+^i$ and $(B.5)$ implies $m_j(\theta) = 1$. Then $A_+^j = A_+^i \triangleq A_+$ and $m_j(\theta) = m_i(\theta) = 1$, $\forall\theta \in A_+$. Similarly we can show that $A_-^j = A_-^i \triangleq A_-$ and $m_j(\theta) = m_i(\theta) = 0$, $\forall\theta \in A_-$. This implies that $A_0^j = A_0^i = A_0$. Also note that by definition $\forall\theta \in A_0$, $\theta - r\cdot(1-m_j(\theta)) = 0 = \theta - r\cdot(1-m_i(\theta))$, i.e. $m_j(\theta) = m_i(\theta)$. Since supp$(P(\cdot)) = A_+ \cup A_- \cup A_0$, we prove that $\Pr(m_i(\theta) = m_j(\theta)) = 1$. ■

A remark to the case of $\mu_i = \mu_j = 0$: this case is actually very rare, in the sense that it only exists under very restrictive conditions. Specifically, the seond part of the preference assumption (A2) (do not collect information that is not used..) implies that a partial separating equilibrium with $\mu_i = \mu_j = 0$ exists iff $\exists \theta_B \in (0, r) \cap$supp$(P(\cdot))$ s.t. $\Pr(\theta = \theta_B) \in (0, 1)$, and $\exists F \subset$supp$(P(\cdot)) \setminus \{\theta_B\}$ s.t. a) $\{\theta \in \text{supp}(P(\cdot)) | \theta > r\} \subset F \subset \{\theta \in \text{supp}(P(\cdot)) | \theta \geq 0\}$; b) $\Pr(F) / [1 - \Pr(\theta = \theta_B)] = 1 - \theta_B/r$; c) $\kappa \geq [1 - \Pr(\theta = \theta_B)] \cdot H(1 - \theta_B/r)$; where $\forall p \in [0, 1]$, $H(p) \triangleq -p \cdot \ln p - (1 - p) \cdot \ln(1 - p)$ is the bivariate Shannon entropy function[17]. If such equilibrium exists, it has the form

$$m(\theta) = \left\{ \begin{array}{ll} 1 & \text{if } \theta \in F \\ 1 - \theta_B/r & \text{if } \theta = \theta_B \\ 0 & \text{otherwise} \end{array} \right\}$$

Since it is such a rare case, we just present the above statement without proof. In the main part of the paper, we do not discuss this strange case and when we mention the partial separating equilibria we always mean the case with binding capacity constraints.

Proof of Proposition 2.

Proof. the proof is a direct application of Lemma 04, 05 and 06. ■

Lemma 07 *the costly information acquisition problem has an equilibrium with at least one player pooling in I (N) iff* $\Pr(\theta \geq 0) = 1$ *(*$\Pr(\theta - r \leq 0) = 1$*).*

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument.

(Sufficiency) If $\Pr(\theta \geq 0) = 1$, both players pooling in I is obvious an equilibrium.

(Necessity) Suppose player i pools in I, i.e. $\forall \theta \in$supp$(P(\cdot))$, $m_i(\theta) = 1$, which implies $\lambda_i(\theta) \geq 0$, $\eta_i(\theta) = 0$, $p_{Ii} = m_i(\theta) = 1$ and thus the first term of the right hand side of (4.5) vanishes. Then (4.5) becomes

$$\forall \theta \in \text{supp}(P(\cdot)), \theta - r \cdot (1 - m_j(\theta)) = \lambda_i(\theta) \geq 0$$

i.e.

$$\forall \theta \in \text{supp}(P(\cdot)), \theta \geq r \cdot (1 - m_j(\theta)) \geq 0$$

i.e. $\Pr(\theta \geq 0) = 1$. ■

Lemma 08 *in an equilibrium of the costly information acquisition problem with one player pooling in I (N), the other player must also pool in I (N).*

[17] Note that $H(0) \triangleq \lim_{p \to 0} H(p) = 0$, $H(1) \triangleq \lim_{p \to 1} H(p) = 0$.

Proof. We only prove the case of pooling in I. The case of pooling in N follows the same argument. Suppose player i pools in I, then Lemma 07 implies $\Pr(\theta \geq 0) = 1$. By the non-triviality assumption, $\Pr(\theta = 0) < 1$. Thus pooling in I always generates a non-negative utility for player j and any other strategy either reduces his expected utility or incurs non-zero cost of information acquisition. Thus player j must pool in I. ■

Lemma 09 *in an equilibrium of the costly information acquisition problem with one player playing a partial separating strategy* $m(\cdot)$, *the other player must play the same strategy. Moreover,* $\Pr(m(\theta) \in (0,1)) = 1$.

Proof. Suppose player i plays a partial separating strategy $m_i(\cdot)$, i.e. $\Pr(m_i(\theta) \in (0,1)) > 0$. We first show that $\Pr(m_i(\theta) \in (0,1)) = 1$. Suppose $\Pr(m_i(\theta) \in (0,1)) \in (0,1)$, then $\Pr(m_i(\theta) = 1) + \Pr(m_i(\theta) = 0) > 0$ and $p_{Ii} \in (0,1)$. In the case of $\Pr(m_i(\theta) = 1) > 0$, $\forall\theta \in$supp$(P(\cdot))$ s.t. $m_i(\theta) = 1$, the first term of right hand side of (4.5) becomes infinite and (4.5) cannot hold. $\Pr(m_i(\theta) = 0) > 0$ leads to the same contradiction. Thus $\Pr(m_i(\theta) \in (0,1)) = 1$ and $m_i(\cdot)$ satisfies

$$\begin{aligned} \forall\theta &\in \text{supp}(P(\cdot)), \\ \theta - r\cdot(1 - m_j(\theta)) &= \mu\cdot\left[\ln\left(\frac{m_i(\theta)}{1-m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right] \qquad (B.6) \end{aligned}$$

Suppose player j pools in I, i.e. $\Pr(m_j(\theta) = 1) = 1$, then Lemma 07 suggests that $\Pr(\theta \geq 0) = 1$. Combined with the non-triviality assumption, we have $\Pr(\theta > 0) > 0$. Thus $(B.6)$ becomes

$$\begin{aligned} \forall\theta &\in \text{supp}(P(\cdot)), \\ &\mu\cdot\left[\ln\left(\frac{m_i(\theta)}{1-m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1-p_{Ii}}\right)\right] \\ &= \theta - r\cdot(1 - m_j(\theta)) \\ &= \theta - r\cdot(1-1) \\ &= \theta \geq 0 \quad (B.7) \end{aligned}$$

i.e. $\Pr(m_i(\theta) \geq p_{Ii}) = 1$ and $\Pr(m_i(\theta) > p_{Ii}) > 0$. Therefore $p_{Ii} = \int m_i(\theta)\cdot dP(\theta) > p_{Ii}$, which is a contradiction.

