FLEXIBLE INFORMATION ACQUISITION IN STRATEGIC SITUATIONS

MING YANG

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF ECONOMICS

ADVISOR: STEPHEN MORRIS

JUNE 2012

Abstract

My dissertation focuses on flexible information acquisition in strategic environments. "Flexible" means that people choose not only the quantity but also the qualitative nature of their information. This is modeled by rational inattention where information acquisition incurs a cost proportional to reduction of entropy. Hence, people only collect information most relevant to their payoffs but be rationally inattentive to other aspects. In strategic environments, people's incentives to acquire information are shaped by their payoff structures, which depend on others' strategies. My dissertation addresses the key questions like what information is acquired and how it affects equilibrium outcomes.

Chapter 1 studies the optimality of securitized debt under flexible information acquisition. A seller designs an asset backed security and a buyer decides whether to buy it to provide liquidity. Rather than treating the seller as an insider endowed with information, we assume no information asymmetry before bargaining. The buyer has an expertise in flexibly acquiring information of the fundamental. She collects the most relevant information determined by the "shape" of the security, which may endogenously generate adverse selection. Hence, the seller deliberately designs the security in order to induce the buyer to acquire information least harmful to the seller's interest. Issuing securitized debt is uniquely optimal in raising liquidity regardless of the stochastic interdependence of underlying assets. Fixed aggregate risk and homogeneous information cost are the key factors driving the results.

Chapter 2 studies flexible information acquisition in a coordination game with binary actions. When information is cheap, this flexibility enables players to acquire information that makes efficient coordination possible, while also leads to multiple equilibria. This result contrasts with the global game literature, where information structure is less flexible and cheap information leads to unique equilibrium with inefficient coordination. Moreover, differing from decision problems with information

iii

acquisition, players could be strictly better off if they can throw away information. We also go beyond the entropic information cost to highlight the key aspects of flexibility and how they drive our results.

Chapter 3 examines flexible information acquisition in linear best-response games. Introducing capacity constraints on information acquisition dampens both people's responses and their incentives to acquire information. We show an equivalence between the games with capacity constraints and the games without such constraints but having lower strategic externalities.

Acknowledgements

I am indebted to my adviser, Stephen Morris, whose guidance, support, and patience made this dissertation possible. Christopher Sims taught me essential modeling techniques to realize my idea. Hyun Shin provided expertise in finance and helped me make connections that move my work forward. I am grateful to Dilip Abreu, Markus Brunnermeier, Sylvain Chassang, Tri Vi Dang, Emmanuel Farhi, Bengt Holmstrom, Benjamin Moll, Wolfgang Pesendorfer, Richard Rogerson, Alp Simsek, Satoru Takahashi, Jean Tirole, Michael Woodford and Wei Xiong for the invaluable advice during the course of my study. I thank Wei Xiong for serving on my examination committee. I have also benefited greatly from advice and support from my teachers and colleagues at various stages of this process. I would especially like to thank Dong Beom Choi, Wei Cui, Liang Dai, Faruk Gul, Jakub Jurek, Juan Ortner, Martin Schmalz, Takuo Sugaya and Yao Zeng. I would also like to acknowledge the generous support of Laura Hedden and Karen Neukirchen. Finally, I would like to thank my parents and Chen Rui for their unconditional love and support.

Contents

Chapter 1

Optimality of Securitized Debt with Endogenous and Flexible Information Acquisition [1]

1.1 Introduction

Pooling assets and issuing asset-backed securities (ABSs), in particular, issuing a securitized debt, is a popular way to raise liquidity. For example, commercial banks pool a large number of individual home mortgages or automobile loans to create a special purpose vehicle (SPV), which then issues ABSs to finance the purchase of these loans. This process can be modeled as the following story. A risk-neutral seller owns some assets generating uncertain future cash flows. She is impatient and wants to raise liquidity by issuing an asset-backed security (ABS) to a risk-neutral buyer. To raise liquidity, the seller proposes an ABS and its price, and sets it as a take-it-or-leave-it offer. Then the buyer decides whether to accept the offer or not.

This simple trading game will serve as a benchmark throughout the chapter, and will be greatly enriched in order to capture our key ideas featuring the optimality of securitized debt.

The security design literature has provided insightful viewpoints in investigating this securitization process as described above. Much of this literature models sellers as "insiders" who are endowed with private information about the assets, which makes potential buyers hesitate to provide liquidity due to adverse selection. In overcoming such adverse selection, this literature considers the possibility of signaling by sellers, where buyers are passive because they cannot acquire any information about the assets. Also, various assumptions on information, assets and feasible securities to be designed are imposed in these models, which lead to various conclusions on the optimality of different forms of securities.

This chapter explores another perspective to look into the problem of adverse selection, which is universal in liquidity provision. Rather than assuming exogenous private information, we consider adverse selection resulted from information acquisition. A justification to this approach is that, in reality, people can naturally acquire different forms of information from different sources, so that the observed information asymmetry essentially lies in the ability to acquire information at the first place. While previous models directly assume the information asymmetry as exogenous, we explicitly model it endogenously. The main finding of this chapter is that, when one party of trading, namely, either the seller or the buyer, designs the contract and her counterparty acquires information about the fundamental, securitized debt should be the uniquely optimal contract for liquidity provision. This result is not limited to any specific case whether the buyer or the seller acquires information. To streamline the presentation of our entire analysis, we consider the case where the buyer can acquire information about the assets according to the security proposed by the seller in the benchmark model, which endogenously generates

adverse selection different from that in classic security design literature. We follow Dang, Gorton, and Holmstrom (2010) in treating the buyer as an "expert" who acquires information accordingly. In reality, buyers involved in ABS transactions are skillful and sophisticated. Their expertise in assessing investment opportunities is better modeled by endogenous information acquisition rather than exogenous information endowment. Here endogeneity means that the agents can choose from a set of information structures according to their investment opportunities. Taking this endogeneity into account, sellers design securities generating least incentive for buyers to acquire information. Dang, Gorton, and Holmstrom (2010) model such information acquisition through the costly state verification approach, in which buyers either acquire a specific signal about the future cash flows of assets or do not acquire any information. In other words, the buyer can only choose from two specific information structures. Based on this rigid information acquisition process, Dang, Gorton, and Holmstrom (2010) show that debt is the least information-sensitive and thus is an optimal contract to provide liquidity. However, there also exist infinitely many other securities, which are called "quasi-debts", as information-sensitive as the standard debt contract. Also, it is identified that some restrictive conditions are required in order to ensure the optimality of these quasi-debts when pooling is considered. As we discuss below, their non-uniqueness result stems from the rigidity of information acquisition inhabits the costly state verification approach.

This chapter differs from Dang, Gorton, and Holmstrom (2010) by allowing for flexible information acquisition, which helps achieve the unique optimality of securitized debt, even if pooling of various assets is taken into account. Similar to Dang, Gorton, and Holmstrom (2010), we assume no information asymmetry at the beginning to focus on the adverse selection resulting from endogenous information acquisition. Given the security backed by the cash flows and its associated price proposed by the seller, flexibility enables the buyer to acquire information accord-

ingly about the underlying assets. Here, specifically, flexibility means that the set of feasible information structures to be acquired by the buyer consists of all conditional distributions of signals on the underlying cash flows. It captures the ability of the buyer to allocate her attention in whatever way she wants. Hence, the buyer chooses not only the quantitative but also the qualitative nature of her information.

We model flexible information acquisition through the paradigm of rational inattention building upon Sims (2003), where any information structure can be acquired at a cost proportional to reduction of entropy. This cost could result from the required time or resource to run models, do statistical tests or write reports. Flexibility enables the buyer to acquire payoff-relevant information accordingly, and the information cost requires her to optimally acquire such information in both quantitative and qualitative aspects. For example, to assess a collateralized debt with face value $1000 and price $800, a potential buyer would like to analyze data more carefully to see when the underlying cash flow possibly varies around $800, but put less attention to check whether the cash flow could reach $2000 or not, since any realization of the underlying cash flow that is above the face value always generates $1000 to the buyer. Similar to Dang, Gorton, and Holmstrom (2010), standard securitized debt is optimal for liquidity provision in our model. But our result is sharper in the sense that securitized debt is the uniquely optimal one. In Dang, Gorton, and Holmstrom (2010), only two extreme information structures are available in the setup of costly state verification but infinite forms of securities can be designed, which inevitably results in the indistinguishability of some securities. In our framework, with help of flexibility, the variety of available information structures matches the variety of potential securities to be designed, and thus the uniqueness of the standard securitized debt could be guaranteed. Quasi-debts are no longer optimal in our model. By reshaping the uneven tail above the price of a quasi-debt to a flat one, not only the buyer's information cost could be saved but also poten-

tial loss of trade from adverse selection could be mitigated. The resulted surplus could be employed by the seller to make both parties better off, and thus ultimately make a better provision of liquidity possible. Moreover, flexible information acquisition provides a unified framework to analyze securitization of multiple assets. We show that pooling and issuing securitized debt is uniquely optimal to raise liquidity, regardless of the stochastic interdependence among the underlying assets and the allocation of bargaining power.

There are two key factors determining the unique optimality of standard securitized debt. The first one is the fixed aggregate risk implicitly specified in the benchmark trading game in the sense that the total cash flows owned by the seller and buyer are invariant with respect to the success or failure of the transaction. This factor and its underlying mechanism towards the unique optimality of securitized debt does not depend on whether the buyer or the seller acquires information, as long as the roles of security design and information acquisition are separately allocated to the two parties of trading. Specifically, in our benchmark model, since the aggregate risk is fixed, information acquisition is not socially valuable, so that acquiring information is no more than waste of money when both parties are considered as a unity. Moreover, this trading game with fixed aggregate risk leads to conflicting interests of the two parties, so that the information acquired by the buyer makes herself better off but at the expense of the seller through adverse selection, which further reduces the potential gain from trade. Since the buyer's incentive to acquire information is shaped by the offer proposed to her, the seller deliberately designs the ABS to optimally discourage information acquisition harmful to her own interests. Due to the limited liability, any feasible ABS is bounded above by the sum of underlying cash flows. When information cost is not too high, the flexibility allows the buyer to distinguish between any states with different payoffs. Hence the seller makes the ABS a constant whenever it is off the boundary to discourage in-

formation acquisition and thus mitigate adverse selection. This consideration gives rise to a flat tail. In states where the underlying cash flows are too low to support such constant, the ABS reaches the boundary and equals the sum of underlying cash flows. Therefore, the flat tail and the boundary component constitute a securitized debt, which is uniquely optimal for liquidity provision. We also use an example with variable aggregate risk to illustrate the importance of fixed aggregate risk in our framework. Consider the seller as an entrepreneur that raises funds from the buyer to take a project with uncertain future cash flows. They jointly expose themselves to an aggregate risk if the buyer accepts the offer, and are not exposed to such risk if the offer is rejected. In this case, information acquisition could be socially valuable and the conflicting interests of the two parties could be partly reconciled. Therefore, the seller could deliberately design a contract to encourage the buyer to acquire information that helps avoid investing in states where cash flows are too low. This increases her benefit from the trade, and also leads to a more socially desirable outcome.

Another key factor is the homogeneity in information acquisition. That is, no state is more special than other states in terms of the difficulty of information acquisition. This feature stems from rational inattention[2] and is the reason why our qualitative result does not depend on the stochastic interdependence among the underlying assets. Intuitively, if information about some assets is much easier to acquire than other assets, the flat part of the securitized debt cannot be preserved in the optimal ABS. We provide an example that illustrates this idea. The above two factors specify the boundary of our theory.

Finally, the origin of the uniqueness of optimal contract is not only from the flexibility itself, but from the double-sided symmetry of flexibility. In principle,

---

[2]We can achieve the same qualitative results under more general information costs possessing such property.

general flexible choice, not necessarily restricted to flexible information acquisition, enables an economic agent to make state-contingent responses. In other words, the agent can make a best response in one state, and can make another best response in another state. Double-sided symmetry of flexibility requires that both parties engaged in a potential trade are endowed with the same level of flexibility.

How this double-sided symmetry of flexibility works can be seen by comparing our framework to Dang, Gorton, and Holmstrom (2010) and the traditional models of costly state verification like Townsend (1979). In all these three models, the contract designer is endowed with flexibility, in the sense that she can assign state-contingent repayment through designing any form of security. What matters to shape the different results regarding uniqueness of the optimal contract relies on the potential flexibility of the other party who decides whether to accept the offer. In our framework, ex-ante symmetric information in the form of a double-sided ignorance prevents the buyer to make a state-contingent choice if she only follows the traditional costly state verification approach to acquire information. However, the buyer in our framework is able to choose state-contingent probability of accepting the offer, namely, she can perform flexible information acquisition. In this sense the buyer enjoys the same level of flexibility as the seller. Given this double-sided symmetry of flexibility in our model, the uniqueness of an optimal contract, which is the standard securitized debt, is guaranteed. In Dang, Gorton, and Holmstrom (2010), however, the buyer can only follow the traditional costly state verification approach to acquire information, in which only two options, namely, to acquire a signal or not, are available. In other words, the buyer in Dang, Gorton, and Holmstrom (2010) cannot make state-contingent decision. Hence, the desired double-sided symmetry of flexibility fails and the uniqueness of the optimal contract fails as a consequence. Interestingly, Townsend (1979) also employs the costly state verification approach with two options to model information acquisition, namely, to audit or not, but the

unique optimality of a standard debt still emerges. Why it is this case? Different from Dang, Gorton, and Holmstrom (2010) and our framework, in Townsend (1979) the entrepreneur has information advantage over the lender in the sense that the entrepreneur knows the realized profit of the project which the lender does not know. Thanks to the revelation principle, the lender who acquire information in the interim stage can decide whether to audit or not in any state based on the truth told by the entrepreneur who has private information. In other words, although the lender in Townsend (1979) still only has two options to acquire information as the buyer in Dang, Gorton, and Holmstrom (2010), such two options in Townsend (1979) are state-contingent while their counterparts in Dang, Gorton, and Holmstrom (2010) are not. Therefore, the double-sided symmetry of flexibility is still established in Townsend (1979), and the uniqueness of the optimal contract, also a standard debt, is ensured in their model as well.

This chapter proceeds as following. Section 1.2 studies flexible information acquisition in a binary choice problem, which provides a solid foundation for analyzing players' behavior in the trading game and liquidity provision. Section 1.3 derives the uniquely optimal contract as the securitized debt in various circumstances and identifies the two key driving forces of this result. We conclude and discuss in Section 1.4.

1.1.1 Relation to the Literature

We model players' information acquisition behavior through the framework of rational inattention building on Sims (1998) and Sims (2003).[3] In applied work, rational inattention is mainly studied in two cases: the linear-quadratic case (e.g.,

---

[3]To learn more about rational inattention, see Sims (2005),Sims (2006), Sims (2010), Luo (2008), Mackowiak and Wiederholt (2011), Van Nieuwerburgh and Veldkamp (2009a), Van Nieuwerburgh and Veldkamp (2009b), Paciello (2009), Matejka (2010), Mondria (2010), Matejka and Sims (2011).

Mackowiak and Wiederholt (2009)), and the binary-action case. A leading example of the latter is Woodford (2009) , where firms acquire information and then decide whether to review their prices. Similar to Yang (2011), this chapter also adopts the binary-action setup in a strategic framework, which is different from the single-person decision problem as employed in Woodford (2009). Compared to Yang (2011) where both players acquire information and move simultaneously, this chapter considers a case in which players move sequentially, and only one party acquires information that results in information asymmetry. Also, this chapter focuses on a specific security design problem, rather than addresses a general coordination game as Yang (2011). Together with Yang (2011), our work makes early attempts to incorporate rational inattention based flexible information acquisition into strategic problems and offers various new results different from this trend of rational inattention literature.