Simillarly, we can exclude the possibility of $\Pr(m_j(\theta) = 0) = 1$. Then according to Lemma 03, player j must play a partial separating strategy, i.e. $\Pr(m_j(\theta) \in (0,1)) > 0$. By the same argument in proving $\Pr(m_i(\theta) \in (0,1)) = 1$, we know that $\Pr(m_j(\theta) \in (0,1)) = 1$. Therefore $m_j(\cdot)$ satisfies

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ \theta - r\cdot\left(1-m_i\left(\theta\right)\right) &= \mu\cdot\left[\ln\left(\frac{m_j\left(\theta\right)}{1-m_j\left(\theta\right)}\right)-\ln\left(\frac{p_{Ij}}{1-p_{Ij}}\right)\right] \qquad (B.8)\end{aligned}$$

Suppose $(m_1(\cdot), m_2(\cdot))$ is a partial separating equilibrium, then it satisfies

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ \theta - r\cdot\left(1-m_1\left(\theta\right)\right) &= \mu\cdot\left[\ln\left(\frac{m_2\left(\theta\right)}{1-m_2\left(\theta\right)}\right)-\ln\left(\frac{p_{I2}}{1-p_{I2}}\right)\right] \qquad (B.9)\\ \theta - r\cdot\left(1-m_2\left(\theta\right)\right) &= \mu\cdot\left[\ln\left(\frac{m_1\left(\theta\right)}{1-m_1\left(\theta\right)}\right)-\ln\left(\frac{p_{I1}}{1-p_{I1}}\right)\right] \qquad (B.10)\end{aligned}$$

$(B.9) - (B.10)$ implies

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ r\cdot\left(m_1\left(\theta\right)-m_2\left(\theta\right)\right) &= \mu\cdot\left[\ln\left(\frac{m_2\left(\theta\right)}{1-m_2\left(\theta\right)}\right)-\ln\left(\frac{m_1\left(\theta\right)}{1-m_1\left(\theta\right)}\right)\right]-\mu\cdot\left[\ln\left(\frac{p_{I2}}{1-p_{I2}}\right)-\ln\left(\frac{p_{I1}}{1-p_{I1}}\right)\right]\end{aligned}$$

i.e.

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ &\left[\ln\left(\frac{p_{I2}}{1-p_{I2}}\right)-\ln\left(\frac{p_{I1}}{1-p_{I1}}\right)\right]\\ &= \left[\ln\left(\frac{m_2\left(\theta\right)}{1-m_2\left(\theta\right)}\right)-\ln\left(\frac{m_1\left(\theta\right)}{1-m_1\left(\theta\right)}\right)\right]+\frac{r}{\mu}\left(m_2\left(\theta\right)-m_1\left(\theta\right)\right) \qquad (B.11)\end{aligned}$$

Note that if $p_{I2} = p_{I1}$, $(B.11)$ becomes

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ 0 &= \left[\ln\left(\frac{m_2\left(\theta\right)}{1-m_2\left(\theta\right)}\right)-\ln\left(\frac{m_1\left(\theta\right)}{1-m_1\left(\theta\right)}\right)\right]+\frac{r}{\mu}\left(m_2\left(\theta\right)-m_1\left(\theta\right)\right)\end{aligned}$$

and we must have $\forall\theta \in$supp$(P(\cdot))$, $m_2(\theta) = m_1(\theta)$ 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 $(B.11)$ becomes

$$\begin{aligned}\forall\theta &\in \operatorname{supp}\left(P\left(\cdot\right)\right),\\ 0 &< z = \left[\ln\left(\frac{m_2\left(\theta\right)}{1-m_2\left(\theta\right)}\right)-\ln\left(\frac{m_1\left(\theta\right)}{1-m_1\left(\theta\right)}\right)\right]+\frac{r}{\mu}\left(m_2\left(\theta\right)-m_1\left(\theta\right)\right) \qquad (B.12)\end{aligned}$$

which suggests that $\Pr\left(m_2\left(\theta\right) > m_1\left(\theta\right)\right) = 1$. Let $\ln\left(\frac{m_2(\theta)}{1-m_2(\theta)}\right) = x(\theta)$ and $\ln\left(\frac{m_1(\theta)}{1-m_1(\theta)}\right) = y(\theta)$. $(B.12)$ implies

$$\forall\theta \in \operatorname{supp}\left(P\left(\cdot\right)\right), x\left(\theta\right)-y\left(\theta\right) = z-\frac{r}{\mu}\left(m_2\left(\theta\right)-m_1\left(\theta\right)\right) < z$$

i.e.

$$\forall\theta\in\text{supp}\left(P\left(\cdot\right)\right),\, x\left(\theta\right)<y\left(\theta\right)+z \qquad (B.13)$$

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

$$\begin{aligned}
z &= \ln\left(\frac{p_{I2}}{1-p_{I2}}\right)-\ln\left(\frac{p_{I1}}{1-p_{I1}}\right)\\
&= \ln\left(\frac{Em_2\left(\theta\right)}{1-Em_2\left(\theta\right)}\right)-\ln\left(\frac{Em_1\left(\theta\right)}{1-Em_1\left(\theta\right)}\right)\\
&= \ln\left(\frac{E\left[\frac{\exp(x(\theta))}{1+\exp(x(\theta))}\right]}{E\left[\frac{1}{1+\exp(x(\theta))}\right]}\right)-\ln\left(\frac{E\left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))}\right]}{E\left[\frac{1}{1+\exp(y(\theta))}\right]}\right)\\
&< \ln\left(\frac{E\left[\frac{\exp(y(\theta)+z)}{1+\exp(y(\theta)+z)}\right]}{E\left[\frac{1}{1+\exp(y(\theta)+z)}\right]}\right)-\ln\left(\frac{E\left[\frac{\exp(y(\theta))}{1+\exp(y(\theta))}\right]}{E\left[\frac{1}{1+\exp(y(\theta))}\right]}\right)
\end{aligned}$$

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

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

i.e.

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

i.e.

$$2\cdot E\left[\frac{\exp\left(y\left(\theta\right)\right)}{1+\exp\left(y\left(\theta\right)+z\right)}\right]\cdot E\left[\frac{1}{1+\exp\left(y\left(\theta\right)\right)}\right]>2\cdot E\left[\frac{1}{1+\exp\left(y\left(\theta\right)+z\right)}\right]\cdot E\left[\frac{\exp\left(y\left(\theta\right)\right)}{1+\exp\left(y\left(\theta\right)\right)}\right]$$

i.e.

$$\begin{aligned}
&\int\frac{\exp\left(y\left(\theta_1\right)\right)}{1+\exp\left(y\left(\theta_1\right)+z\right)}dP\left(\theta_1\right)\cdot\int\frac{1}{1+\exp\left(y\left(\theta_2\right)\right)}dP\left(\theta_2\right)\\
&\quad+\int\frac{\exp\left(y\left(\theta_2\right)\right)}{1+\exp\left(y\left(\theta_2\right)+z\right)}dP\left(\theta_2\right)\cdot\int\frac{1}{1+\exp\left(y\left(\theta_1\right)\right)}dP\left(\theta_1\right)\\
>&\int\frac{1}{1+\exp\left(y\left(\theta_1\right)+z\right)}dP\left(\theta_1\right)\cdot\int\frac{\exp\left(y\left(\theta_2\right)\right)}{1+\exp\left(y\left(\theta_2\right)\right)}dP\left(\theta_2\right)\\
&\quad+\int\frac{1}{1+\exp\left(y\left(\theta_2\right)+z\right)}dP\left(\theta_2\right)\cdot\int\frac{\exp\left(y\left(\theta_1\right)\right)}{1+\exp\left(y\left(\theta_1\right)\right)}dP\left(\theta_1\right)
\end{aligned}$$

i.e.

$$\begin{aligned}
&\int\left[\frac{\exp\left(y\left(\theta_1\right)\right)}{1+\exp\left(y\left(\theta_1\right)+z\right)}\frac{1}{1+\exp\left(y\left(\theta_2\right)\right)}+\frac{\exp\left(y\left(\theta_2\right)\right)}{1+\exp\left(y\left(\theta_2\right)+z\right)}\frac{1}{1+\exp\left(y\left(\theta_1\right)\right)}\right]dP\left(\theta_1\right)dP\left(\theta_2\right)\\
>&\int\left[\frac{1}{1+\exp\left(y\left(\theta_1\right)+z\right)}\frac{\exp\left(y\left(\theta_2\right)\right)}{1+\exp\left(y\left(\theta_2\right)\right)}+\frac{1}{1+\exp\left(y\left(\theta_2\right)+z\right)}\frac{\exp\left(y\left(\theta_1\right)\right)}{1+\exp\left(y\left(\theta_1\right)\right)}\right]dP\left(\theta_1\right)dP\left(\theta_2\right)
\end{aligned}$$

i.e.

$$\int \frac{\begin{array}{c}\exp\left(y\left(\theta_1\right)\right)\left[1+\exp\left(y\left(\theta_2\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_1\right)\right)\right] \\ +\exp\left(y\left(\theta_2\right)\right)\left[1+\exp\left(y\left(\theta_1\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_2\right)\right)\right] \\ -\exp\left(y\left(\theta_2\right)\right)\left[1+\exp\left(y\left(\theta_2\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_1\right)\right)\right] \\ -\exp\left(y\left(\theta_1\right)\right)\left[1+\exp\left(y\left(\theta_1\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_2\right)\right)\right]\end{array}}{\left[1+\exp\left(y\left(\theta_1\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_2\right)\right)\right]\left[1+\exp\left(y\left(\theta_2\right)+z\right)\right]\left[1+\exp\left(y\left(\theta_1\right)\right)\right]} dP\left(\theta_1\right)dP\left(\theta_2\right) > 0 \quad (B.14)$$

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

$$\begin{aligned}\text{numerator} &= e^u\left[1+e^ve^z\right]\left[1+e^u\right]+e^v\left[1+e^ue^z\right]\left[1+e^v\right]-e^v\left[1+e^ve^z\right]\left[1+e^u\right]-e^u\left[1+e^ue^z\right]\left[1+e^v\right] \\ &= \left[e^u-e^v\right]\left[1+e^ve^z\right]\left[1+e^u\right]-\left[e^u-e^v\right]\left[1+e^ue^z\right]\left[1+e^v\right] \\ &= \left[e^u-e^v\right]\left[1+e^ve^z+e^u++e^ue^ve^z-1-e^ue^z-e^v-e^ve^ue^z\right] \\ &= \left[e^u-e^v\right]^2\left[1-e^z\right]<0\end{aligned}$$

where the last inequality follows the fact that $z>0$. Therefore, the left side of $(B.14)$ is strictly negative, which is a contradiction. Therefore, $\Pr\left(m_1\left(\theta\right)=m_2\left(\theta\right)=m\left(\theta\right)\right)=1$ and $\Pr\left(m\left(\theta\right)\in(0,1)\right)=1$. ■

Proof of Proposition 7.

Proof. the proof is a direct application of Lemma 03, 07, 08 and 09. ■

Proof of Proposition 10.

Proof. Since $\widetilde{r}=\frac{r}{4\cdot\mu}>1$, there are infinitely many possible shapes of $m\in M\left(r,\mu\right)$ as shown in Subsection 4.2. Let $m_1=\left(1+\left(1-\widetilde{r}^{-1}\right)^{1/2}\right)/2$ and $m_2=\left(1-\left(1-\widetilde{r}^{-1}\right)^{1/2}\right)/2$, then the upper increasing component is above m_1, the lower increasing component is below m_2 and the decreasing component is within $[m_1,m_2]$. We use $t\in[0,1]$ to index all the shapes satisfying MLRP (i.e. $m\left(\theta\right)$ increasing in θ). Specifically, let $m_t\left(\theta\right)\in[0,m_2]$ if $\theta\in\left(-\infty,\theta_1+t\cdot\left(\theta_2-\theta_1\right)\right]$; $m_t\left(\theta\right)\in[m_1,1]$ if $\theta\in\left(\theta_1+t\cdot\left(\theta_2-\theta_1\right),+\infty\right)$, where $\theta_1=\mu\cdot\ln\left(\frac{1+(1-\widetilde{r}^{-1})^{1/2}}{1-(1-\widetilde{r}^{-1})^{1/2}}\right)-\frac{r}{2}\cdot\left(1-\widetilde{r}^{-1}\right)^{1/2}$ and $\theta_2=\mu\cdot\ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right)+\frac{r}{2}\cdot\left(1-\widetilde{r}^{-1}\right)^{1/2}$ are defined in Corollary 9.4. For example, the corresponding indices for figure 06, 07 and 08 are $t=0$, $1/2$ and 1, respectively.

Define $\widehat{p}_I\left(\theta_0\right)$ by (4.7), i.e.

$$\theta_0=r/2-\mu\cdot\ln\left(\frac{\widehat{p}_I\left(\theta_0\right)}{1-\widehat{p}_I\left(\theta_0\right)}\right)$$

then $\widehat{p}_I\left(\theta_0\right)$ is a continuous function of θ_0. Note that $\left(\theta_0,m\right)\in\mathbb{R}\times M\left(r,\mu\right)$ is a partial separating equilibrium iff $\int m\left(\theta-\theta_0\right)\cdot dP\left(\theta\right)=\widehat{p}_I\left(\theta_0\right)$.