This chapter is also closely related to the security design literature, in much of which sellers are modeled as "insiders" exogenously endowed with private information. Sellers' information advantage over buyers results in adverse selection which further leads to inefficient trade. In order to deal with the adverse selection problem given that buyers cannot acquire information, sellers want to signal their private information out in order to partly retrieve efficient trade. In this process, appropriate security design matters. This is because signaling is costly, so that to design a security that is less information sensitive than the original asset could save the signaling cost, which in turn adds to the profit of sellers. This consideration is plausible and insightful results have been well established in literature, but there may also be other interesting possibilities worth investigating. Also, various assumptions are imposed in this literature to deliver various results. In this chapter, buyers in financial markets may also actively acquire information, which could result in different interplay between the two parties and different results of security design, and

we can get clearer results from a single assumption.

The key difference between our approach and much of the security design literature could be clearly seen in discussing some of their assumptions and results in details. Gorton and Pennacchi (1990) shows that splitting assets into debt and equity mitigates the lemon problem between outsiders and insiders. They directly assume the existence of debt rather than considering a security design problem. In DeMarzo and Duffie (1999), informed sellers signal the quality of assets to competitive liquidity suppliers through retaining part of the cash flows. Equity is issued when the contractible information is not very sensitive to sellers' private information. Standard debt is optimal within the set of non-decreasing securities if the information structure allows a uniform worst case. Biais and Mariotti (2005) studies the effects of market power on market liquidity. They derive both the optimal security and trading mechanism through the approach of mechanism design. Debt contract turns out to be optimal under distributional conditions of underlying cash flows. DeMarzo (2005) focuses on the consequences of pooling and tranching. Pooling has an information destruction effect that destroys the seller's ability to signal the quality of her assets separately. When tranching is possible, pooling may also have a risk diversification effect that reduces information sensitivity of the senior claim. Under specific distributional assumptions of the noise structure, DeMarzo (2005) shows that the risk diversification effect dominates the information destruction effect as the number of underlying assets goes to infinity. In this limit case, pooling and tranching become optimal. These models also restrict their attention to non-decreasing securities[4]. Innes (1990) provides a standard motivation for this constraint. When the security is not monotone, a seller may cheat through borrowing from a third party, reporting a high cash flow to reduce her repayment and then

---

[4]Biais and Mariotti (2005) also assume dual monotonicity, i.e., both the security and the residual cash flow are non-decreasing.

repaying the side loan. The validity of this argument depends on the context. In the case of publicly traded stocks or bonds, this kind of cheat is unlikely to happen because it is difficult or even illegal for seller to manipulate the cash flows. Moreover, when the security is written on multiple underlying assets, even the concept of monotonicity is not well defined. Our framework is free of these limits.

It is also interesting to contrast our work to Axelson (2007). Different from the signaling literature of security design as DeMarzo and Duffie (1999) and subsequent work, Axelson (2007) considers a security design problem when the buyer rather than the seller have private information about the asset. The benchmark model of Axelson (2007) could be viewed as the case where there is variable aggregate risk, the seller designs contract and the buyer has information advantage in my framework, so that it is natural to expect equity rather than debt to be a better solution to the seller's financing problem. Our framework is different from Axelson (2007) in a deeper sense that we view the buyer's private information as endogenous, which results from endogenous and flexible information acquisition. Also, as we emphasized above, our results are more general in the sense that they are not limited to the case where the buyer is able to obtain private information, but only require that one party of trading designs contract and the other acquires information.

1.2 Binary Choice with Endogenous and Flexible Information Acquisition

Before introducing the economic environment of security design problem, we review the logic of binary choice with flexible information acquisition, which will play a key role in the following analysis. The readers mainly interested in the security design problem can skip this section and go back to it when needed.

In our leading example, a buyer faces a take-it-or-leave-it offer. She has to

acquire information and then make a binary choice. We first focus on information structures with binary signals and then show that it suffices to do so.

1.2.1 Decision Problem

Consider an agent who has to choose an action $a \in \{0, 1\}$ and will receive a payoff $u(a, \theta)$, where $\theta \in \Theta \subset \mathbb{R}$ is an unknown state distributed according to a continuous probability measure P over Θ.

The agent has access to the set of binary-signal information structures. In particular, she observes signals $x \in \{0, 1\}$ parameterized by measurable function $m : \Theta \to [0, 1]$, where $m(\theta)$ is the probability of observing signal 1 if the true state is θ (and so $1 - m(\theta)$ is the probability of observing signal 0). The conditional probability function $m(\theta)$ describes the agent's information acquisition strategy. By choosing different functional forms for $m(\theta)$, the agent can make her signal covary with fundamental in any way she would like. Intuitively, if her welfare is sensitive to fluctuation of the state within some range $A \subset \Theta$, she would pay much attention to this event by letting $m(\theta)$ be highly sensitive to $\theta \in A$. In this sense, choosing an information structure can be interpreted as hiring an analyst to write a report with emphasis on your interests. This idea will be made more clear through an example later in this section.

Quantity and Cost of Information

Following Sims (2003), we measure the quantity of information according to information theory building on Shannon (1948). Information conveyed by an information structure $m(\cdot)$ is defined as the expected reduction of uncertainty through observing signals generated by $m(\cdot)$, where the uncertainty associated with a distribution is measured by Shannon's entropy.

Before observing her signal, the agent's uncertainty about θ is given by Shannon's entropy of her prior[5]

$$H\left(prior\right) = -\int_{\Theta} p\left(\theta\right) \ln p\left(\theta\right) d\theta \ ,$$

where p is the density function of prior P[6]. After observing signal 1, the agent forms a posterior of θ

$$\frac{m\left(\theta\right) p\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}$$

and her posterior uncertainty upon receiving signal 1 is measured by her posterior entropy

$$
\begin{aligned}
H\left(posterior|1\right) &= -\int_{\Theta} \frac{m\left(\theta\right) p\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)} \ln\left(\frac{m\left(\theta\right) p\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}\right) d\theta \\
&= -\int_{\Theta} \frac{m\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)} \ln\left(\frac{m\left(\theta\right) p\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}\right) dP\left(\theta\right) \ .
\end{aligned}
$$

Similarly, observing signal 0 leads to a posterior

$$\frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}$$

and posterior entropy

$$H\left(posterior|0\right) = -\int_{\Theta} \frac{1 - m\left(\theta\right)}{1 - \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)} \ln\left(\frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}\right) dP\left(\theta\right) \ .$$

Then the agent's expected posterior entropy through choosing information structure

---

[5]This is essentially the unique measure of uncertainty given three axioms. See Cover and Thomas (1991) for detailed discussion.

[6]Following the convention of information theory, we let $0 \cdot \ln 0 = 0$. This is reasonable since $\lim_{x \to 0} x \cdot \ln x = 0$.

$m\left(\cdot\right)$ is

$$H\left(posterior\right)$$
$$= \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right) \cdot H\left(posterior|1\right) + \left[1 - \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)\right] \cdot H\left(posterior|0\right)$$
$$= -\int_{\Theta} m\left(\theta\right) \ln\left(\frac{m\left(\theta\right) p\left(\theta\right)}{\int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}\right) dP\left(\theta\right) - \int_{\Theta} \left[1 - m\left(\theta\right)\right] \ln\left(\frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \int_{\Theta} m\left(\theta'\right) dP\left(\theta'\right)}\right) dP\left(\theta\right) \ .$$

Let $I\left(m\right)$ denote the quantity of information gained through $m\left(\cdot\right)$, which equals the difference between the agent's prior entropy and expected posterior entropy, i.e.,

$$I\left(m\right) = H\left(prior\right) - H\left(posterior\right)$$
$$= \left[\int_{\Theta} g\left(m\left(\theta\right)\right) dP\left(\theta\right) - g\left(\int_{\Theta} m\left(\theta\right) dP\left(\theta\right)\right)\right] , \qquad (1.1)$$

where

$$g\left(x\right) = x \cdot \ln x + \left(1 - x\right) \cdot \ln\left(1 - x\right) \ .$$

In information theory, $I\left(m\right)$ is called mutual information. It measures the quantity of information about θ that is conveyed by the signal.

Write

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

for the set of binary-signal information structures. Let $c : M \to \mathbb{R}_+$ be the cost (in terms of utility) of acquiring information. We assume that the cost is proportional to the quantity of information gained, i.e.,

$$c\left(m\right) = \mu \cdot I\left(m\right) , \qquad (1.2)$$

where $\mu > 0$ is the marginal cost of information acquisition. It measures the difficulty in acquiring information. When $\mu = 0$, information acquisition incurs no cost

and the agent can directly observe the true state. When $\mu \to \infty$, the agent cannot acquire any information at all.

It is worth noting that mutual information $I(m)$ measures function m's variability, which reflects the informativeness of actions to the fundamental. For example, when $m(\theta)$ is constant, the actions convey no information about θ and the corresponding mutual information is zero. This is because function g is strictly convex and thus $I(m)$ is zero if and only if $m(\theta)$ is constant. Hence, a nice property of our technology of information acquisition is that there exists information acquisition if and only if $m(\theta)$ varies over θ, if and only if information cost is positive. Also note that the "shape" (functional form) of m determines not only the quantity but also the qualitative nature of information. For instance, an agent can concentrate her attention to some event through making $m(\theta)$ highly sensitive to θ within such event. In this sense, our technology of information acquisition is flexible since the agent can decide both the quantity and quality of their information through freely choosing from M. It is also worth noting that $c(\cdot)$ is convex, i.e.,

$$c\left(t \cdot m_1 + (1-t) \cdot m_2\right) \leq t \cdot c\left(m_1\right) + (1-t) \cdot c\left(m_2\right)$$

for all $m_1, m_2 \in M$ and $t \in [0,1]$. This convexity is strict when at least one of m_1 and m_2 is not a constant in θ.

Solving Binary Decision Problem with Information Acquisition

Now we are interested in the problem of an agent choosing an information structure $m \in M$ and a stochastic decision rule $f : \{0,1\} \to [0,1]$ to maximize her expected

utility

$$V\left(m,f\right)=\int_{\Theta}\left\{\begin{array}{c}\left[m\left(\theta\right)f\left(1\right)+\left(1-m\left(\theta\right)\right)f\left(0\right)\right]\cdot u\left(1,\theta\right)\\+\left[m\left(\theta\right)\left(1-f\left(1\right)\right)+\left(1-m\left(\theta\right)\right)\left(1-f\left(0\right)\right)\right]\cdot u\left(0,\theta\right)\end{array}\right\}dP\left(\theta\right)-c\left(m\right)\ .$$

(1.3)

Without loss of generality, we can let $f = f^*$ where $f^*(1) = 1$ and $f^*(0) = 0$. This simplification is based on the following observation. If we let

$$m^*\left(\theta\right) = m\left(\theta\right)f\left(1\right) + \left(1 - m\left(\theta\right)\right)f\left(0\right),$$

then $V\left(m^*, f^*\right) \geq V\left(m, f\right)$. This is because the first term of (1.3) remains the same, while the information cost becomes smaller due to the convexity of $c\left(\cdot\right)$.[7]

Fixing $f = f^*$, we can interpret m as a joint information structure and decision rule specifying that the agent will take action 1 with probability $m\left(\theta\right)$ in state θ.

---

[7]A simple proof: the convexity of $c\left(\cdot\right)$ implies

$$\begin{aligned}c\left(\alpha\cdot m\right) &= c\left(\alpha\cdot m + \left(1-\alpha\right)\cdot 0\right)\\ &< \alpha\cdot c\left(m\right) + \left(1-\alpha\right)\cdot c\left(0\right)\\ &= \alpha\cdot c\left(m\right)\end{aligned}$$

for $\alpha \in [0, 1)$. Without loss of generality, let $\Delta f = f\left(1\right) - f\left(0\right) \geq 0$. Note that if $f\left(1\right) = 0$ or $f\left(1\right) = 1$ and $f\left(0\right) = 0$, we are done. Let $\alpha = \Delta f / f\left(1\right)$. Thus at least one of $f\left(1\right)$ and α is strictly less than 1.Then

$$\begin{aligned}c\left(m^*\right) &= c\left(f\left(1\right)\cdot\left[\alpha\cdot m + 1 - \alpha\right]\right)\\ &\leq f\left(1\right)\cdot c\left(\left[\alpha\cdot m + 1 - \alpha\right]\right)\\ &\leq f\left(1\right)\cdot\left(\left[\alpha\cdot c\left(m\right) + 0\right]\right)\\ &\leq \Delta f\cdot c\left(m\right) < c\left(m\right)\ .\end{aligned}$$

Now the agent's problem is to choose $m \in M$ to maximize

$$V^*(m) = \int_\Theta [m(\theta) \cdot u(1,\theta) + [1 - m(\theta)] \cdot u(0,\theta)] \, dP(\theta) - c(m)$$

$$= \int_\Theta m(\theta) \cdot [u(1,\theta) - u(0,\theta)] \, dP(\theta) - c(m) + \int_\Theta u(0,\theta) \, dP(\theta) .$$

Since $\int_\Theta u(0,\theta) \, dP(\theta)$ is a constant that does not depend on m, we can redefine the agent's objective as

$$\max_{m \in M} V^*(m) = \int_\Theta \Delta u(\theta) \cdot m(\theta) \, dP(\theta) - c(m) ,$$

where

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

is the payoff gain from taking action 1 over action 0. It shapes the agent's incentive of information acquisition.

The following lemma characterizes the optimal strategy m for the agent.[8]

Proposition 1.1 [9]*Let* $\Pr(\Delta u(\theta) \neq 0) > 0$ *to exclude the trivial case that the agent is always indifferent between the two actions. Let* $m \in M$ *be an optimal strategy and*

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

be the corresponding unconditional probability of taking action 1. Then,

i) the optimal strategy is unique;

ii) there are three possibilities for the optimal strategy:

---

[8]We became aware of the related work Woodford (2008) while working on this paper. Here we use Lemma 2 of Woodford (2008) to characterize the optimal strategy. To maintain the completeness of our paper, we give a proof in our context.

[9]We do not have to require $\Theta \subset \mathbb{R}$. This proposition holds for any probability space Θ.

a) $p_1 = 1$ (i.e., $m(\theta) = 1$ almost surely) if and only if

$$\int_\Theta \exp\left(-\mu^{-1}\Delta u(\theta)\right) dP(\theta) \leq 1 \; ; \tag{1.4}$$

b) $p_1 = 0$ (i.e., $m(\theta) = 0$ almost surely) if and only if

$$\int_\Theta \exp\left(\mu^{-1}\Delta u(\theta)\right) dP(\theta) \leq 1 \; ; \tag{1.5}$$

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

$$\int_\Theta \exp\left(\mu^{-1}\Delta u(\theta)\right) dP(\theta) > 1 \ and \ \int_\Theta \exp\left(-\mu^{-1}\Delta u(\theta)\right) dP(\theta) > 1 \; ; \tag{1.6}$$

in this case, the optimal strategy m is characterized by

$$\Delta u(\theta) = \mu \cdot \left[g'(m(\theta)) - g'(p_1)\right] \tag{1.7}$$

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

$$g'(x) = \ln\left(\frac{x}{1-x}\right) \; .$$

Proof. See Appendix 1.5. ∎

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

In case a), condition (1.4) holds if action 1 is very likely the ex ante best action and the cost of acquiring information is sufficiently high. Hence the agent just takes action 1 without acquiring any information. Similarly, case b) implies that if action 0 is ex ante very likely to dominate action 1 and the information cost is

sufficiently high, the agent always takes action 0. In this two cases, marginal benefit of acquiring information is less than the marginal cost. Hence the decision maker chooses not to acquire any information.

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

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

When neither action is dominant, i.e.,

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

the marginal cost of information acquisition μ plays a role.

On the one hand,

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

Hence Proposition 1.1 predicts that no information is acquired if μ is high enough.