Define $\widetilde{p}_I\left(\theta_0,t\right)\triangleq\int m_t\left(\theta-\theta_0\right)\cdot dP\left(\theta\right)$. Since $P\left(\cdot\right)$ is absolutely continuous, $\widetilde{p}_I\left(\theta_0,t\right)$ is a continuous function with respect to θ_0 and t (even if $m_t\left(\theta-\theta_0\right)$ is not continuous in θ).

Also note that $\text{supp}(P(\cdot)) = \mathbb{R}$ implies that $\widetilde{p}_I(\theta_0, t)$ is strictly decreasing in t, thus $\forall p_I \in [\widetilde{p}_I(\theta_0, 1), \widetilde{p}_I(\theta_0, 0)]$, there exists some $t \in [0,1]$ s.t. $p_I = \int m_t(\theta - \theta_0) \cdot dP(\theta)$. On the other hand, for any possible shape $m \in M(r, \mu)$, no matter whether it satisfies MLRP or not, we have $p_I = \int m(\theta - \theta_0) \cdot dP(\theta) \in [\widetilde{p}_I(\theta_0, 1), \widetilde{p}_I(\theta_0, 0)]$. Hence, the costly information acquisition problem has a partial separating equilibrium iff $\exists \theta_0 \in \mathbb{R}$ s.t. $\widehat{p}_I(\theta_0) \in [\widetilde{p}_I(\theta_0, 1), \widetilde{p}_I(\theta_0, 0)]$. According to Proposition 6 and 8, there exists a partial separating equilibrium $(\theta_0^*, m^*) \in \mathbb{R} \times M(r, \mu)$ since $\Pr(\theta > r) > 0$ and $\Pr(\theta < 0) > 0$. This implies $\widehat{p}_I(\theta_0^*) \in [\widetilde{p}_I(\theta_0^*, 1), \widetilde{p}_I(\theta_0^*, 0)]$.

If $\widehat{p}_I(\theta_0^*) \in (\widetilde{p}_I(\theta_0^*, 1), \widetilde{p}_I(\theta_0^*, 0))$, there exists $\delta > 0$ s.t. $\widehat{p}_I(\theta_0^* + \delta) \in (\widetilde{p}_I(\theta_0^* + \delta, 1), \widetilde{p}_I(\theta_0^* + \delta, 0))$, since $\widehat{p}_I(\theta_0)$, $\widetilde{p}_I(\theta_0, 1)$ and $\widetilde{p}_I(\theta_0, 0)$ are continuous in θ_0. Then $\exists t \in [0,1]$ s.t. $p_I = \int m_t(\theta - \theta_0^* - \delta) \cdot dP(\theta) = \widehat{p}_I(\theta_0^* + \delta)$. Since any possible shape $m \in M(r, \mu)$ is not invariant with respect to translation, this suggests that $(\theta_0^* + \delta, m_t) \in \mathbb{R} \times M(r, \mu)$ is another partial separating equilibrium.

Now we consider the boundary situations $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*, 1)$ and $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*, 0)$.

Suppose $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*, 1)$ and (θ_0^*, m_1) is the unique equilibrium, then $\forall \theta_0 \in \mathbb{R} \backslash \{\theta_0^*\}$, $\widehat{p}_I(\theta_0^*) < \widetilde{p}_I(\theta_0^*, 1) < \widetilde{p}_I(\theta_0^*, 0)$, since $\widehat{p}_I(\theta_0)$, $\widetilde{p}_I(\theta_0, 1)$ and $\widetilde{p}_I(\theta_0, 0)$ are continuous in θ_0. This suggests that for any shape $m \in M(r, \mu)$ and any $\theta_0 < \theta_0^*$, $\widehat{p}_I(\theta_0) < \int m(\theta - \theta_0) \cdot dP(\theta)$ and thus $g(\theta_0, m) = r/2 - \mu \cdot \ln\left(\frac{\int m(\theta - \theta_0) \cdot dP(\theta)}{1 - \int m(\theta - \theta_0) \cdot dP(\theta)}\right) < r/2 - \mu \cdot \ln\left(\frac{\widehat{p}_I(\theta_0)}{1 - \widehat{p}_I(\theta_0)}\right) = \theta_0$. Therefore $\theta_0 = -\infty$ is a fixed point of the mapping $g(\theta_0, m)$ and there exists an equilibrium pooling in I. This is a contradiction to the assumption $\Pr(\theta < 0) > 0$ according to Proposition 8. In otherwords, there must be multiple equilibria.

The multiplicity in the other boundary situation $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*, 0)$ can be proved in a similar way.

Moreover, there are infinitely many partial separating equilibria.

In the interior case $\widehat{p}_I(\theta_0^*) \in (\widetilde{p}_I(\theta_0^*, 1), \widetilde{p}_I(\theta_0^*, 0))$, there is a continuum of θ_0 s.t. $\widehat{p}_I(\theta_0) \in (\widetilde{p}_I(\theta_0, 1), \widetilde{p}_I(\theta_0, 0))$, since $\widehat{p}_I(\theta_0)$, $\widetilde{p}_I(\theta_0, 1)$ and $\widetilde{p}_I(\theta_0, 0)$ are all continuous in θ_0. For each such θ_0, $\exists t \in (0,1)$ s.t. $(\theta_0, m_t) \in \mathbb{R} \times M(r, \mu)$ is a partial separating equilibrium satisfying MLRP. There also exist infinitely many partial separating equilibria that do not satisfy MLRP, since for each such (θ_0, m_t) we can construct a new equilibrium by removing some weight from the upper and the lower increasing components to the decreasing component without changing the corresponding $\widehat{p}_I(\theta_0)$.

In the boundary cases, e.g. $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*, 1)$, suppose there are finitely many equilibria. The above analysis implies that all these equilibria are on the boundary $\widehat{p}_I(\theta_0) = \widetilde{p}_I(\theta_0, 1)$, otherwise there must be some interior case which induces infinitely many equilibria. Suppose there are n equilibria, then the set of equilibria is $\left\{\theta_0^k, m_1\right\}_{k=1}^n$. Let $\theta_0^1 = \min\left\{\theta_0^k\right\}_{k=1}^n$, then for any shape $m \in M(r,\mu)$ and any $\theta_0 < \theta_0^*$, $\widehat{p}_I(\theta_0) < \int m(\theta-\theta_0)\cdot dP(\theta)$ and thus $g(\theta_0,m) = r/2 - \mu\cdot\ln\left(\frac{\int m(\theta-\theta_0)\cdot dP(\theta)}{1-\int m(\theta-\theta_0)\cdot dP(\theta)}\right) < r/2-\mu\cdot\ln\left(\frac{\widehat{p}_I(\theta_0)}{1-\widehat{p}_I(\theta_0)}\right) = \theta_0$. As shown before, this implies the existence of an equilibrium pooling in I, which is a contradiction. The existence of infinitely many partial separating equilibria in the other boundary situation $\widehat{p}_I(\theta_0^*) = \widetilde{p}_I(\theta_0^*,0)$ can be proved in a similar way. ■

Proof of Proposition 11.

Proof. We write $g_\beta(\cdot)$ for the density function over signals induced by precision β signals, 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's conjecture over his opponent's play is the probability he attaches to his opponent investing as a function of θ, which is a function $m:\mathbb{R}\rightarrow[0,1]$. For the same reason in the last paragraph of Section 2, we restrict our attention to $m \in \widetilde{\Omega} = \{m\in\Omega|\forall\theta\in\operatorname{supp}(P(\cdot)),\, m(\theta)\in[0,1]\}$. Note that $\widetilde{\Omega}$ is a compact functional space.

If a player chooses (β,s) against conjecture m, his expected utility is

$$V(\beta,s,m) = \int_x s(x)\cdot\left[\int_\theta(\theta-r\cdot(1-m(\theta)))\cdot l_\beta(\theta|x)\cdot d\theta\right]\cdot g_\beta(x)\cdot dx$$

With an optimal choice of $s(\cdot)$ this gives

$$\begin{aligned} V^*(\beta,m) &= \int_x \max\left\{0,\int_\theta(\theta-r\cdot(1-m(\theta)))\cdot l_\beta(\theta|x)\cdot d\theta\right\}\cdot g_\beta(x)\cdot dx \\ &= \int_x \max\left\{0,\int_\theta(\theta-r\cdot(1-m(\theta)))\cdot\beta^{1/2}\cdot f\left(\beta^{1/2}(x-\theta)\right)\cdot p(\theta)\cdot d\theta\right\}\cdot dx \qquad (B.15)\end{aligned}$$

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 Dirac function. Then $(B.15)$ implies

$$\begin{aligned} V^{**}(m) &\triangleq \lim_{\beta\to\infty}V^*(\beta,m) \\ &= \int_x \max\{0,(x-r\cdot(1-m(x)))\cdot p(x)\}\cdot dx \\ &= \int_\theta \max\{0,(\theta-r\cdot(1-m(\theta)))\cdot p(\theta)\}\cdot d\theta \qquad (B.16)\end{aligned}$$

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

We first show that $\forall m \in \widetilde{\Omega}$, $\forall \beta > 0$, $V^{**}(m) > V^*(\beta, m)$. By the convexity of $\max\{0, \cdot\}$, $(B.15)$ implies

$$\begin{aligned} V^*(\beta, m) &= \int_x \max\left\{0, \int_\theta (\theta - r \cdot (1 - m(\theta))) \cdot \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot p(\theta) \cdot d\theta\right\} \cdot dx \\ &\leq \int_x \int_\theta \max\{0, (\theta - r \cdot (1 - m(\theta))) \cdot p(\theta)\} \cdot \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot d\theta \cdot dx \\ &= \int_\theta \max\{0, (\theta - r \cdot (1 - m(\theta))) \cdot p(\theta)\} \cdot \int_x \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot dx \cdot d\theta \\ &= \int_\theta \max\{0, (\theta - r \cdot (1 - m(\theta))) \cdot p(\theta)\} \cdot 1 \cdot d\theta \\ &= V^{**}(m) \qquad (B.17) \end{aligned}$$

where the last equality follows $(B.16)$. The above inequality is strict iff

$$L\left(\left\{x \in \mathbb{R} | \int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{(\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) > 0\}} \cdot d\theta \in (0, 1)\right\}\right) > 0 \qquad (B.18)$$

where $L(\cdot)$ is the Lebesque measure over $\mathbb{R}$ and 1_A denotes the characteristic function of set A. Since $\forall \theta \in$supp$(P(\cdot))$, $m(\theta) \in [0, 1]$, we have $\Pr((\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) > 0) \geq \Pr((\theta - r) \cdot p(\theta) > 0) = \Pr(\theta > r) > 0$ and $\Pr((\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) \leq 0) \geq \Pr(\theta \cdot p(\theta) \leq 0) = \Pr(\theta < 0) > 0$. Then supp$(f(\cdot)) = \mathbb{R}$ implies that $\forall x \in \mathbb{R}$,

$$\int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{(\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) > 0\}} \cdot d\theta \geq \int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{\theta > r\}} \cdot d\theta > 0$$

and

$$\int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{(\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) \leq 0\}} \cdot d\theta \geq \int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{\theta < 0\}} \cdot d\theta > 0$$

i.e. $\int_\theta \beta^{1/2} \cdot f\left(\beta^{1/2}(x - \theta)\right) \cdot 1_{\{(\theta - r \cdot (1 - m(\theta))) \cdot p(\theta) > 0\}} \cdot d\theta \in (0, 1)$. Thus $(B.18)$ holds and $(B.17)$ becomes

$$\forall m \in \widetilde{\Omega}, \forall \beta > 0, V^*(\beta, m) < V^{**}(m) \qquad (B.19)$$

Now we prove another intermediate result: $\forall \beta > 0$, $\exists \beta' > 0$ s.t. $\exists \delta(\beta, \beta') > 0$ s.t. $\forall m \in \widetilde{\Omega}$, $V^*(\beta', m) - V^*(\beta, m) > \delta(\beta, \beta')$.

Suppose this result does not hold, then $\exists \beta > 0$, s.t. $\forall \beta' > 0$, $\forall n > 0$, $\exists m^n_{\beta,\beta'} \in \widetilde{\Omega}$, s.t. $V^*\left(\beta', m^n_{\beta,\beta'}\right) - V^*\left(\beta, m^n_{\beta,\beta'}\right) \leq 1/n$. Thus $\forall \beta' > 0$, there exists a $m_{\beta,\beta'} \in \widetilde{\Omega}$ and a subsequence $\{n_{k,\beta'}\}_{k=1}^\infty$ s.t. $\lim_{k\to\infty} m^{n_{k,\beta'}}_{\beta,\beta'} = m_{\beta,\beta'}$ and $V^*(\beta', m_{\beta,\beta'}) - V^*(\beta, m_{\beta,\beta'}) \leq 0$, since $\widetilde{\Omega}$ is compact and $V^*(\beta, m)$ is a continuous functional of m for all $\beta > 0$. However, $(B.19)$ implies that $V^*(\beta', m_{\beta,\beta'}) - V^*(\beta, m_{\beta,\beta'}) > 0$ if β' is large enough, which is a contradiction.