On the other hand, since

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

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

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

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

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

$$= +\infty \ ,$$

we have

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

A similar argument leads to

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

Therefore, Proposition 1.1 reads that there must exist information acquisition if the marginal cost of information is sufficiently low. This interpretation coincides with our intuition that the agent rationally decides whether to acquire information through comparing the cost to the benefit of information acquisition.

When neither action is dominant and the marginal cost of information acquisition takes intermediate values, the agent finds it optimal to acquire some information to make her action (partially, in a random manner) contingent on θ. This is the case specified by condition (1.6). Since g' is strictly increasing, (1.7) implies that $m\left(\theta\right)$, the conditional probability of choosing action 1, is increasing with respect to payoff gain $\Delta u\left(\theta\right)$. This is intuitive. The left hand side of (1.7) represents the marginal benefit of increasing $m\left(\theta\right)$, while the right hand side of (1.7) is the marginal cost of

information when increasing $m(\theta)$. Therefore, if deciding to acquire information, the agent will equate her marginal benefit with her marginal cost of doing so.

An Example

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

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

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

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

Let $t = 0$, then the agent finds it optimal to acquire some information. According to (1.7), the optimal information acquisition strategy $m(\theta)$ satisfies

$$\theta/\mu = g'(m(\theta)) - g'\left(\int_\Theta m(\theta)\, dP(\theta)\right), \tag{1.8}$$

where

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

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

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

Hence

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

and (1.8) becomes

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

Therefore,

$$m\left(\theta\right) = \frac{1}{1 + \exp\left(-\theta/\mu\right)} \ . \tag{1.9}$$

First note that

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

Step function $a\left(\theta\right)$ captures the agent's choice under complete information. In this case, the agent can observe the exact value of θ. When $\mu > 0$, the best response is characterized by (1.9). Since information is no longer free, the agent has to allow some mistake in her response. The conditional probability of mistake is given by

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

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

Second, parameter μ measures the difficulty in acquiring information. Figure 1.1 shows how $m\left(\theta\right)$ varies with this parameter.

When $\mu = 0$, information acquisition incurs no cost and the agent's response is a step function. She never makes mistake. When μ becomes larger, she starts to compromise the accuracy of her decision to save information cost. Larger μ leads to flatter $m\left(\theta\right)$. Finally, when μ is extremely large, $m\left(\theta\right)$ is almost constant and the agent almost stops acquiring information.
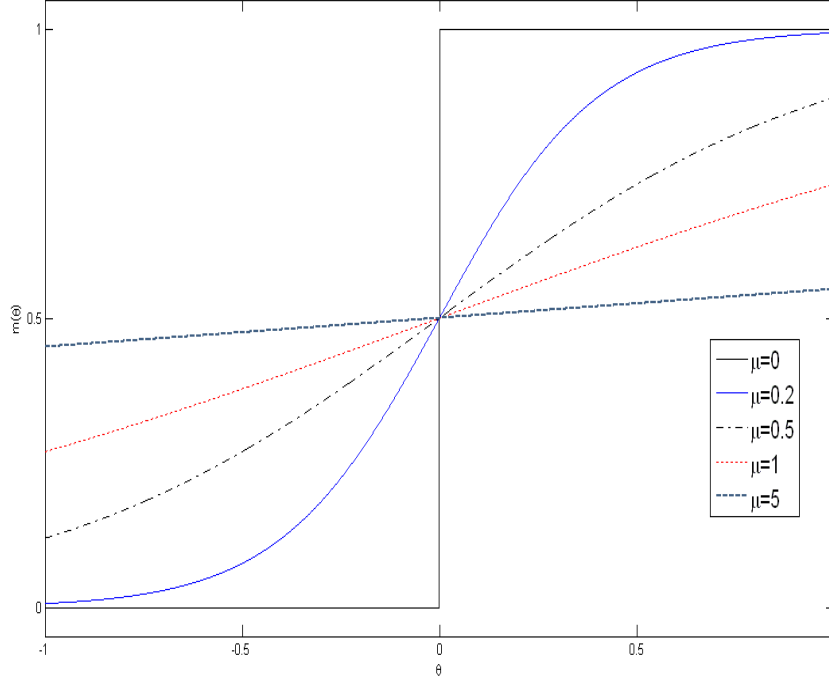
Figure 1.1: Information acquisition under various information costs

Third, since the agent's action is highly sensitive to θ where slope $\left|\frac{dm(\theta)}{d\theta}\right|$ is large, $\left|\frac{dm(\theta)}{d\theta}\right|$ reflects her attentiveness around θ. Under this interpretation, Figure 1.1 reveals that the agent actively collects information for intermediate values of the fundamental but is rationally inattentive to values at the tails. This result coincides with our intuition. When θ is too high (low), the agent should take action 1 (action 0) anyway. Hence the information about θ on the tails are not so relevant to her payoff. When θ takes intermediate values, the agent's payoff gain from taking action 1 over action 0 depends crucially on the sign of θ. Therefore, the information about θ around zero is payoff-relevant and attracts most of her attention.

We have been focusing on binary-signal information structures. Next subsection justifies this setup.

1.2.2 Justifying the Binary-signal Information Structure

Generally, an agent can purchase any information structure $((X, \sigma), \pi)$. Here X is the set of realizations of the signal, σ is a σ-algebra on X, and $\forall \theta \in \Theta$, $\pi(\cdot|\theta)$ is a probability measure on X. $\pi(\cdot|\cdot)$ conveys information about state θ in the sense that for any event $A \subset X$, $\pi(A|\theta)$ specifies the conditional probability of A given θ. Before making a decision, the agent can acquire information about the state in the form of an information structure. An information structure specifies both the quantity and qualitative nature of the information.

The binary-signal information structure analyzed above is a special case with $X = \{0, 1\}$ and $\pi(1|\theta) = m(\theta)$ (and so $\pi(0|\theta) = 1 - m(\theta)$). For binary choice problem with flexible information acquisition, it suffices to restrict our attention to this special class of information structures. To see this, let $((X, \sigma), \pi)$ be any information structure chosen by the agent. Given $((X, \sigma), \pi)$, the agent optimally chooses her action rule as $a : X \to [0, 1]$, where $a(x)$ is the probability of taking action 1 upon receiving signal x. Let

$$
\begin{aligned}
X_1 &= \{x \in X : a(x) = 1\}, \\
X_0 &= \{x \in X : a(x) = 0\},
\end{aligned}
$$

and

$$
X_{ind} = \{x \in X : a(x) \in (0, 1)\}.
$$

X_1 (X_0) is the set of signal realizations such that the agent definitely takes action 1 (0). She is indifferent when her signal belongs to X_{ind}. Then (X_1, X_0, X_{ind}) forms a partition of X. Since the only use of the signal is to make a binary decision, a signal differentiating more finely among the states just conveys redundant information and

wastes the agent's attention. Hence the agent will not discern signal realizations within any of X_1, X_0 and X_{ind}. In addition, because she is indifferent between action 0 and 1 upon event X_{ind}, she would rationally pay no attention to distinguish this event from other realizations. Hence, the agent always play pure strategies upon receiving her signal. Therefore, the agent always prefers binary-signal information structures.[10]

1.3 Security Design with Information Acquisition

1.3.1 Basic Setup

We consider a two-period game with two players. One player is a seller that owns N assets at period 0. These assets generate verifiable random cash flows $\overrightarrow{\theta} \in \Theta \subset \mathbb{R}_+^N$ in period 1[11]. The other player is a potential buyer holding consumption goods (money) at period 0. Player i's utility function is given by

$$u_i = c_{i0} + \delta_i \cdot c_{i1}, \tag{1.10}$$

where c_{it} denotes player i's consumption at period t and $\delta_i \in [0,1]$ is her subjective discount factor, $i \in \{s,b\}$($\{s,b\}$ stands for $\{seller, buyer\}$). We assume $\delta_b > \delta_s$ to represent that the seller has a better investment opportunity than the buyer. This assumption creates the trading demand. Both agents may benefit from transferring some goods to the seller at date 0 and compensating the buyer with repayment backed by the random cash flows $\overrightarrow{\theta}$ at date 1.

Similar to Dang, Gorton, and Holmstrom (2010), we assume no information

---

[10]Woodford (2009) has a similar argument that the agent only needs to acquire a "yes/no" signal.

[11]Here the assumption of verifiable cash flows is natural, since we generally have third parties monitor and collect the underlying loans and distribute the cash flows to the holders of asset backed securities.

asymmetry at period 0 to focus on the adverse selection resulting from endogenous information acquisition. Hence the two agents start with identical information about $\overrightarrow{\theta}$, which is represented by a full support common prior P over Θ. Without loss of generality, we assume that P is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}_+^N$.

A security backed by $\overrightarrow{\theta}$, the cash flows of the N assets, is a mapping $s : \Theta \to \mathbb{R}_+$ such that $\forall \overrightarrow{\theta} \in \Theta$, $s\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^N \theta_n\right]$. A contract $(s(\cdot), q)$ is a security $s(\cdot)$ associated with a price $q > 0$. Throughout the chapter, we focus on the case where one player proposes a take-it-or-leave-it contract $(s(\cdot), q)$ to her opponent, who then acquires information and decides whether to accept it. This setup captures the idea that some agents in the markets of securitized assets are less sophisticated than others and cannot produce private information about the underlying cash flows. This separation between bargaining power and ability of information acquisition also makes our problem tractable.[12]

We first study the case where the seller designs the contract and the buyer acquires information. We then highlight two key factors driving the unique optimality of issuing securitized debt. We finally exchange the bargaining power and the ability of information acquisition to show the robustness of our main results.

---

[12]We would have to study a much more complicated signaling game if the issuer can produce private information before her proposal. In that case, the set of possible signals consists of all contracts, which is a functional space. To the best of our knowledge, this kind of signaling games are rarely studied before. DeMarzo, Kremer, and Skrzypacz (2005) does consider a security design problem where potential signals are securities. But their approach does not fit our framework of flexible information acquisition. In the literature, either the informed agent chooses finite-dimension signals (e.g., the level of debt in Ross (1977), the retaining fraction of the equity in Leland and Pyle (1977), etc.), or the issuer designs the security before obtaining her information (e.g., DeMarzo and Duffie (1999), Biais and Mariotti (2005)).

1.3.2 Optimal Contract when the Seller Designs

Consider the particular binary choice problem where the agent is a risk neutral buyer with utility (1.10). Action 1 corresponds to buying the ABS $s\left(\overrightarrow{\theta}\right)$ at price q and action 0 corresponds to not buying. Write $m_{s,q}$ for the buyer's optimal strategy when facing contract (s, q). Let

$$\overline{p}_{s,q} = \int_{\Theta} m_{s,q}\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)$$

be the buyer's unconditional probability of accepting the offer. The seller thus enjoys an expected utility

$$W\left(s,q\right) = \int_{\Theta} m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot s\left(\overrightarrow{\theta}\right)\right] dP\left(\overrightarrow{\theta}\right) . \qquad (1.11)$$

The seller's problem is to choose a contract (s, q) satisfying $s\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right]$ to maximize $W\left(s, q\right)$. Let $\left(s^*\left(\cdot\right), q^*\right)$ denote the optimal contract and

$$\overline{p}_{s^*,q^*} = \int_{\Theta} m_{s^*,q^*}\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)$$

be the corresponding probability of trade.

According to Proposition 1.1, there are three possible cases: a) $\overline{p}_{s^*,q^*} = 1$; b) $\overline{p}_{s^*,q^*} = 0$; and c) $\overline{p}_{s^*,q^*} \in (0, 1)$. We first argue that case b) is impossible.

Proposition 1.2 $\overline{p}_{s^*,q^*} > 0$, *i.e., trade happens with positive probability.*

Proof. We prove by constructing a securitized debt that generates positive expected payoff to the seller. Let $\beta \in \left(\delta_s \delta_b^{-1}, 1\right)$ and

$$f\left(q\right) = \int_{\Theta} \min\left(\sum_{n=1}^{N} \theta_n, \beta \delta_s^{-1} q\right) dP\left(\overrightarrow{\theta}\right) .$$

Since P is a continuous distribution and $\beta^{-1}\delta_s\delta_b^{-1} < 1$, there exists $q_0 > 0$ s.t.

$$\Pr\left(\sum_{n=1}^{N}\theta_n \geq \beta\delta_s^{-1}q\right) > \beta^{-1}\delta_s\delta_b^{-1}$$

for all $q \in [0, q_0]$. Hence for any $q \in (0, q_0)$,

$$
\begin{aligned}
f'(q) &= \beta\delta_s^{-1}\int_{\left\{\overrightarrow{\theta}\in\Theta:\sum_{n=1}^{N}\theta_n\geq\beta\delta_s^{-1}q\right\}} 1 \cdot dP\left(\overrightarrow{\theta}\right) \\
&= \Pr\left(\sum_{n=1}^{N}\theta_n \geq \beta\delta_s^{-1}q\right) \cdot \beta\delta_s^{-1} \\
&> \beta^{-1}\delta_s\delta_b^{-1} \cdot \beta\delta_s^{-1} = \delta_b^{-1}.
\end{aligned}
$$

Note that

$$f(0) = 0 \ ,$$

which implies that

$$f(q) > \delta_b^{-1}q$$

for all $q \in (0, q_0)$.

Consider a securitized debt

$$s\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N}\theta_n, D\right)$$

with face value $D = \beta\delta_s^{-1}q$ and price $q \in (0, q_0)$. The buyer's payoff gain from accepting this offer over rejecting it is

$$\Delta u\left(\overrightarrow{\theta}\right) = \delta_b \cdot s\left(\overrightarrow{\theta}\right) - q. \tag{1.12}$$

28

By Jensen's inequality,

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

$$
\geq \exp\left(\mu^{-1}\int_\Theta \Delta u\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)\right)
$$

$$
= \exp\left(\mu^{-1}\left[\delta_b \cdot \int_\Theta \min\left(\sum_{n=1}^{N}\theta_n, \beta\delta_s^{-1}q\right) dP\left(\overrightarrow{\theta}\right) - q\right]\right)
$$

$$
= \exp\left(\mu^{-1}\left[\delta_b \cdot f\left(q\right) - q\right]\right)
$$

$$
> \exp\left(0\right) = 1,
$$

where the last inequality comes from (1.12). Hence according to Proposition 1.1, $\overline{p}_{s,q} > 0$. Then, the seller's expected payoff from this contract is

$$
W\left(s, q\right) = \int_\Theta m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot s\left(\overrightarrow{\theta}\right)\right] dP\left(\overrightarrow{\theta}\right)
$$

$$
= \int_\Theta m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot \min\left(\sum_{n=1}^{N}\theta_n, \beta\delta_s^{-1}q\right)\right] dP\left(\overrightarrow{\theta}\right)
$$

$$
\geq \int_\Theta m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot \beta\delta_s^{-1}q\right] dP\left(\overrightarrow{\theta}\right)
$$

$$
= \left(1 - \beta\right)q \cdot \overline{p}_{s,q} > 0 .
$$

By definition, the seller's expected payoff through the optimal contract is $W\left(s^*, q^*\right) \geq W\left(s, q\right) > 0$. This directly implies $\overline{p}_{s^*,q^*} > 0$ since $\overline{p}_{s^*,q^*} = 0$ always generates zero expected payoff to the seller. This concludes the proof. ∎

The key of the proof is to show that the seller can always enjoy a positive expected payoff through proposing a securitized debt. Hence her optimal contract must also generate a positive expected payoff, which can be achieved only through a successful trade. Although facing adverse selection, the seller always prefers trade. This is because she owns all bargaining power. She is able to minimize the negative effect of information acquisition through appropriately designing a contract and

thus enjoy the benefit from trade.

According to Proposition 1.2, only case a) and c) are possible. In case a) $\overline{p}_{s^*,q^*} = 1$ and the buyer does not acquire any information. In case c), $\overline{p}_{s^*,q^*} \in (0,1)$ and the buyer does acquire some information. We first study the seller's optimal contract in case a).