According to this intermediate result, we conclude that $\forall \beta > 0$, $\exists \beta' > \beta$, s.t. $\forall m \in \widetilde{\Omega}$, $V^*(\beta', m) - c \cdot h(\beta') > V^*(\beta, m) - c \cdot h(\beta)$ if $c < \bar{c} \triangleq \frac{\delta(\beta,\beta')}{h(\beta') - h(\beta)} > 0$.

Therefore, the players would like to acquire information of precision at least β. ■

Proof of Proposition 12.

Proof. Let $\mathbf{P}_{\mathbb{R}}$ be the space of all the probability density functions over $\mathbb{R}$ and $\mathbf{P}_{[\epsilon,r-\epsilon]} \triangleq \{p \in \mathbf{P}_{\mathbb{R}} : \text{supp}\,(p) \subset [\epsilon, r-\epsilon]\}$. Define a distance $\lambda(\cdot,\cdot)$ on $\mathbf{P}_{\mathbb{R}}$ as $\lambda(p_1,p_2) = \int |p_1(\theta) - p_2(\theta)| \cdot d\theta$ for any pair $(p_1,p_2) \in \mathbf{P}_{\mathbb{R}} \times \mathbf{P}_{\mathbb{R}}$. Define a set of functions

$$M \triangleq \left\{ m(\cdot) \in \widetilde{\Omega} : \theta = \mu \cdot \ln\left(\frac{m(\theta)}{1-m(\theta)}\right) + r \cdot \left(\frac{1}{2} - m(\theta)\right)\right\} \qquad (B.20)$$

M is actually the set of all possible shapes of the equilibrium strategies. Note that $\#M = \left\{\begin{array}{ll} 1 & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} \le 1 \\ \infty & \text{if } \widetilde{r} = \frac{r}{4\cdot\mu} > 1 \end{array}\right\}$. Define a mapping $g : \mathbb{R} \times \mathbf{M} \times \mathbf{P}_{\mathbb{R}} \to \mathbb{R}$ as

$$g(\theta_0, m, p) \triangleq r/2 - \mu \cdot \ln\left(\frac{\int m(\theta-\theta_0)\cdot p(\theta)\cdot d\theta}{1-\int m(\theta-\theta_0)\cdot p(\theta)\cdot d\theta}\right) \qquad (B.21)$$

then $g(\theta_0, m, p)$ is a continuous function of θ_0 for given $(m,p) \in \mathbf{M} \times \mathbf{P}_{\mathbb{R}}$, a continuous functional of m for given $(\theta_0, p) \in \mathbb{R} \times \mathbf{P}_{\mathbb{R}}$ and a continuous functional of p for given $(\theta_0, m) \in \mathbb{R} \times \mathbf{M}$.

We first prove the following lemma.

Lemma 010 *$\exists \overline{\theta}_0, \underline{\theta}_0 \in \mathbb{R}$, s.t. $\overline{\theta}_0 > \underline{\theta}_0$, $\forall \theta_0 \ge \overline{\theta}_0$, $\forall (m,p) \in \mathbf{M} \times \mathbf{P}_{[\epsilon,r-\epsilon]}$, $g(\theta_0,m,p) - \theta_0 > \epsilon/2$ and $\forall \theta_0 \le \underline{\theta}_0$, $\forall (m,p) \in \mathbf{M} \times \mathbf{P}_{[\epsilon,r-\epsilon]}$, $g(\theta_0,m,p) - \theta_0 < -\epsilon/2$.*

Proof. If $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$, then $\forall \theta_0 \ge r - \epsilon + \mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2}\cdot(1-\widetilde{r}^{-1})^{1/2}$, Proposition 9 implies $\forall m \in M$, $m(\theta - \theta_0)$ is a strictly increasing function for $\theta \in [\epsilon, r-\epsilon]$. If $\widetilde{r} = \frac{r}{4\cdot\mu} \le 1$, $m(\theta-\theta_0)$ is strictly increasing in θ for all $\theta_0 \in \mathbb{R}$. Thus $\forall \theta_0 \ge r - \epsilon + \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2}\cdot(1-\widetilde{r}^{-1})^{1/2}, 0\right\}$, $\forall p \in \mathbf{P}_{[\epsilon,r-\epsilon]}$, we have $\Pr(\{\theta \in \text{supp}\,(p) : m(\theta-\theta_0) \le m(r -$
1 and

$$\begin{aligned} g(\theta_0,m,p) - \theta_0 &= r/2 - \mu \cdot \ln\left(\frac{\int m(\theta-\theta_0)\cdot p(\theta)\cdot d\theta}{1-\int m(\theta-\theta_0)\cdot p(\theta)\cdot d\theta}\right) - \theta_0 \\ &\ge r/2 - \mu \cdot \ln\left(\frac{\int m(r-\epsilon-\theta_0)\cdot p(\theta)\cdot d\theta}{1-\int m(r-\epsilon-\theta_0)\cdot p(\theta)\cdot d\theta}\right) - \theta_0 \\ &= r/2 - \mu \cdot \ln\left(\frac{m(r-\epsilon-\theta_0)}{1-m(r-\epsilon-\theta_0)}\right) - \theta_0 \qquad (B.22) \end{aligned}$$

$(B.20)$ implies

$$\theta - \theta_0 = \mu \cdot \ln\left(\frac{m(\theta-\theta_0)}{1-m(\theta-\theta_0)}\right) + r \cdot \left(\frac{1}{2} - m(\theta-\theta_0)\right) \qquad (B.23)$$

then

$$\mu \cdot \ln\left(\frac{m(r-\epsilon-\theta_0)}{1-m(r-\epsilon-\theta_0)}\right) = r - \epsilon - \theta_0 - r \cdot \left(\frac{1}{2} - m(r-\epsilon-\theta_0)\right) \qquad (B.24)$$

Plugging $(B.24)$ into $(B.22)$ leads to

$$\begin{aligned} g(\theta_0, m, p) - \theta_0 &\geq r/2 - \left[r - \epsilon - \theta_0 - r \cdot \left(\frac{1}{2} - m(r - \epsilon - \theta_0)\right)\right] - \theta_0 \\ &= \epsilon - r \cdot m(r - \epsilon - \theta_0) \qquad (B.25) \end{aligned}$$

Since $\lim_{\theta_0 \to +\infty} m(r - \epsilon - \theta_0) = 0$, there exists $\overline{\theta}_0 > r - \epsilon + \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$ s.t. $g(\theta_0, m, p) - \theta_0 > \epsilon/2$ for all $\theta_0 \geq \overline{\theta}_0$. Note that $(B.25)$ holds for all $p \in \mathbf{P}_{[\epsilon, r-\epsilon]}$ and all $m \in M$[18], thus this $\overline{\theta}_0$ is uniform over $\mathbf{M} \times \mathbf{P}_{[\epsilon, r-\epsilon]}$.

If $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, then $\forall \theta_0 \leq \epsilon - \left(\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}\right)$, Proposition 9 implies $\forall m \in M$, $m(\theta - \theta_0)$ is a strictly increasing function for $\theta \in [\epsilon, r - \epsilon]$. If $\widetilde{r} = \frac{r}{4 \cdot \mu} \leq 1$, $m(\theta - \theta_0)$ is strictly increasing in θ for all $\theta_0 \in \mathbb{R}$. Thus $\forall \theta_0 \leq \epsilon - \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$, $\forall p \in \mathbf{P}_{[\epsilon, r-\epsilon]}$, we have $\Pr(\{\theta \in \mathrm{supp}(p) : m(\theta - \theta_0) \geq m(\epsilon - \theta_0)\}) = 1$ and

$$\begin{aligned} g(\theta_0, m, p) - \theta_0 &= r/2 - \mu \cdot \ln\left(\frac{\int m(\theta - \theta_0) \cdot p(\theta) \cdot d\theta}{1 - \int m(\theta - \theta_0) \cdot p(\theta) \cdot d\theta}\right) - \theta_0 \\ &\leq r/2 - \mu \cdot \ln\left(\frac{\int m(\epsilon - \theta_0) \cdot p(\theta) \cdot d\theta}{1 - \int m(\epsilon - \theta_0) \cdot p(\theta) \cdot d\theta}\right) - \theta_0 \\ &= r/2 - \mu \cdot \ln\left(\frac{m(\epsilon - \theta_0)}{1 - m(\epsilon - \theta_0)}\right) - \theta_0 \qquad (B.26) \end{aligned}$$

by $(B.23)$

$$\mu \cdot \ln\left(\frac{m(\epsilon - \theta_0)}{1 - m(\epsilon - \theta_0)}\right) = \epsilon - \theta_0 - r \cdot \left(\frac{1}{2} - m(\epsilon - \theta_0)\right) \qquad (B.27)$$

Plugging $(B.27)$ into $(B.26)$ leads to

$$\begin{aligned} g(\theta_0, m, p) - \theta_0 &\leq r/2 - \left[\epsilon - \theta_0 - r \cdot \left(\frac{1}{2} - m(\epsilon - \theta_0)\right)\right] - \theta_0 \\ &= r \cdot [1 - m(\epsilon - \theta_0)] - \epsilon \qquad (B.28) \end{aligned}$$

Since $\lim_{\theta_0 \to -\infty} m(\epsilon - \theta_0) = 1$, there exists $\underline{\theta}_0 < \epsilon - \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$ s.t. $g(\theta_0, m, p) - \theta_0 < -\epsilon/2$ for all $\theta_0 \leq \underline{\theta}_0$. Note that $(B.28)$ holds for all $p \in \mathbf{P}_{[\epsilon, r-\epsilon]}$ and all $m \in M$[19], thus this $\overline{\theta}_0$ is uniform over $\mathbf{M} \times \mathbf{P}_{[\epsilon, r-\epsilon]}$. Also note that $\overline{\theta}_0 > \underline{\theta}_0$ by definition. ■

$\forall \delta > 0$, let $A_\delta \triangleq \left\{q \in \mathbf{P}_{\mathbb{R}} : \exists p \in \mathbf{P}_{[\epsilon, r-\epsilon]} \text{ s.t. } \lambda(q, p) < 2 \cdot \delta\right\}$. Now we prove another lemma.

Lemma 011 *$\exists \delta > 0$, s.t. $\forall q \in A_\delta$, $\forall m \in M$, $g(\overline{\theta}_0, m, q) - \overline{\theta}_0 > 0$ and $g(\underline{\theta}_0, m, q) - \underline{\theta}_0 < 0$.*

Proof. $\overline{\theta}_0 > r - \epsilon + \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$ implies that $g(\overline{\theta}_0, m, p) - \overline{\theta}_0$ does not vary over $m \in M$. Thus $g(\overline{\theta}_0, m, p) - \overline{\theta}_0$ is actually a continuous functional of p.

[18] Note that $m(r - \epsilon - \theta_0)$ does not vary over $m \in M$ since $\theta_0 \geq r - \epsilon + \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$.

[19] Note that $m(\epsilon - \theta_0)$ does not vary over $m \in M$ since $\theta_0 \leq \epsilon - \max\left\{\mu \cdot \ln\left(\frac{1-(1-\widetilde{r}^{-1})^{1/2}}{1+(1-\widetilde{r}^{-1})^{1/2}}\right) + \frac{r}{2} \cdot (1 - \widetilde{r}^{-1})^{1/2}, 0\right\}$.

By Lemma 010, we have

$$\forall (m,p) \in \mathbf{M} \times \mathbf{P}_{[\epsilon, r-\epsilon]},\, ,\ \epsilon/2 < g\left(\overline{\theta}_0, m, p\right) - \overline{\theta}_0 = r/2 - \mu \cdot \ln\left(\frac{\int m\left(\theta - \overline{\theta}_0\right) \cdot p(\theta) \cdot d\theta}{1 - \int m\left(\theta - \overline{\theta}_0\right) \cdot p(\theta) \cdot d\theta}\right) - \overline{\theta}_0$$

i.e.

$$\forall (m,p) \in \mathbf{M} \times \mathbf{P}_{[\epsilon, r-\epsilon]},\ p_I\left(\overline{\theta}_0, m, p\right) \triangleq \int m\left(\theta - \overline{\theta}_0\right) \cdot p(\theta) \cdot d\theta < \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1} \qquad (B.29)$$

Note that $\forall (m,q) \in \mathbf{M} \times A_\delta$, by definition $\exists p \in \mathbf{P}_{[\epsilon, r-\epsilon]}$ s.t. $\lambda(q,p) < 2 \cdot \delta$ and

$$\begin{aligned}
& |p_I\left(\overline{\theta}_0, m, q\right) - p_I\left(\overline{\theta}_0, m, p\right)| \\
= \ & |\int m\left(\theta - \overline{\theta}_0\right) \cdot q(\theta) \cdot d\theta - \int m\left(\theta - \overline{\theta}_0\right) \cdot p(\theta) \cdot d\theta| \\
\leq \ & \int m\left(\theta - \overline{\theta}_0\right) \cdot |q(\theta) - p(\theta)| \cdot d\theta \\
\leq \ & \int 1 \cdot |q(\theta) - p(\theta)| \cdot d\theta \\
= \ & \lambda(q,p) < 2 \cdot \delta \qquad (B.30)
\end{aligned}$$

where the last inequality comes from the fact $\forall \theta \in \mathbb{R}$, $m\left(\theta - \overline{\theta}_0\right) \in [0,1]$.