Optimal Contract without Inducing Information Acquisition

A direct application of Proposition 1.1 suggests that any contract (s,q) that does not induce information acquisition must satisfy

$$\mathbf{E} \exp\left(-\mu^{-1}\left[\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q\right]\right) \leq 1 \ ,$$

i.e.,

$$q \leq -\mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s\left(\overrightarrow{\theta}\right)\right) \ . \tag{1.13}$$

Intuitively, the buyer just accepts the offer when the price is low enough relative to the repayment of the security. This inequality must bind for seller's optimal contract, otherwise she can benefit from increasing the price q. Hence, (1.13) reduces to

$$q = -\mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s\left(\overrightarrow{\theta}\right)\right) \ . \tag{1.14}$$

Since the contract is always accepted, the seller's expected payoff becomes

$$\begin{aligned}
&\int_{\Theta}\left[q - \delta_s \cdot s\left(\overrightarrow{\theta}\right)\right] dP\left(\overrightarrow{\theta}\right) \\
=\ & q - \delta_s \cdot \mathbf{E}s\left(\overrightarrow{\theta}\right) \\
=\ & -\mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s\left(\overrightarrow{\theta}\right)\right) - \delta_s \cdot \mathbf{E}s\left(\overrightarrow{\theta}\right).
\end{aligned}$$

Hence the seller's problem can be formalized as

$$\min_{s(\cdot)} \mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s\left(\overrightarrow{\theta}\right)\right) + \delta_s \cdot \mathbf{E} s\left(\overrightarrow{\theta}\right)$$

subject to the feasibility condition

$$s\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right] . \tag{1.15}$$

Proposition 1.3 *If the seller's optimal contract induces the buyer to always accept it without acquiring information, it must be a securitized debt*

$$s^*\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N} \theta_n, D^*\right)$$

with price q^*, where the face value is determined by

$$
\begin{aligned}
D^* &= D\left(q^*\right) \\
&= \mu\delta_b^{-1} \cdot \left[\ln \delta_b - \ln \delta_s\right] + \delta_b^{-1} q^* ,
\end{aligned}
$$

$q^* > 0$ is the unique fixed point of

$$h\left(q\right) = -\mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot \min\left(\sum_{n=1}^{N} \theta_n, D\left(q\right)\right)\right)$$

and the expectation is taken under common prior P.

Proof. See Appendix 1.5. ∎

First note that the face value has a lower bound, i.e.,

$$D^* > \mu\delta_b^{-1} \cdot \left[\ln \delta_b - \ln \delta_s\right] .$$

31

Hence if the maximal cash flow

$$\sup \left\{ \sum_{n=1}^{N} \theta_n : \overrightarrow{\theta} \in \Theta \right\} \leq \mu \delta_b^{-1} \cdot [\ln \delta_b - \ln \delta_s] \ ,$$

the optimal security is actually the pool of all assets. This could happen when the seller has an extremely good investment opportunity relative to the buyer (i.e., $\ln \delta_b - \ln \delta_s \gg 1$) or it is too hard for the buyer to acquire information (i.e., $\mu \gg 1$). As a direct implication, when the buyer cannot acquire any information (i.e., $\mu \to \infty$), the seller just sells the pool of all assets at price

$$\delta_b \cdot \mathbf{E} \left[\sum_{n=1}^{N} \theta_n \right]$$

and enjoys the maximal trading surplus

$$(\delta_b - \delta_s) \cdot \mathbf{E} \left[\sum_{n=1}^{N} \theta_n \right] \ .$$

Another interesting observation comes from equation (1.14), which implies

$$\begin{aligned} q^* &= -\mu \ln \mathbf{E} \exp \left(-\mu^{-1} \delta_b \cdot s^* \left(\overrightarrow{\theta} \right) \right) \\ &\leq -\mu \ln \left(\exp \left(-\mu^{-1} \delta_b \cdot \mathbf{E} s^* \left(\overrightarrow{\theta} \right) \right) \right) \\ &= \delta_b \cdot \mathbf{E} s^* \left(\overrightarrow{\theta} \right) \ , \end{aligned}$$

where the inequality follows Jensen's inequality. Since the offer induces no information acquisition, both parties remain symmetrically informed and the seller should have charged the buyer $\delta_b \cdot \mathbf{E} s^* \left(\overrightarrow{\theta} \right)$. However, the seller finds it optimal to charge a lower price q^* to bribe the buyer not to acquire information.

In the rest of this section, we show that securitized debt remains uniquely optimal even if there is information acquisition.

Optimal Contract with Information Acquisition

According to Proposition 1.1, any contract $(s(\cdot), q)$ that induces the buyer to acquire information must satisfy

$$\mathbf{E} \exp\left(\mu^{-1} \left[\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q\right]\right) > 1 \tag{1.16}$$

and

$$\mathbf{E} \exp\left(-\mu^{-1} \left[\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q\right]\right) > 1 \ , \tag{1.17}$$

where the expectation is taken according to common prior P. That is, neither accepting nor rejecting the offer is dominant ex ante, and thus the buyer finds it optimal to acquire some information.

Given such a contract, Proposition 1.1 prescribes that the buyer's optimal strategy $m_{s,q}$ is uniquely characterized by

$$\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q = \mu \cdot \left[g'\left(m_{s,q}\left(\overrightarrow{\theta}\right)\right) - g'\left(\overline{p}_{s,q}\right)\right] \ , \tag{1.18}$$

where

$$\overline{p}_{s,q} = \int_\Theta m_{s,q}\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)$$

is the buyer's unconditional probability of accepting the offer.

Taking into account of the buyer's response $m_{s,q}$, the seller chooses $(s(\cdot), q)$ to maximize her expected payoff

$$W(s, q) = \int_\Theta m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot s\left(\overrightarrow{\theta}\right)\right] dP\left(\overrightarrow{\theta}\right)$$

subject to (1.16), (1.17), (1.18) and the feasibility condition

$$s\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right] . \qquad (1.19)$$

It is worth noting that both (1.16) and (1.17) should not bind for the optimal contract, otherwise no information will be acquired according to Proposition 1.1. Hence, conditional on the fact that the optimal contract does induce information acquisition, these two constraints could be ignored during optimization.

We derive the optimal contract $(s^*(\cdot), q^*)$ through calculus of variations. That is, see how the seller's expected payoff responds to the perturbation of her optimal contract.

Let $s\left(\overrightarrow{\theta}\right) = s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right)$ be an arbitrary perturbation of $s^*(\cdot)$. The buyer's best response $m_{s,q^*}(\cdot)$ is implicitly determined by $s(\cdot)$ through functional equation (1.18). Hence we need first characterize how $m_{s,q^*}(\cdot)$ varies with respect to the perturbation of $s^*(\cdot)$.

Lemma 1.1 *For any perturbation* $s\left(\overrightarrow{\theta}\right) = s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right)$, *the response of the buyer's strategy* $m_{s,q^*}(\cdot)$ *is characterized by*

$$
\begin{aligned}
\left.\frac{dm_{s,q^*}\left(\overrightarrow{\theta}\right)}{d\alpha}\right|_{\alpha=0} =\ & \mu^{-1}\delta_b \cdot \left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \varepsilon\left(\overrightarrow{\theta}\right) \\
& + \frac{\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \mu^{-1}\delta_b \int_{\Theta}\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \varepsilon\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)}{\left[g''\left(\overline{p}_{s^*,q^*}\right)\right]^{-1} - \int_{\Theta}\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} dP\left(\overrightarrow{\theta}\right)}
\end{aligned}
\qquad (1.20)
$$

Proof. See Appendix 1.5. ∎

The first term of the right hand side of (1.20) is the buyer's local response to $\varepsilon\left(\overrightarrow{\theta}\right)$. It is of the same sign as the perturbation $\varepsilon\left(\overrightarrow{\theta}\right)$. When the repayment increases at state $\overrightarrow{\theta}$, the buyer is more likely to accept the offer at this state. The

second term measures the buyer's average response to perturbation $\varepsilon\left(\overrightarrow{\theta}\right)$ over all states. It is straightforward to verify that the denominator is positive due to Jensen's inequality. Hence, if on average the perturbation increases her repayment, the buyer would like to accept the offer more often.

Now we can calculate the variation of the seller's expected payoff $W\left(s,q^*\right)$. Taking derivative with respect to α at $\alpha=0$ for both sides of (1.11) leads to

$$\left.\frac{dW\left(s,q^*\right)}{d\alpha}\right|_{\alpha=0}=\left.\int_{\Theta}\frac{dm_{s,q^*}\left(\overrightarrow{\theta}\right)}{d\alpha}\right|_{\alpha=0}\left[q^*-\delta_s s^*\left(\overrightarrow{\theta}\right)\right]dP\left(\overrightarrow{\theta}\right)-\delta_s\int_{\Theta}m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\varepsilon\left(\overrightarrow{\theta}\right)dP\left(\overrightarrow{\theta}\right)$$

$$\tag{1.21}$$

Substitute (1.20) into (1.21) and manipulate we get

$$\left.\frac{dW\left(s,q^*\right)}{d\alpha}\right|_{\alpha=0}=\int_{\Theta}r\left(\overrightarrow{\theta}\right)\cdot\varepsilon\left(\overrightarrow{\theta}\right)dP\left(\overrightarrow{\theta}\right)\ ,\tag{1.22}$$

where

$$r\left(\overrightarrow{\theta}\right)=-\delta_s m_{s^*,q^*}\left(\overrightarrow{\theta}\right)+\mu^{-1}\delta_b\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1}\left(q^*-\delta_s\cdot s^*\left(\overrightarrow{\theta}\right)+w\right)$$

$$\tag{1.23}$$

and

$$w=\frac{\int_{\Theta}\left[q^*-\delta_s\cdot s^*\left(\overrightarrow{\theta}\right)\right]\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1}dP\left(\overrightarrow{\theta}\right)}{\left[g''\left(\overline{p}_{s^*,q^*}\right)\right]^{-1}-\int_{\Theta}\left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1}dP\left(\overrightarrow{\theta}\right)}.$$

Note that w is a constant that does not depend on $\overrightarrow{\theta}$. Its value is endogenously determined in equilibrium. Here $r\left(\overrightarrow{\theta}\right)$ is the Frechet derivative[13] of $W\left(s,q^*\right)$ at s^*, it measures the marginal contribution of any perturbation to the seller's expected payoff. The first term of (1.23) is the direct contribution of perturbing s^* disregarding the variation of $m_{s^*,q^*}\left(\overrightarrow{\theta}\right)$. The second term measures the indirect

---

[13]For the readers not familiar with this concept, just think of Frechet derivative as the gradient of $W\left(s,q^*\right)$ at "vector" s. Indeed, the gradient is a special case of Frechet derivative when $\#\Theta$ is finite.

contribution through the variation of $m_{s^*,q^*}\left(\overrightarrow{\theta}\right)$. This expression represents the chain rule of the calculus of variations.

Let

$$A_0 = \left\{ \overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) = 0 \right\},$$

$$A_1 = \left\{ \overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) \in \left(0, \sum_{n=1}^{N} \theta_n \right) \right\}$$

and

$$A_2 = \left\{ \overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) = \sum_{n=1}^{N} \theta_n \right\}.$$

In regions A_0 and A_2, $s^*\left(\cdot\right)$ is bounded by its lower bound and upper bound, respectively. In region A_1, $s^*\left(\cdot\right)$ is off the boundaries. Then $\{A_0, A_1, A_2\}$ is a partition of $\Theta \backslash \left\{ \overrightarrow{0} \right\}$. Since $s^*\left(\cdot\right)$ is the optimal security,

$$\left. \frac{dW\left(s, q^*\right)}{d\alpha} \right|_{\alpha=0} \leq 0$$

holds for any feasible[14] perturbation $\varepsilon\left(\overrightarrow{\theta}\right)$. Hence (1.22) implies

$$r\left(\overrightarrow{\theta}\right) \begin{cases} \leq 0 & \text{if } \overrightarrow{\theta} \in A_0 \\ = 0 & \text{if } \overrightarrow{\theta} \in A_1 \\ \geq 0 & \text{if } \overrightarrow{\theta} \in A_2 \end{cases}. \tag{1.24}$$

---

[14]A perturbation ε is feasible with respect to s^* if $\exists \alpha > 0$, s.t. $\forall \overrightarrow{\theta} \in \Theta$, $s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n \right]$.

Since g is strictly convex, $g'' > 0$ and (1.24) can be rewritten as

$$r\left(\overrightarrow{\theta}\right) \cdot g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)$$

$$= -\delta_s m_{s^*,q^*}\left(\overrightarrow{\theta}\right) g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right)$$

$$\begin{cases} \leq 0 & \text{if } \overrightarrow{\theta} \in A_0 \\ = 0 & \text{if } \overrightarrow{\theta} \in A_1 \\ \geq 0 & \text{if } \overrightarrow{\theta} \in A_2 \end{cases} . \tag{1.25}$$

Recall that given the optimal contract $(s^*(\cdot), q^*)$, the buyer's best response $m_{s^*,q^*}\left(\overrightarrow{\theta}\right)$ is characterized by

$$\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) - q^* = \mu \cdot \left[g'\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) - g'\left(\overline{p}_{s^*,q^*}\right)\right] , \tag{1.26}$$

where

$$\overline{p}_{s^*,q^*} = \int_{\Theta} m_{s^*,q^*}\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)$$

is the buyer's unconditional probability of accepting the optimal contract $(s^*(\cdot), q^*)$.

Then, $(1.25)^{15}$ together with (1.26) determines the optimal contract $(s^*(\cdot), q^*)$. Let $m = f_1(s)$ and $m = f_2(s)$ be the two continuous functions implicitly defined by

$$-\delta_s \cdot m \cdot g''(m) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s + w\right) = 0 \tag{1.27}$$

and

$$\delta_b \cdot s - q^* = \mu \cdot \left[g'(m) - g'\left(\overline{p}_{s^*,q^*}\right)\right] , \tag{1.28}$$

respectively. We have $f_1'(s) < 0$ and $f_2'(s) > 0$ since $[m \cdot g''(m)]' > 0$ and $g''(m) >$

---

[15]One may criticize that Equation (1.25) is just the first order condition of the seller's optimization problem. It only characterizes the critical points. In principle, we should characterize the largest critical point. However, our argument works for any critical point and thus our results are immune to this critique.

0. By definition,

$$m_{s^*,q^*}\left(\overrightarrow{\theta}\right) = f_1\left(s^*\left(\overrightarrow{\theta}\right)\right) \text{ implies } r\left(\overrightarrow{\theta}\right) = 0 \ .$$

Also note that $m_{s^*,q^*}\left(\overrightarrow{\theta}\right) = f_2\left(s^*\left(\overrightarrow{\theta}\right)\right)$ for all $\overrightarrow{\theta} \in \Theta$. Now we can characterize the optimal security through analyzing f_1 and f_2 together.

Proposition 1.4 $\Pr\left(A_0\right) = 0$, where $A_0 = \left\{\overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) = 0\right\}$.

Proof. See Appendix 1.5. ∎

This proposition states that constraint $s\left(\overrightarrow{\theta}\right) \geq 0$ never binds. The logic underlying the proof is that on the boundary $s\left(\overrightarrow{\theta}\right) = 0$, although an increment of $s\left(\overrightarrow{\theta}\right)$ increases the seller's repayment, it increases the probability of trading even more. Hence the seller on average gains through deviating from the lower boundary. As its implication, it is not optimal to issue equity residual/call option to raise liquidity.

For those states in A_1, where the limited liability constraint

$$s\left(\overrightarrow{\theta}\right) \leq \sum_{n=1}^{N} \theta_n$$

does not bind either, both

$$m_{s^*,q^*}\left(\overrightarrow{\theta}\right) = f_1\left(s^*\left(\overrightarrow{\theta}\right)\right)$$

and

$$m_{s^*,q^*}\left(\overrightarrow{\theta}\right) = f_2\left(s^*\left(\overrightarrow{\theta}\right)\right)$$

must hold. Since $f_1'\left(s\right) < 0$ and $f_2'\left(s\right) > 0$, $f_1\left(s\right)$ and $f_2\left(s\right)$ intersect at most once. Hence $s^*\left(\overrightarrow{\theta}\right)$ should be a constant and the buyer has no incentive to acquire information within region A_1. This result coincides with our intuition. If

the limited liability constraint never binds, the seller would issue a security with constant repayment to avoid the buyer's information acquisition. However, once the underlying cash flows are too low to support such constant, $s^*\left(\overrightarrow{\theta}\right)$ reaches the limited liability boundary and equals $\sum_{n=1}^{N}\theta_n$. The next proposition shows that the optimal security must be a securitized debt.

Proposition 1.5 *If the seller's optimal contract induces the buyer to acquire information, it must be a securitized debt* $s^*\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N}\theta_n, D^*\right)$.

Proof. See Appendix 1.5. ■

Together with Proposition 1.2 and 1.3, this proposition enables us to conclude that pooling the assets and issuing a senior tranche is always the uniquely optimal way to raise liquidity. Pooling is directly derived from the seller's desire to maximize liquidity. It has nothing to do with the consideration of risk diversification since both agents are risk-neutral. The flat tail of the optimal security results from the seller's effort to minimize her opponent's information acquisition. In contrast to the non-uniqueness result in Dang, Gorton, and Holmstrom (2010), we can show the unique optimality of debt because of our flexible information acquisition framework. In Dang, Gorton, and Holmstrom (2010), only two extreme information structures are available in the setup of costly state verification while infinite forms of securities can be designed, which inevitably results in the indistinguishability of some securities. In our framework, with help of flexibility, the variety of available information structures matches the variety of potential securities to be designed, and thus the uniqueness of the standard securitized debt could be guaranteed. Quasi-debts are no longer optimal in our model. By reshaping the uneven tail above the price of a quasi-debt to a flat one, not only the buyer's information cost could be saved but also potential loss of trade from adverse selection could be mitigated. The resulted surplus could be employed by the seller to make both parties better off, and thus ultimately make

a better provision of liquidity possible. Moreover, this flexibility also enables us to show the optimality of pooling and tranching in a broader class of environments than Dang, Gorton, and Holmstrom (2010) and without assuming a sufficiently large number of underlying assets as in DeMarzo (2005)[16].

In addition, while most models in literature are built upon specific assumptions about the cash flows, our qualitative result does not rely on such distributional details of underlying assets. Since the stochastic interdependence among the underlying assets could be complex and violate such assumptions, our model provides a better explanation for the prevalence of securitization in financial markets.

The security design literature usually assumes Monotone Likelihood Ratio Property (MLRP) or similar conditions to guarantee a meaningful result. Our framework justifies this assumption through endogenizing the information structure. According to Proposition 1.5, the optimal security $s^* \left(\overrightarrow{\theta} \right)$ is non-decreasing in the sum of cash flows. Proposition 1.1 implies that the best information structure $m_{s^*,q^*} \left(\overrightarrow{\theta} \right)$ is increasing in the payoff gain $\delta_b \cdot s^* \left(\overrightarrow{\theta} \right) - q^*$. Hence $m_{s^*,q^*} \left(\overrightarrow{\theta} \right)$ is also non-decreasing in the sum of the cash flows. Therefore, the larger the cash flows, the higher the probability that the buyer gets a signal asking her to accept. This can be interpreted as a generalized MLRP for multi-dimensional states.

To facilitate the analysis, the security design literature usually restrict their attention to the set of "regular" securities, which are non-decreasing in the underlying cash flows (e.g., DeMarzo and Duffie (1999), DeMarzo (2005)). We do not have such restriction, but show that the optimal security naturally turns out to be non-decreasing.

Finally, Dang, Gorton, and Holmstrom (2010) get debt contract uniquely optimal when their fixed information cost is zero. This can be viewed as a special case

---

[16]DeMarzo (2005) shows that the benefit of pooling achieves a theoretical maximum as the number of underlying assets approaches infinity.

of our model where marginal cost of information acquisition vanishes.

Understanding the Origin of Uniqueness

For readers familiar with the approach of costly state verification (CSV), a question naturally arises regarding the uniqueness of the optimal contract. Both Townsend (1979) and Dang, Gorton, and Holmstrom (2010) employ CSV, why does the former but not the latter get debt uniquely optimal? In last subsection, we have attributed the non-uniqueness in Dang, Gorton, and Holmstrom (2010) to the rigidity of CSV. This argument is correct when comparing Dang, Gorton, and Holmstrom (2010) to our model, but not fully convincing when Townsend (1979) is also considered. To fully understand the different results in Dang, Gorton, and Holmstrom (2010), Townsend (1979) and our model, we first highlight the essence of flexibility. In principle, general flexible choice, not necessarily restricted to flexible information acquisition, enables an economic agent to make state-contingent responses. In other words, the agent can make a best response in one state, and can make another best response in another state. In all these three models, the contract designer is endowed with flexibility, in the sense that she can assign state-contingent repayment through designing any form of security. What matters to shape the different results regarding uniqueness of the optimal contract relies on the potential flexibility of the other party who decides whether to accept the offer. Through comparing these three models, we argue that the origin of the uniqueness is not only from the flexibility itself, but from the double-sided symmetry of flexibility. Here, double-sided symmetry of flexibility requires that both parties engaged in a potential trade are endowed with the same level of flexibility.

In our framework, ex-ante symmetric information in the form of a double-sided ignorance prevents the buyer to make a state-contingent choice if she only follows the traditional CSV approach to acquire information. However, the buyer in our

41

framework is able to choose state-contingent probability (i.e., $m\left(\overrightarrow{\theta}\right)$) of accepting the offer, namely, she can perform flexible information acquisition. In this sense the buyer enjoys the same level of flexibility as the seller. Given this double-sided symmetry of flexibility in our model, the uniqueness of an optimal contract, which is the standard securitized debt, is guaranteed. In Dang, Gorton, and Holmstrom (2010), however, the buyer can only follow the traditional CSV approach to acquire information, in which only two options, namely, to acquire a signal or not, are available. Moreover, ex-ante symmetric ignorance precludes the possibility of conditioning the action on any private information. Hence the CSV makes the buyer in Dang, Gorton, and Holmstrom (2010) unable to make state-contingent decision. As a result, the desired double-sided symmetry of flexibility fails and the uniqueness of the optimal contract fails as a consequence. Interestingly, Townsend (1979) also employs the costly state verification approach with two options to model information acquisition, namely, to audit or not, but the unique optimality of a standard debt still emerges. Why it is this case? Different from Dang, Gorton, and Holmstrom (2010) and our framework, in Townsend (1979) the entrepreneur has information advantage over the lender in the sense that the entrepreneur knows the realized profit of the project which the lender does not know. Thanks to the revelation principle, the lender who acquire information in the interim stage can decide whether to audit or not in any state based on the truth told by the entrepreneur who has private information. In other words, although the lender in Townsend (1979) still only has two options to acquire information as the buyer in Dang, Gorton, and Holmstrom (2010), such two options in Townsend (1979) are state-contingent while their counterparts in Dang, Gorton, and Holmstrom (2010) are not. Therefore, the double-sided symmetry of flexibility is still established in Townsend (1979), and the uniqueness of the optimal contract, also a standard debt, is ensured in their model as well. Figure 1.2 shows the relation among these three models.

	flexible info. acquisition $\Longrightarrow$ state-contingent response $m(\theta)$	ex-ante symmetric info./double-sided ignorance
OUR MODEL		
DANG, GORTON & HOLMSTROM (2010)	CSV: rigid info. acquisition	ex-ante symmetric info./double-sided ignorance
TOWNSEND (1979)	CSV: rigid info. acquisition	asymmetric info. + revelation principle $\Longrightarrow$ state-contingent audit

Figure 1.2: Relation among our model, dang, Gorton & Holmstrom (2010) and townsend (1979)

This subsection explores the origin of uniqueness of the optimal contract. We address the optimality of securitized debt in next subsection.

Two Key Factors Driving the Optimality of Securitized Debt

Although our model explains the popularity of securitized debt contracts, it is important to figure out the boundary of our theory. In this subsection, we propose two key factors that drive our results. We show that issuing securitized debt is no longer optimal in absence of these factors.

The first feature of our model is its fixed aggregate risk. Before designing the contract, the seller has already owned assets $\overrightarrow{\theta}$. Hence the assets owned by the seller and the buyer as a whole is invariant with respect to the success or failure of the transaction. This fixed aggregate risk leads to conflicting interests of the two parties, where any information acquired by the buyer makes herself better off but hurts the seller's benefit through adverse selection. That is, the buyer attempts to

43

acquire information that helps her reject the offer once the repayment is lower than the price and accept the offer in the opposite case. However, whatever quantity and quality of information is acquired has nothing to do with their aggregate risk.

The importance of this factor can be seen clearly in our derivation of the optimal security. Since the buyer's incentive to acquire information and the seller's incentive to design the security are totally shaped by their payoff gains from the success over the failure of the transaction, it makes sense to examine their payoff gains. Conditional on $\overrightarrow{\theta}$, the buyer's and seller's payoff gains are

$$\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q$$

and

$$q - \delta_s \cdot s\left(\overrightarrow{\theta}\right) \;,$$

respectively. Both these payoff gains do not explicitly depend on $\overrightarrow{\theta}$. The future cash flows $\overrightarrow{\theta}$ can affect their incentives only through the security $s\left(\overrightarrow{\theta}\right)$. This is the reason that we can define the functions $m = f_1(s)$ and $m = f_2(s)$ rather than $m = f_1\left(s, \overrightarrow{\theta}\right)$ or $m = f_2\left(s, \overrightarrow{\theta}\right)$ in (1.27) and (1.28). The simple shape of securitized debt comes from this independence of f_1 and f_2 on $\overrightarrow{\theta}$.

To make our point more clear, we consider a similar problem with variable aggregate risk. The seller is an entrepreneur who wants to raise capital q to take a project that generates cash flow θ. The project requires a total investment $\overline{q}$, which is financed by a bank as well as the entrepreneur's own capital $\overline{q} - q$. As before, she designs a security $s(\theta)$ and proposes a take-it-or-leave-it offer (s, q) to the bank, who is the buyer that acquires information in the present problem. The entrepreneur's project gets funded and generates future cash flow θ only if the bank accepts the offer. Hence, the aggregate risk depends on whether the transaction succeeds. In this case, the buyer's payoff gain remains the same but the seller's

payoff gain becomes

$$\delta_s \cdot [\theta - s(\theta)] - (\bar{q} - q) \ ,$$

which explicitly depends on θ. As a result, we have $m = f_1(s, \theta)$ rather than $m = f_1(s)$ and the flat part of the debt is no longer optimal. Even if $s(\theta)$ is off the boundaries, the seller would like to fluctuate $s(\theta)$ to induce the buyer to acquire some information. In general, information acquisition benefits the buyer and seller as a whole. It prevents the project to be taken when the cash flow is too low. In fact, this is a story of consulting. The seller designs a state contingent repayment to elicit information from the buyer. Their incentives are aligned rather than opposite to each other.

The second factor that drives our results is homogeneous information acquisition. That is, no state is more special than other states in terms of the difficulty of information acquisition. This property stems from rational inattention[17] and is the reason why our qualitative result does not depend on the stochastic interdependence among the underlying assets. Recall the binary decision problem in Section 1.2, the decision maker's optimal strategy m is characterized by equation (1.7)

$$\Delta u(\theta) = \mu \cdot [g'(m(\theta)) - g'(p_1)] \ ,$$

where

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

The right hand side of equation (1.7) is the Frechet derivative[18] of information cost. It does not explicitly depends on θ. This is the homogeneity we referred to. As an

---

[17]There are many information cost functions satisfying this property. For example, any strictly concave and symmetric function g in (1.1) corresponds to an information cost with this property.

[18]For the readers not familiar with this concept, just think of the Frechet derivative as the gradient of the cost function.

example, homogeneity fails if we replace the term

$$g'(m(\theta)) - g'(p_1)$$

with

$$g'(m(\theta)) - g'(p_1) + k(\theta)$$

for some non-constant function $k(\theta)$. In this case, we should define $m = f_2(s, \theta)$ instead of $m = f_2(s)$ in (1.28). This dependence reflects the buyer's varying difficulties in discerning different states. Hence the optimal contract may not have a flat part as in debt.

We use a non-homogeneous information cost to illustrate our idea. Specifically, let $\theta \in [0, 1]$ and

$$c(m) = \frac{\mu}{\Pr(\theta \in [0, a])} \cdot \left[\int_{[0,a]} g(m(\theta)) \, dP(\theta) - g\left(\int_{[0,a]} m(\theta) \, dP(\theta) \right) \right]$$

for some $a \in (0, 1)$. Hence the state is directly observable for $\theta \in (a, 1]$. For $\theta \in [0, a]$, the buyer can acquire its information at marginal cost $\frac{\mu}{\Pr(\theta \in [0,a])}$. Let $\delta_b = 1$ and the seller's optimal contract be (s, q). Given this contract, the buyer's optimal strategy is characterized by

$$s(\theta) - q = \mu \cdot [g'(m(\theta)) - g'(p_1)] \text{ if } \theta \in [0, a] \ ,$$

and

$$m(\theta) = \begin{cases} 1 \text{ if } \theta \in (a, 1] \text{ and } s(\theta) - q \geq 0 \\ 0 \text{ if } \theta \in (a, 1] \text{ and } s(\theta) - q < 0 \end{cases},$$

where

$$p_1 = \frac{\int_{[0,a]} m(\theta) \, dP(\theta)}{\Pr(\theta \in [0, a])} \ .$$
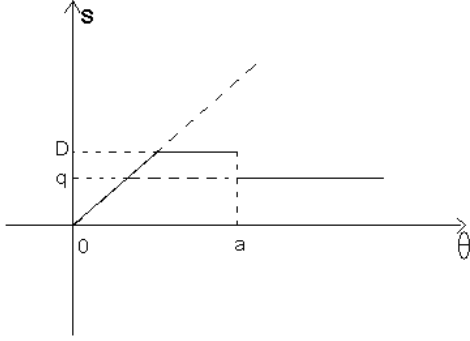
Figure 1.3: optimal contract under non-homogeneous information cost

For $\theta \in (a, 1]$, the buyer accepts the offer if and only if $s(\theta) - q \geq 0$, thus we must have

$$s(\theta) = q$$

for $\theta \in (a, 1]$. Information remains costly in region $[0, a]$, thus a debt contract is optimal within this region according to our previous argument. However, the optimal contract on interval $[0, 1]$ is no longer a debt, as shown in Figure 1.3.

1.3.3 Allocation of Bargaining Power

One may wonder if our results are sensitive to the allocation of bargaining power. The answer is no. This subsection introduces the case where the buyer owns bargaining power and then presents the main results. Due to the similarity between the two cases, we omit most proofs here.

Suppose the buyer proposes the contract $(s(\cdot), q)$ and the seller acquires information. Write $m_{s,q}$ for the seller's optimal strategy. The uninformed buyer thus enjoys expected payoff

$$W(s, q) = \int m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q\right] dP\left(\overrightarrow{\theta}\right).$$

47

The buyer's problem is to choose a feasible contract (s, q) satisfying $s\left(\overrightarrow{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right]$ to maximize $W(s, q)$. Let $(s^*(\cdot), q^*)$ denote the optimal contract for the buyer and

$$\overline{p}_{s^*, q^*} = \int_{\Theta} m_{s^*, q^*}\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right)$$

be the corresponding probability of trade.