Let

$$\overline{\delta} = \frac{1}{4} \cdot \left[\frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1} - \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1}\right] > 0$$

then $(B.29)$ and $(B.30)$ implies $\forall (m,q) \in \mathbf{M} \times A_{\overline{\delta}}$, $\exists p \in \mathbf{P}_{[\epsilon, r-\epsilon]}$ s.t. $\lambda(q,p) < 2 \cdot \overline{\delta}$ and

$$\begin{aligned}
p_I\left(\overline{\theta}_0, m, q\right) &< p_I\left(\overline{\theta}_0, m, p\right) + 2 \cdot \overline{\delta} \\
&= p_I\left(\overline{\theta}_0, m, p\right) + \frac{1}{2} \cdot \left[\frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1} - \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1}\right] \\
&< \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1} \\
&\quad + \frac{1}{2} \cdot \left[\frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1} - \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1}\right] \\
&= \frac{1}{2} \cdot \left[\frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1} + \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2 + \epsilon/2\right)/\mu\right) + 1}\right] \\
&< \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1} \qquad (B.31)
\end{aligned}$$

Note that $(B.31)$ is equivalent to

$$\begin{aligned}
\forall (m,q) &\in \mathbf{M} \times A_{\overline{\delta}}, \\
g\left(\overline{\theta}_0, m, q\right) - \overline{\theta}_0 &= r/2 - \mu \cdot \ln\left(\frac{\int m\left(\theta - \overline{\theta}_0\right) \cdot q(\theta) \cdot d\theta}{1 - \int m\left(\theta - \overline{\theta}_0\right) \cdot q(\theta) \cdot d\theta}\right) - \overline{\theta}_0 \\
&> r/2 - \mu \cdot \ln\left(\frac{\frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1}}{1 - \frac{1}{\exp\left(\left(\overline{\theta}_0 - r/2\right)/\mu\right) + 1}}\right) - \overline{\theta}_0 = 0
\end{aligned}$$

thus we prove the first part of Lemma 010. By the same argument, we can choose $\underline{\delta} > 0$, s.t. $\forall q \in A_{\underline{\delta}}$, $\forall m \in M$, $g\left(\underline{\theta}_0, m, q\right) - \underline{\theta}_0 < 0$. Finally, let $\delta = \min\left\{\underline{\delta}, \overline{\delta}\right\}$. ■

Now choose $\delta > 0$ as suggested by Lemma 010. $\forall p \in \mathbf{P}_{\mathbb{R}}$ s.t. $\Pr\left(\theta \in [\epsilon, r-\epsilon]\right) > 1 - \delta$, let $\widetilde{p}\left(\theta\right) = \left\{ \begin{array}{c} 0 \;\; \text{if } \theta \in \mathbb{R} \backslash \left[\epsilon, r-\epsilon\right] \\ \frac{p(\theta)}{\Pr(\theta \in [\epsilon, r-\epsilon])} \;\; \text{if } \theta \in \left[\epsilon, r-\epsilon\right] \end{array} \right\}$, then $\widetilde{p} \in \mathbf{P}_{[\epsilon, r-\epsilon]}$. The distance between p and $\widetilde{p}$ is

$$
\begin{aligned}
\lambda\left(p, \widetilde{p}\right) &= \int \left|p\left(\theta\right) - \widetilde{p}\left(\theta\right)\right| \cdot d\theta \\
&= \int_{\theta \in \mathbb{R} \backslash [\epsilon, r-\epsilon]} \left|p\left(\theta\right) - \widetilde{p}\left(\theta\right)\right| \cdot d\theta + \int_{\theta \in [\epsilon, r-\epsilon]} \left|p\left(\theta\right) - \widetilde{p}\left(\theta\right)\right| \cdot d\theta \\
&= \int_{\theta \in \mathbb{R} \backslash [\epsilon, r-\epsilon]} \left|p\left(\theta\right) - 0\right| \cdot d\theta + \int_{\theta \in [\epsilon, r-\epsilon]} \left|p\left(\theta\right) - \frac{p\left(\theta\right)}{\Pr\left(\theta \in [\epsilon, r-\epsilon]\right)}\right| \cdot d\theta \\
&= 1 - \Pr\left(\theta \in [\epsilon, r-\epsilon]\right) + \frac{1 - \Pr\left(\theta \in [\epsilon, r-\epsilon]\right)}{\Pr\left(\theta \in [\epsilon, r-\epsilon]\right)} \cdot \int_{\theta \in [\epsilon, r-\epsilon]} p\left(\theta\right) \cdot d\theta \\
&= 1 - \Pr\left(\theta \in [\epsilon, r-\epsilon]\right) + \frac{1 - \Pr\left(\theta \in [\epsilon, r-\epsilon]\right)}{\Pr\left(\theta \in [\epsilon, r-\epsilon]\right)} \cdot \Pr\left(\theta \in [\epsilon, r-\epsilon]\right) \\
&= 2 \cdot \left[1 - \Pr\left(\theta \in [\epsilon, r-\epsilon]\right)\right] < 2 \cdot \delta
\end{aligned}
$$

this suggests $p \in A_{\delta}$. Then by Lemma 011, we have $g\left(\overline{\theta}_0, m, p\right) - \overline{\theta}_0 > 0$ and $g\left(\underline{\theta}_0, m, p\right) - \underline{\theta}_0 < 0$ for all $m \in M$. Since $g\left(\theta_0, m, p\right)$ is a continuous function of θ_0 for any given $(m, p) \in \mathbf{M} \times \mathbf{P}_{\mathbb{R}}$ and $\overline{\theta}_0 > \underline{\theta}_0$ as shown in Lemma 010, $\forall m \in M$, $\exists \theta_0^m \in \left(\underline{\theta}_0, \overline{\theta}_0\right)$ s.t. $g\left(\theta_0^m, m, p\right) = \theta_0^m$. Here $m\left(\theta - \theta_0^m\right)$ is a partial separating equilibrium. For any given $m \in M$, if $g\left(\theta_0, m, p\right) - \theta_0 \leq 0$ for some $\theta_0 \in \left(\overline{\theta}_0, \infty\right)$, then we have another partial separating equilibrium. Otherwise, $g\left(\theta_0, m, p\right) - \theta_0 > 0$ for all $\theta_0 \in \left(\overline{\theta}_0, \infty\right)$, which implies an equilibrium pooling in N. The same argument applies for $\theta_0 \in \left(-\infty, \underline{\theta}_0\right)$, which leads to other partial separating equilibria or the pooling (in I) equilibrium. Thus we prove the multiplicity for each $m \in M$. Moreover, if $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, $\#M = \infty$ and we have infinitely many equilibria.

This concludes the proof of Proposition 12.