Proposition 1.6 $\overline{p}_{s^*, q^*} > 0$, *i.e., trade happens with positive probability.*

Proof. See Appendix 1.5. ∎

Proposition 1.7 *If the buyer's optimal contract induces the seller to always accept it without acquiring information, it must be a securitized debt*

$$s^*\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N} \theta_n, D^*\right)$$

with price q^*, where

$$D^* = \mu \delta_s^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_s^{-1} q^* ,$$

q^* is the unique fixed point of

$$h(q) = \mu \ln \mathbf{E} \exp\left(\mu^{-1} \delta_s \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu \delta_s^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_s^{-1} q\right)\right)$$

and the expectation is taken according to common prior P.

Proof. The proof is very similar to that of Proposition 1.3 and is omitted here. ∎

Proposition 1.8 *If the buyer's optimal contract induces the seller to acquire information, it must be a securitized debt $s^*\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N} \theta_n, D^*\right)$.*

Proof. The proof is very similar to that of Proposition 1.5 and is omitted here. ∎

Proposition 1.3, 1.5, 1.7 and 1.8 show that the optimal security is always a securitized debt, no matter who owns bargaining power.[19] This result is consistent with our previous analysis. Exchanging bargaining power does not change the facts that aggregate risk is fixed and information acquisition is homogeneous.

1.4   Conclusions and Discussions

This chapter studies liquidity provision in presence of endogenous and flexible information acquisition. In our model, there is no information asymmetry before bargaining. Also, the buyer has an expertise in acquiring information of the fundamental in the manner of rational inattention. She collects the most payoff-relevant information according to the contract proposed to her, which may endogenously generate adverse selection. Hence, the seller deliberately designs the security in order to induce the buyer to acquire information least harmful to the seller's interest. It is shown that pooling and issuing securitized debt is the uniquely optimal way to raise liquidity, regardless of the stochastic interdependence among the underlying assets and the allocation of bargaining power. Compared to the security design literature, our results are clearer. We neither restrict our attention to non-decreasing securities nor impose various assumptions on information structures like MLRP. Instead, these properties of the optimal security are justified in equilibrium. Our results are driven by two key factors. The one is the fixed aggregate risk and the other is homogeneous information cost, without which the securitized debt may not be optimal.

The role of fixed aggregate risk sheds light on a general classification of information, namely, to classify what information is socially valuable and what information

---

[19]However, reallocating the bargaining power does affect the face value and price of the debt, and thus affects the agents' expected payoffs.

is not. In particular, flexibility enables economic agents to acquire these two types of information separately, which results in different welfare implications of information acquisition. At the level of the society, acquisition of information that is not socially valuable not only wastes social resource but also leads to endogenous adverse selection, which in turn harms social welfare. Hence, desired organizational form of the society should deter acquisition of such information. On the contrary, acquisition of socially valuable information generally increases social welfare and thus should be encouraged in principle. In our model with fixed aggregate risk, none of information is socially valuable, so that securitized debt is optimal because it best deters information acquisition. On the other hand, as the example mentioned with variable aggregate risk, some certain information is socially valuable as it helps prevent investing in bad states. Consequently, acquisition of such socially valuable information should be encouraged, and thus securitized debt may not be the optimal contract. This classification of information also provides a new perspective to look into the mutual existence of debt and equity, both as popular forms of financial contracts in reality. For start-ups and projects with high risk, issuing equity could be more desirable because it encourages acquisition of socially valuable information, which helps to screen projects and control the aggregate risk of the entire society. In contrast, for mature corporations with robust growth, in which the provision of liquidity is of the priority, debt could be more desirable as it deters unnecessary acquisition of information that is not socially valuable. This consideration is partly consistent with the well-known packing-order theory, and future work may further unify the life-cycle evolution of capital structure of corporations along the line of flexible information acquisition.

Under a similar mentality, flexibility also helps revisit the endogenous determination of capital structure in literature by specializing information acquisition. Given flexible information acquisition, agents who monitor may have different incentives

in acquiring different information regarding various forms of financial contracts. Hence, different layers of financial contracts in certain capital structure enable a specialization of information acquisition. In other words, layers of capital structure correspond to specialized layers of information to be acquired. This specialization may in turn affect production of information as well as efficiency of monitoring, and further reshape the optimal capital structure. In this way, it is seen that flexibility plays an role in determining the capital structure, and more results regarding its effects on corporate finance as well as social welfare are to be expected.

1.5 Appendix

Proof of Proposition 1.1.

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

$$
\begin{aligned}
&V^* (m + \alpha \cdot \varepsilon) \\
= \ &\int_\Theta (m(\theta) + \alpha \cdot \varepsilon(\theta)) \cdot \Delta u(\theta) \, dP(\theta) \\
&- \mu \cdot \left[\int_\Theta g(m(\theta) + \alpha \cdot \varepsilon(\theta)) \, dP(\theta) - g\left(\int_\Theta [m(\theta) + \alpha \cdot \varepsilon(\theta)] \, dP(\theta) \right) \right] ,
\end{aligned}
$$

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

$$
\begin{aligned}
\left. \frac{dV^*(m + \alpha \cdot \varepsilon)}{d\alpha} \right|_{\alpha = 0} = \ &\int_\Theta \varepsilon(\theta) \cdot \Delta u(\theta) \, dP(\theta) \\
&- \mu \cdot \left[\int_\Theta \varepsilon(\theta) \cdot g'(m(\theta)) \, dP(\theta) - g'\left(\int_\Theta m(\theta) \, dP(\theta) \right) \cdot \int_\Theta \varepsilon(\theta) \, dP(\theta) \right] \\
= \ &\int_\Theta \varepsilon(\theta) \cdot [\Delta u(\theta) - \mu \cdot (g'(m(\theta)) - g'(p_1))] \, dP(\theta) \ .
\end{aligned}
$$

Note that

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

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

$$\left\{\begin{array}{c} \widetilde{m} \in M : V^*\left(\widetilde{m}\right) - V^*\left(m\right) = \int_\Theta \left[\Delta u\left(\theta\right) - \mu g'\left(m\left(\theta\right)\right) + \mu g'\left(\int_\Theta m\left(\theta\right) dP\left(\theta\right)\right)\right] \\ \cdot \left(\widetilde{m}\left(\theta\right) - m\left(\theta\right)\right) dP\left(\theta\right) \end{array}\right\}.$$

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

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

This inequality is strict when

$$m \in M^o \triangleq M \backslash \left\{m \in M : m\left(\theta\right) \text{ is a constant almost surely}\right\}$$

and $\Pr\left(\widetilde{m}\left(\theta\right) \neq m\left(\theta\right)\right) > 0$, since $V^*\left(\cdot\right)$ is strictly concave on M^o. We will use this observation later in this proof.

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

$$\Delta u\left(\theta\right) - \mu \cdot \left(g'\left(m\left(\theta\right)\right) - g'\left(p_1\right)\right) \left\{\begin{array}{ll} \geq 0 & \text{if } m\left(\theta\right) = 1 \\ = 0 & \text{if } m\left(\theta\right) \in \left(0, 1\right) \\ \leq 0 & \text{if } m\left(\theta\right) = 0 \end{array}\right. . \tag{1.29}$$

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

$$p_1 = \int_{\Theta} m\left(\theta\right) dP\left(\theta\right) < 1$$

and for $\theta \in B = \{\theta \in \Theta : m\left(\theta\right) = 1\}$,

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

since $\lim_{x \to 1} g'\left(x\right) = \infty$. Then $\varepsilon\left(\theta\right) = -1_B$ is a feasible perturbation and

$$
\begin{aligned}
\frac{dV^*\left(m + \alpha \cdot \varepsilon\right)}{d\alpha}\bigg|_{\alpha=0} \\
= \int_{\Theta} \left[\Delta u\left(\theta\right) - \mu \cdot \left(g'\left(m\left(\theta\right)\right) - g'\left(p_1\right)\right)\right] \cdot \varepsilon\left(\theta\right) dP\left(\theta\right) \\
= \int_B \left(-\infty\right) \cdot \left(-1\right) dP\left(\theta\right) \\
= +\infty \, ,
\end{aligned}
$$

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

We first search for the sufficient condition for scenario c). According to (1.29), we have

$$\Delta u\left(\theta\right) - \mu \cdot \left(g'\left(m\left(\theta\right)\right) - g'\left(p_1\right)\right) = 0 \; a.s.. \tag{1.30}$$

By definition,

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

thus (1.30) implies

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

Let

$$M_1 = \left\{ m(\theta, p) = \frac{p}{p + (1-p) \cdot \exp\left(-\mu^{-1}\Delta u(\theta)\right)} : p \in [0, 1] \right\} \qquad (1.31)$$

and

$$J(p) = \int_\Theta m(\theta, p)\, dP(\theta) \ ,$$

then there exists $p_1 \in [0, 1]$ such that $m(\cdot, p_1) \in M_1 \subset M$ is an optimal strategy.

Note that $J(p_1) = p_1$ is a necessary condition.

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

$$\max_{p \in [0,1]} V^*(m(\cdot, p)) = \int_\Theta \Delta u(\theta) \cdot m(\theta, p)\, dP(\theta) - c(m(\cdot, p)) \ .$$

The first order derivative with respect to p is

$$\frac{dV^*(m(\cdot, p))}{dp}$$
$$= \int_\Theta \Delta u(\theta) \cdot \frac{\partial m(\theta, p)}{\partial p} dP(\theta)$$
$$\quad -\mu \cdot \left[\int_\Theta g'(m(\theta, p)) \frac{\partial m(\theta, p)}{\partial p} dP(\theta) - g'\left(\int_\Theta m(\theta, p)\, dP(\theta) \right) \int_\Theta \frac{\partial m(\theta, p)}{\partial p} dP(\theta) \right]$$
$$= \int_\Theta \left[\Delta u(\theta) - \mu \cdot g'(m(\theta, p)) + \mu \cdot g'(J(p)) \right] \cdot \frac{\partial m(\theta, p)}{\partial p} dP(\theta) \ .$$

By definition,

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

thus

$$\frac{dV^{*}\left(m\left(\cdot,p\right)\right)}{dp} = \int_{\Theta}\left[-\mu\cdot g'\left(p\right)+\mu\cdot g'\left(J\left(p\right)\right)\right]\cdot\frac{\partial m\left(\theta,p\right)}{\partial p}dP\left(\theta\right)$$

$$= \mu\cdot\left[g'\left(J\left(p\right)\right)-g'\left(p\right)\right]\cdot\int_{\Theta}\frac{\partial m\left(\theta,p\right)}{\partial p}dP\left(\theta\right) \ . \qquad (1.32)$$

Since

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

$$= \left[p\cdot\exp\left(\frac{1}{2}\mu^{-1}\Delta u\left(\theta\right)\right)+\left(1-p\right)\cdot\exp\left(-\frac{1}{2}\mu^{-1}\Delta u\left(\theta\right)\right)\right]^{-2}$$

$$> 0$$

for all $\theta\in\Theta$,

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

if and only if

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

Since g' is strictly increasing in its argument, we have

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

if and only if

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

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

$$J\left(p_1\right)=p_1 \ . \qquad (1.33)$$

But (1.33) is not sufficient. The sufficient condition for $m\left(\cdot,p_1\right)$ to be a local max-

imum within M_1 is

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

$$\text{s.t. } J(p) \geq p \text{ for all } p \in (p_1 - \beta, p_1]$$

$$\text{and } J(p) \leq p \text{ for all } p \in [p_1, p_1 + \beta).$$

Note that

$$J(0) = 0, J(1) = 1,$$

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

and

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

Case i):

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

and

$$\int_\Theta \exp\left(-\mu^{-1} \Delta u(\theta)\right) dP(\theta) > 1.$$

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

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

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

$$= 0$$

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

observation mentioned above implies

$$V^* (\widetilde{m}) - V^* (m (\cdot, p_1))$$

$$< \int_\Theta \left[\Delta u (\theta) - \mu \cdot g' (m (\cdot, p_1)) + \mu \cdot g' \left(\int_\Theta m (\cdot, p_1) \, dP (\theta) \right) \right] (\widetilde{m} (\theta) - m (\cdot, p_1)) \, dP (\theta)$$

$$= \int_\Theta [\Delta u (\theta) - \mu \cdot g' (m (\cdot, p_1)) + \mu \cdot g' (J (p_1))] (\widetilde{m} (\theta) - m (\cdot, p_1)) \, dP (\theta)$$

$$= 0$$

for all $\widetilde{m} \in M$ such that $\Pr (\widetilde{m} (\theta) \neq m (\theta, p_1)) > 0$. Hence, $V^* (m (\cdot, p_1))$ is strictly higher than the values achieved at any other $\widetilde{m} \in M$, i.e., $\{p \in (0, 1) : J (p) = p\} = \{p_1\}$ and $m (\cdot, p_1)$ is the *unique* global maximum. This actually proves (1.6).

Case ii):

$$\int_\Theta \exp \left(\mu^{-1} \Delta u (\theta) \right) dP (\theta) > 1 \tag{1.34}$$

and

$$\int_\Theta \exp \left(-\mu^{-1} \Delta u (\theta) \right) dP (\theta) \leq 1 . \tag{1.35}$$

(1.34) implies $J (p) > p$ for p close enough to 0. Note that

$$\left. \frac{d^2 J}{dp^2} \right|_{p=1} = -2 \cdot \int_\Theta \left[\exp \left(-\mu^{-1} \Delta u (\theta) \right) - \exp \left(-2\mu^{-1} \Delta u (\theta) \right) \right] dP (\theta)$$

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

where the expectation is taken according to prior P. Since

$$f (x) = x^2$$

is a strictly convex function, Jensen's inequality implies

$$\mathbf{E} \exp \left(-\mu^{-1} \Delta u (\theta) \right) \geq \mathbf{E} \exp \left(-2\mu^{-1} \Delta u (\theta) \right) .$$

The inequality is not strict only if $\Delta u\left(\theta\right)$ =constant almost surely. Since

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

this constant must be non-negative. Moreover, since $\Pr\left(\Delta u\left(\theta\right)\ne 0\right)>0$, this constant must be strictly positive. Hence

$$\mathbf{E}\exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right)>\mathbf{E}\exp\left(-2\mu^{-1}\Delta u\left(\theta\right)\right)$$

and

$$\left.\frac{d^2 J}{dp^2}\right|_{p=1}<0\ . \tag{1.36}$$

Together with (1.35), (1.36) implies $J\left(p\right)>p$ for p close enough to 1. Hence there exists $\epsilon>0$, s.t. $J\left(p\right)>p$ for all $p\in\left[0,\epsilon\right]\cup\left[1-\epsilon,1\right]$.

We claim that $J\left(p\right)>p$ for all $p\in\left(0,1\right)$. If this is not true, let $p_1=\sup\left\{p\in\left(0,1\right):J\left(p\right)\le p\right\}$. The continuity of $J\left(p\right)$ implies $J\left(p_1\right)=p_1$. Hence $m\left(\cdot,p_1\right)\in M^o$ and it is a critical point of functional $V^*\left(\cdot\right)$. By the same argument as in Case i), $m\left(\cdot,p_1\right)$ is the unique global maximum. However, by definition, $p_1<1-\epsilon$ and $J\left(p\right)>p$ for all $p\in\left(p_1,1\right)$. Then $V^*\left(m\left(\cdot,p\right)\right)>V^*\left(m\left(\cdot,p_1\right)\right)$ for all $p\in\left(p_1,1\right)$ since $\frac{dV^*\left(m\left(\cdot,p\right)\right)}{dp}$ is of the same sign as $J\left(p\right)-p$. This contradicts the unique optimality of $m\left(\cdot,p_1\right)$. Therefore, $J\left(p\right)>p$ for all $p\in\left(0,1\right)$ and the optimal strategy cannot be an interior point of M (i.e., it cannot be the case $p_1\in\left(0,1\right)$.) Then according to our previous discussion, only scenarios a) that $p_1=1$ and scenario b) that $p_1=0$ are possible. Since we have shown $J\left(p\right)>p$ for all $p\in\left(0,1\right)$, we know that

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

Hence, $p_1=1$, i.e., $m\left(\theta\right)=1$ *a.s.* is the *unique* optimal strategy. This actually proves (1.4).

58

case iii):

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

and

$$\int_{\Theta} \exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right) > 1 \ .$$

In this case, by the same argument as in case ii), $m\left(\theta\right) = 0$ a.s. is the *unique* optimal strategy. This actually proves (1.5).

Now we show that it is impossible to have the case

$$\int_{\Theta} \exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right) \leq 1 \tag{1.37}$$

and

$$\int_{\Theta} \exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right) \leq 1 \ . \tag{1.38}$$

Since

$$f\left(x\right) = x^{-1}$$

is strictly convex for $x > 0$, Jensen's inequality implies

$$\int_{\Theta} \exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right) \geq \left[\int_{\Theta} \exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right)\right]^{-1} \ .$$

The inequality is not strict only if $\Delta u\left(\theta\right)$ =constant almost surely. If this is true, (1.37) and (1.38) implies $\Delta u\left(\theta\right) = 0$ almost surely. This is the trivial case excluded by our assumption. Hence

$$\int_{\Theta} \exp\left(-\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right) > \left[\int_{\Theta} \exp\left(\mu^{-1}\Delta u\left(\theta\right)\right) dP\left(\theta\right)\right]^{-1}$$

and (1.37) and (1.38) cannot be simultaneously satisfied.

Since cases i), ii) and iii) exhaust all possibilities, for each case, the corresponding

conditions are not only sufficient but also necessary.

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

This concludes the proof. ∎

Proof of Proposition 1.3.

Proof. Let $s\left(\overrightarrow{\theta}\right) = s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right)$ be an arbitrary perturbation of the optimal security $s^*\left(\cdot\right)$. Let

$$J\left(\alpha\right) = \mu \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s\left(\overrightarrow{\theta}\right)\right) + \delta_s \cdot \mathbf{E}s\left(\overrightarrow{\theta}\right) .$$

Taking first order variation leads to

$$
\begin{aligned}
\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} \\
= & -\delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \mathbf{E}\left[\exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right)\right) \cdot \varepsilon\left(\overrightarrow{\theta}\right)\right] + \delta_s \cdot \mathbf{E}\varepsilon\left(\overrightarrow{\theta}\right) \\
= & \mathbf{E}\left[\left(\delta_s - \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right)\right)\right) \cdot \varepsilon\left(\overrightarrow{\theta}\right)\right] \\
\triangleq & \mathbf{E}\left[r\left(\overrightarrow{\theta}\right) \cdot \varepsilon\left(\overrightarrow{\theta}\right)\right] . & (1.39)
\end{aligned}
$$

Let

$$A_0 = \left\{\overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) = 0\right\},$$

$$A_1 = \left\{\overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) \in \left(0, \sum_{n=1}^{N}\theta_n\right)\right\}$$

and

$$A_2 = \left\{\overrightarrow{\theta} \in \Theta : \overrightarrow{\theta} \neq \overrightarrow{0}, s^*\left(\overrightarrow{\theta}\right) = \sum_{n=1}^{N}\theta_n\right\}.$$

Then $\{A_0, A_1, A_2\}$ is a partition of $\Theta \backslash \left\{\overrightarrow{0}\right\}$. Since $s^*\left(\cdot\right)$ is the optimal security,

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} \geq 0$$

60

holds for any feasible perturbation $\varepsilon\left(\overrightarrow{\theta}\right)$. Hence, we have

$$
r\left(\overrightarrow{\theta}\right)\begin{cases} \geq 0 & \text{if } \overrightarrow{\theta} \in A_0 \\ = 0 & \text{if } \overrightarrow{\theta} \in A_1 \\ \leq 0 & \text{if } \overrightarrow{\theta} \in A_2 \end{cases} . \tag{1.40}
$$

For any $\overrightarrow{\theta}' \in A_0$, (1.40) implies $r\left(\overrightarrow{\theta}'\right) \geq 0$, i.e.,

$$
\begin{aligned}
\delta_s &\geq \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) \right) \right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}'\right) \right) \\
&= \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) \right) \right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot 0 \right) \\
&= \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) \right) \right]^{-1} ,
\end{aligned}
$$

i.e.,

$$
\begin{aligned}
\ln \delta_s &\geq \ln \delta_b - \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) \right) \\
&= \ln \delta_b + \mu^{-1} q^* ,
\end{aligned}
$$

where the last equality comes from (1.14). Hence,

$$
\mu^{-1} q^* \leq \ln \delta_s - \ln \delta_b < 0 ,
$$

which is a contradiction. Therefore,

$$
\Pr\left(A_0\right) = 0 . \tag{1.41}
$$

For any $\overrightarrow{\theta}' \in A_1$, (1.40) implies $r\left(\overrightarrow{\theta}'\right) = 0$, i.e.,

$$
\delta_s = \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}\right) \right) \right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot s^*\left(\overrightarrow{\theta}'\right) \right) ,
$$

61

i.e.,

$$
\begin{aligned}
\ln \delta_s &= \ln \delta_b - \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}\right)\right) - \mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}'\right) \\
&= \ln \delta_b + \mu^{-1}q^* - \mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}'\right) ,
\end{aligned}
$$

where the last equality follows (1.14). Therefore,

$$
s^* \left(\overrightarrow{\theta}'\right) = \mu\delta_b^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_b^{-1}q^* \tag{1.42}
$$

is a constant for all $\overrightarrow{\theta}' \in A_1$.

For any $\overrightarrow{\theta}' \in A_2$, (1.40) implies $r\left(\overrightarrow{\theta}'\right) \leq 0$, i.e.,

$$
\begin{aligned}
\delta_s &\leq \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}\right)\right)\right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}'\right)\right) \\
&= \delta_b \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}\right)\right)\right]^{-1} \exp\left(-\mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta'_n\right) ,
\end{aligned}
$$

i.e.,

$$
\begin{aligned}
\ln \delta_s &\leq \ln \delta_b - \ln \mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot s^* \left(\overrightarrow{\theta}\right)\right) - \mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta'_n \\
&= \ln \delta_b + \mu^{-1}q^* - \mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta'_n ,
\end{aligned}
$$

where the last equality comes from (1.14). Therefore,

$$
\sum_{n=1}^{N} \theta'_n \leq \mu\delta_b^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_b^{-1}q^*. \tag{1.43}
$$

Let

$$
D^* = \mu\delta_b^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_b^{-1}q^*.
$$

Then, (1.41), (1.42) and (1.43) imply that

$$s^* \left(\overrightarrow{\theta} \right) = \min \left(\sum_{n=1}^{N} \theta_n, D^* \right) \ ,$$

i.e., the optimal security must be a securitized debt.

Finally, let

$$h\left(q \right) = -\mu \ln \mathbf{E} \exp \left(-\mu^{-1} \delta_b \cdot \min \left(\sum_{n=1}^{N} \theta_n, \mu \delta_b^{-1} \cdot \left[\ln \delta_b - \ln \delta_s \right] + \delta_b^{-1} q \right) \right) \ .$$

We show that $q^* > 0$ and it is the unique fixed point of $h\left(q \right)$.

By (1.14), we have

$$
\begin{aligned}
q^* &= -\mu \ln \mathbf{E} \exp \left(-\mu^{-1} \delta_b \cdot s^* \left(\overrightarrow{\theta} \right) \right) \\
&= -\mu \ln \mathbf{E} \exp \left(-\mu^{-1} \delta_b \cdot \min \left(\sum_{n=1}^{N} \theta_n, D^* \right) \right) \\
&= -\mu \ln \mathbf{E} \exp \left(-\mu^{-1} \delta_b \cdot \min \left(\sum_{n=1}^{N} \theta_n, \mu \delta_b^{-1} \cdot \left[\ln \delta_b - \ln \delta_s \right] + \delta_b^{-1} q^* \right) \right) \\
&= h\left(q^* \right) \ .
\end{aligned}
$$

Hence q^* is a fixed point of $h(q)$. First note $h(0) > 0$. Second note that

$$
\begin{aligned}
& h'(q) \\
& = \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right) \right) \right]^{-1} \\
& \quad \cdot \mathbf{E} \left[\begin{array}{c} \exp\left(-\mu^{-1}\delta_b \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right) \right) \\ \cdot \mathbf{1}_{\left\{ \sum_{n=1}^{N} \theta_n \geq \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right\}} \end{array} \right] \\
& \leq \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right) \right) \right]^{-1} \\
& \quad \cdot \mathbf{E} \left[\exp\left(-\mu^{-1}\delta_b \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right) \right) \cdot 1 \right] \\
& = 1
\end{aligned}
$$

and

$$
\begin{aligned}
& \lim_{q\to\infty} h'(q) \\
& = \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \right]^{-1} \cdot \mathbf{E} \left[\begin{array}{c} \exp\left(-\mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \\ \cdot \lim_{q\to\infty} \mathbf{1}_{\left\{ \sum_{n=1}^{N} \theta_n \geq \mu\delta_b^{-1} \cdot [\ln\delta_b - \ln\delta_s] + \delta_b^{-1}q \right\}} \end{array} \right] \\
& = \left[\mathbf{E} \exp\left(-\mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \right]^{-1} \cdot \mathbf{E} \left[\exp\left(-\mu^{-1}\delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \cdot 0 \right] \\
& = 0 .
\end{aligned}
$$

Hence, $h(q)$ has a unique fixed point $q^* > 0$. This concludes the proof. ∎

Proof of Lemma 1.1.

Proof. Taking derivative with respect to α at $\alpha = 0$ for both sides of (1.18) leads

to

$$\mu^{-1}\delta_b \cdot \varepsilon\left(\overrightarrow{\theta}\right) = g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) \cdot \left.\frac{dm_{s,q^*}\left(\overrightarrow{\theta}\right)}{d\alpha}\right|_{\alpha=0}$$

$$-g''\left(\overline{p}_b\right) \cdot \int_{\Theta} \left.\frac{dm_{s,q^*}\left(\overrightarrow{\theta}\right)}{d\alpha}\right|_{\alpha=0} dP\left(\overrightarrow{\theta}\right) \ .$$

Take integral of both sides and manipulate we get

$$\int_{\Theta} \left.\frac{dm_{s,q^*}\left(\overrightarrow{\theta}\right)}{d\alpha}\right|_{\alpha=0} dP\left(\overrightarrow{\theta}\right)$$

$$= \mu^{-1}\delta_b \left[1 - \int_{\Theta} \left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} dP\left(\overrightarrow{\theta}\right) \cdot g''\left(\overline{p}_{s^*,q^*}\right)\right]^{-1}$$

$$\cdot \int_{\Theta} \left[g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \varepsilon\left(\overrightarrow{\theta}\right) dP\left(\overrightarrow{\theta}\right) \ .$$

Combining the above two equations leads to (1.20). ∎

Proof of Proposition 1.4.

Proof. We first prove $f_1(0) > f_2(0)$. If not, $f_1(s) < f_2(s)$ for all $s > 0$. Hence $\forall \overrightarrow{\theta} \neq \overrightarrow{0}$,

$$r\left(\overrightarrow{\theta}\right) \cdot g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right)$$

$$= -\delta_s m_{s^*,q^*}\left(\overrightarrow{\theta}\right) g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right)$$

$$= -\delta_s \cdot f_2\left(s^*\left(\overrightarrow{\theta}\right)\right) g''\left(f_2\left(s^*\left(\overrightarrow{\theta}\right)\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right)$$

$$< -\delta_s \cdot f_1\left(s^*\left(\overrightarrow{\theta}\right)\right) g''\left(f_1\left(s^*\left(\overrightarrow{\theta}\right)\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right)$$

$$= 0 \ ,$$

where the inequality holds since $[m \cdot g''(m)]' > 0$. Then (1.25) implies $s^*\left(\overrightarrow{\theta}\right) = 0$ almost surely. Therefore, there is no trade, which contradicts Proposition 1.2.

Now we know $f_1(0) > f_2(0)$. $\forall \overrightarrow{\theta} \in A_0$,

$$
\begin{aligned}
r\left(\overrightarrow{\theta}\right) &\cdot g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) \\
&= -\delta_s m_{s^*,q^*}\left(\overrightarrow{\theta}\right) g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right) \\
&= -\delta_s \cdot f_2(0)\, g''\left(f_2(0)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot 0 + w\right) \\
&> -\delta_s \cdot f_1(0)\, g''\left(f_1(0)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot 0 + w\right) \\
&= 0\,,
\end{aligned}
$$

where the second equality follows the definition that $s^*\left(\overrightarrow{\theta}\right) = 0$ for $\overrightarrow{\theta} \in A_0$, the last equality comes from the definition of $f_1(s)$, and the inequality holds since $[m \cdot g''(m)]' > 0$. This result contradicts (1.25), which states $r\left(\overrightarrow{\theta}\right) \cdot g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) \leq 0$ for $\overrightarrow{\theta} \in A_0$. This concludes the proof. $\blacksquare$

Proof of Proposition 1.5.

Proof. Let $(\bar{s}, \bar{m})$ be the unique intersection of $f_1(s)$ and $f_2(s)$. $\forall \overrightarrow{\theta}$ such that $\sum_{n=1}^{N} \theta_n < \bar{s}$,

$$
m_{s^*,q^*}\left(\overrightarrow{\theta}\right) = f_2\left(s^*\left(\overrightarrow{\theta}\right)\right) < f_2(\bar{s}) = f_1(\bar{s}) < f_1\left(s^*\left(\overrightarrow{\theta}\right)\right)\,.
$$

Then

$$
\begin{aligned}
r\left(\overrightarrow{\theta}\right) &\cdot g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) \\
&= -\delta_s m_{s^*,q^*}\left(\overrightarrow{\theta}\right) g''\left(m_{s^*,q^*}\left(\overrightarrow{\theta}\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right) \\
&> -\delta_s \cdot f_1\left(s^*\left(\overrightarrow{\theta}\right)\right) \cdot g''\left(f_1\left(s^*\left(\overrightarrow{\theta}\right)\right)\right) + \mu^{-1}\delta_b\left(q^* - \delta_s \cdot s^*\left(\overrightarrow{\theta}\right) + w\right) \\
&= 0,
\end{aligned}
$$

where the inequality holds since $[m \cdot g''(m)]' > 0$. According to (1.25), $s^*\left(\overrightarrow{\theta}\right) = \sum_{n=1}^{N} \theta_n$ for all $\overrightarrow{\theta}$ such that $\sum_{n=1}^{N} \theta_n < \bar{s}$.

For any $\overrightarrow{\theta}$ such that $\sum_{n=1}^{N} \theta_n > \overline{s}$, if $s^* \left(\overrightarrow{\theta} \right) = \sum_{n=1}^{N} \theta_n$, then (1.25) implies

$$
\begin{aligned}
0 \ \le \ & r \left(\overrightarrow{\theta} \right) \cdot g'' \left(m_{s^*, q^*} \left(\overrightarrow{\theta} \right) \right) \\
= \ & -\delta_s m_{s^*, q^*} \left(\overrightarrow{\theta} \right) g'' \left(m_{s^*, q^*} \left(\overrightarrow{\theta} \right) \right) + \mu^{-1} \delta_b \left(q^* - \delta_s \cdot s^* \left(\overrightarrow{\theta} \right) + w \right) \\
= \ & -\delta_s \cdot f_2 \left(s^* \left(\overrightarrow{\theta} \right) \right) \cdot g'' \left(f_2 \left(s^* \left(\overrightarrow{\theta} \right) \right) \right) + \mu^{-1} \delta_b \left(q^* - \delta_s \cdot s^* \left(\overrightarrow{\theta} \right) + w \right) \\
< \ & -\delta_s \cdot f_2 \left(\overline{s} \right) \cdot g'' \left(f_2 \left(\overline{s} \right) \right) + \mu^{-1} \delta_b \left(q^* - \delta_s \cdot s^* \left(\overrightarrow{\theta} \right) + w \right) \\
< \ & -\delta_s \cdot f_1 \left(s^* \left(\overrightarrow{\theta} \right) \right) \cdot g'' \left(f_1 \left(s^* \left(\overrightarrow{\theta} \right) \right) \right) + \mu^{-1} \delta_b \left(q^* - \delta_s \cdot s^* \left(\overrightarrow{\theta} \right) + w \right) \\
= \ & 0 \, ,
\end{aligned}
$$

which is a contradiction. Hence Proposition 1.4 implies $s^* \left(\overrightarrow{\theta} \right) = \overline{s}$ for all $\overrightarrow{\theta}$ such that $\sum_{n=1}^{N} \theta_n > \overline{s}$.

For any $\overrightarrow{\theta}$ such that $\sum_{n=1}^{N} \theta_n = \overline{s}$, $s^* \left(\overrightarrow{\theta} \right) = \overline{s}$ is a direct implication of Proposition 1.4.

Therefore, the optimal security is a securitized debt with face value $\overline{s}$, i.e., $s^* \left(\overrightarrow{\theta} \right) = \min \left(\sum_{n=1}^{N} \theta_n, \overline{s} \right)$.

It is also possible that $\overline{s} = \infty$, i.e., $f_1(s)$ and $f_2(s)$ never intersects. Then the optimal security

$$
s^* \left(\overrightarrow{\theta} \right) = \min \left(\sum_{n=1}^{N} \theta_n, \infty \right) = \sum_{n=1}^{N} \theta_n
$$

is a special securitized debt, i.e., equity. This concludes the proof. ∎

Proof of Proposition 1.6.

Proof. Let $\beta \in \left(\delta_s \delta_b^{-1}, 1 \right)$ and

$$
f(q) = \delta_b \cdot \mathbf{E} \min \left(\sum_{n=1}^{N} \theta_n, \beta \delta_s^{-1} q \right) \, ,
$$

where the expectation is taken according to common prior P. Since P is a continuous

distribution and $\beta^{-1}\delta_s\delta_b^{-1} < 1$, there exists $q_0 > 0$ s.t.

$$\Pr\left(\sum_{n=1}^{N}\theta_n \geq \beta\delta_s^{-1}q\right) > \beta^{-1}\delta_s\delta_b^{-1}$$

for all $q \in [0, q_0]$. Hence for any $q \in (0, q_0)$,

$$
\begin{aligned}
f'(q) &= \beta\delta_b\delta_s^{-1}\int_{\left\{\overrightarrow{\theta}\in\Theta:\sum_{n=1}^{N}\theta_n\geq\beta\delta_s^{-1}q\right\}}1\cdot dP\left(\overrightarrow{\theta}\right) \\
&= \beta\delta_b\delta_s^{-1}\cdot\Pr\left(\sum_{n=1}^{N}\theta_n\geq\beta\delta_s^{-1}q\right) \\
&> \beta\delta_b\delta_s^{-1}\cdot\beta^{-1}\delta_s\delta_b^{-1} = 1.
\end{aligned}
$$

Note that

$$f(0) = 0 ,$$

which implies that

$$f(q) > q$$

for all $q \in (0, q_0)$.

Consider a securitized debt

$$s\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N}\theta_n, D\right)$$

with face value $D = \beta\delta_s^{-1}q$ and price $q \in (0, q_0)$. Since the seller's payoff gain from

accepting this offer over rejecting it is

$$q - \delta_s s\left(\overrightarrow{\theta}\right)$$

$$= q - \delta_s \min\left(\sum_{n=1}^{N}\theta_n, \beta\delta_s^{-1}q\right)$$

$$\geq q - \delta_s \cdot \beta\delta_s^{-1}q$$

$$= (1 - \beta) \cdot q > 0$$

for all $\overrightarrow{\theta} \in \Theta$, the seller will accept this offer without acquiring any information. Hence the buyer's expected payoff from proposing $(s\left(\cdot\right), q)$ is

$$W\left(s, q\right) = \delta_b \cdot \mathbf{E}\min\left(\sum_{n=1}^{N}\theta_n, \beta\delta_s^{-1}q\right) - q$$

$$= f\left(q\right) - q$$

$$> 0.$$

By definition, the seller's expected payoff through the optimal contract is $W\left(s^*, q^*\right) \geq W\left(s, q\right) > 0$. This directly implies $\overline{p}_{s^*,q^*} > 0$ since $\overline{p}_{s^*,q^*} = 0$ always generates zero expected payoff to the buyer. ∎

Chapter 2

Coordination with Flexible

Information Acquisition [1]

2.1 Introduction

This chapter studies a coordination game where players can flexibly acquire infor-
mation. Coordination requires common knowledge among the players of the payoffs.
In practice, players may not be able to observe payoffs perfectly and so lack such
common knowledge. This imperfection could result from the complex nature of
the environment. The global game literature attempts to model this issue through
allowing players to receive private signals of payoffs while fixing the information
structure. Accordingly researchers obtain a well-known limit unique equilibrium
which features inefficient coordination (e.g., Carlsson and Damme (1993), Frankel,
Morris, and Pauzner (2003)). This chapter explores another model of the lack of
common knowledge. Rather than equipping the players with exogenous signals, we
allow them to acquire information about the payoffs through choosing information

---
[1]This paper was presented at North American Winter Meeting of the Economet-
ric Society, Denver, January 2011.

structures. On the one hand, an information structure exclusively specifies both the amount of information, which is measured by reduction of Shannon's entropy, and the substance of which the information is relevant. Hence, information acquisition is flexible in the sense that the players choose not only the quantity but also the qualitative nature of their information. On the other hand, in our model, information acquisition also incurs a cost proportional to the amount of information acquired. As a consequence, players will focus on the information most relevant to their welfare and be rationally inattentive to other aspects of the fundamental. This chapter addresses the following questions centered on flexible information acquisition: what is the information acquired under this flexibility, and how does this flexibility affect welfare in terms of efficiency of coordination? We also go beyond the entropic information cost to highlight the key aspects of this flexibility and how they drive our results.

Throughout the chapter, we use the following story to illustrate our idea. Two players coordinate in investing in a risky project. The project's future cash flow is driven by a randomly fluctuating fundamental. Each player must decide whether to invest or not, and her payoff depends on her opponent's action as well as on the realized fundamental. Given her opponent's action, a player's payoff from investing increases in the fundamental. Moreover, given any realization of the fundamental, a player's gain from 'invest' over 'not invest' is strictly higher when her opponent also invests. Therefore, players' actions are strategic complements. Before making a decision, each player can independently purchase private information about the fundamental in the form of an information structure, i.e., the conditional distribution of her signal given the fundamental. She then takes action according to her realized signal.

In our benchmark model, the players' flexible information acquisition is modeled through the framework of rational inattention building on Sims (2003). The basic

idea of rational inattention is that people face informational capacity constraints defined by Shannon's information theory. That is, there are limited bits that can be used to reduce the uncertainty of some exogenous variables. We deviate from the standard setup of rational inattention by replacing the capacity constraint with an information cost, while maintaining the players' flexibility in choosing information structures. As a result, players can and will collect information most relevant to their welfare but be rationally inattentive to other aspects of the fundamental.

Specifically, there are two factors affecting players' information acquisition strategy. The first one is the fundamental effect. Intuitively, given her opponent's action, a player is willing to collect information about the fundamental that helps her invest with high (low) probability in the high (low) states. The second but more interesting factor is the players' consideration of their opponents' information acquisition. Since my payoff depends on your actions and your actions are motivated by your beliefs, upon receiving my signal, I attempt to infer not only the fundamental, but also your signal, your beliefs about my signal, your beliefs concerning my beliefs induced by my signal, and so on. This reasoning never stops and applies to both players, thus makes higher-order beliefs of all orders relevant for both players' decisions. Moreover, considering that these higher-order beliefs are ultimately determined by the information acquired, both players have an incentive to match each other's informational choice in order to minimize the probability of miscoordination. As a result, the strategic complementarity between actions induces coordination motive in information acquisition. Indeed, this motive can evolve to coordination in information acquisition since information acquisition is flexible, especially when information cost is low. Hence, lowering information cost makes efficient coordination in investing possible through efficient coordination in acquiring information, while also gives rise to multiple equilibria due to multiple ways of coordinating information acquisition.

In order to highlight the indispensable mechanism of flexibility, it is instructive

to contrast the role of information acquisition in our framework from that in an extended global game model, where players are allowed to endogenously choose the precision of their information. In this extended model, player i observes her signal $x_i = \theta + \beta_i^{-1/2} \cdot \varepsilon_i$, where θ is the fundamental, ε_i is a noise with density f and β_i is the precision of her signal. Player i can purchase more accurate signals by increasing precision β_i at some cost. Intuitively, as such cost decreases, players acquire signals with higher precision. Due to the private nature of such signals, higher precision weakens players' approximate common knowledge of payoffs. Hence, in the limit we attain the inefficient and unique equilibrium commonly seen in standard global game models. The reason why efficient coordination fails and multiplicity disappears in this extended global game model lies in its rigidity of information acquisition. This rigidity is represented by the restriction that players can only pay equal attention to every possible realization of the fundamental. That is, the distribution of player i's observational error $\beta_i^{-1/2} \cdot \varepsilon_i$ is invariant in θ. As a result, players only coordinate in choosing the overall precision but cannot coordinate their attention allocation for different levels of θ. This mechanism of rigidity sharply contrasts with its counterpart when information acquisition is flexible as discussed above.

An important implication of our approach is that flexible information acquisition, together with the fact that information is costly, helps players in strategic circumstances only acquire information valuable for efficient coordination and refrain themselves from information harmful to coordination even if information cost goes to zero, and thus become strictly better off than in the case of rigid information acquisition; while an agent in a single person decision problem cannot enjoy such benefit by doing so. In the extended global game model, an efficient outcome can not be sustained in any equilibrium, but if players could throw away information with an explicit commitment device, efficiency could be retrieved in equilibrium. In contrast, in our framework with flexible information acquisition, since the feasible

information structure to be acquired is flexible and information is costly, players can choose the quantitative and qualitative nature of information. Thus they act as if they commit to throwing away information harmful to coordination. Hence, efficient coordination could be supported in equilibrium. In other words, flexibility of information acquisition creates an implicit commitment device. It should be highlighted that this contrast could not be seen in single person decision problems with information acquisition. This is not only because, literally, single person decision problems do not involve coordination, but also due to the fact that more information is always more desirable disregarding its cost, no matter whether information acquisition is flexible or rigid.

To explore the essence of flexibility, it is important to first focus on the fact that different forms of information acquisition could be exclusively captured by different schemes of information cost. If information structure is endowed exogenously, it could be viewed that only this endowed information structure can be acquired at finite cost while all other information structures incur an infinite cost. If information acquisition is endogenous but rigid as in the extended global game model described above, only information structures following the form of $x = \theta + \beta^{-1/2} \cdot \varepsilon$ are associated with finite cost. In our benchmark model where information acquisition follows rational inattention, any information structure is associated with a cost proportional to the resulted reduction of entropy. In all, analysis on schemes of information cost covers any consideration on the forms of information acquisition. It is worth noting that information cost given by rational inattention respects Blackwell's ordering[2], due to which it suffices for players to only consider binary information structures; and further features uniform boundedness for those binary information structures,

---

[2]An information cost respects Blackwell's ordering if it assigns lower cost to less informative information structures. An information structure is less informative than the other if it can be obtained from the other by adding garbling noise in the sense of Blackwell (1953).

which ensures the availability of any such essential information. These two aspects will help us understand the essence of flexibility.

Having abstracted the two aspects from rational inattention, we next go beyond the entropic information cost to define the flexibility in our context, and identify its key aspect driving our results. Information acquisition is flexible if the information cost respects Blackwell's ordering and features uniform boundedness for the binary information structures. Blackwell's ordering precludes the information structures that contain information of no potential value. The uniform boundedness specifies a common upper bound on the costs of all binary information structures. It guarantees the availability of all potentially valuable information structures. We show that our qualitative results remain valid for any information cost satisfying the flexibility. On the other hand, information acquisition is rigid if it is not flexible. As typical examples with rigid information acquisition, standard global game models and our extended global game model satisfy neither of the two conditions. More importantly, the second aspect plays the main role in shaping our results. Suppose the first condition holds and thus we can focus on binary information structures, then for almost any such information structure, we can make it supported in limit unique equilibrium by assigning infinite cost to preclude some binary information structures. Therefore, the rigidity resulted from the violation of the second aspect leads to the failure of our qualitative results. In this sense, we can interpret the second aspect in a way that the selection of some specific equilibrium could be done by designing its corresponding rigidity.

This chapter proceed as following. Section 2.2 sets up the model and prepares some simple facts about information acquisition behavior in equilibrium. In Section 2.3, we characterize the equilibria and gain some intuition through comparative static analysis. Section 2.4 conveys our main results. It first compares our approach to an extended global game model and explores the origin of the difference. This

75

section then compares the welfare implications of flexible information acquisition in strategic settings and non-strategic settings. It also extends the concept of flexibility and discuss its essence. Section 2.5 examines effects of public information through a comparative static analysis with respect to the common prior. We conclude in Section 2.6 by discussing several extensions of the benchmark model. Most proofs are relegated to the appendix.

2.1.1   Relation to Literature

Throughout our benchmark model, rational inattention illustrates the flexibility of players' information acquisition. Sims (1998) pioneers rational inattention to model price stickiness, where the capacity constraints dampen and delay people's responses to shocks. Sims (2003) and Sims (2005) further develop the theory to accommodate dynamic programming in both linear-quadratic and non-linear-quadratic cases. Later Matejka (2010) shows that a perfectly attentive seller sets discrete and rigid prices to stimulate a rationally inattentive buyer to consume more. Recently, Mackowiak and Wiederholt (2009) examine rational inattention in a dynamic stochastic general equilibrium model. They show that prices respond strongly and quickly to idiosyncratic shocks but weakly and slowly to aggregate shocks.[3]

In applied work, rational inattention is mainly studied in two cases: the linear-quadratic case (e.g., Mackowiak and Wiederholt (2009)), and the binary-action case. A leading example of the latter is Woodford (2009), where firms acquire information and then decide whether to review their prices. Compared to the standard "Ss" model, the data of individual price changes are better explained through a large information cost. Our model also adopts the binary-action setup.

---

[3]To learn more about rational inattention, see Sims (2006), Sims (2010), Luo (2008), Mackowiak and Wiederholt (2011), Van Nieuwerburgh and Veldkamp (2009a), Van Nieuwerburgh and Veldkamp (2009b), Paciello (2009), Mondria (2010), Matejka and Sims (2011).

The literature of global games has contributed a great deal to our understanding of coordination under incomplete information, but our findings vary due to different approaches. As introduced by Carlsson and Damme (1993), the approach of global games is a natural refinement to remove multiplicity in two-player, two-action games with common knowledge of payoffs. In a global game model, players are not sure about the payoffs but can make inference according to their privately observed signals. This uncertainty about payoffs and other players' beliefs weakens approximate common knowledge among players, and thus facilitates uniqueness. For example, when each player observes a private signal equal to the fundamental plus independent noise, a unique strategy profile survives iterative dominance as the signal noise vanishes Frankel, Morris, and Pauzner (2003). Regaining the predictive power through removing the multiplicity, global game models are widely applied in the study of currency attacks Morris and Shin (1998), debt pricing Morris and Shin (2004) and bank runs Goldstein and Pauzner (2005), etc. All these models are characterized by their exogenous information structures. Researchers assume the signals to equal the fundamental plus some independent noise and so players only passively respond to such signals. Researchers also let the noise vanish while fixing the additive structure to get the limit unique equilibrium with inefficient outcome. In other words, exogeneity and rigidity are the key features of global game analysis. As will be shown, it is this rigidity that leads to the limit uniqueness result.

Another related strand of work is characterized by endogenous but rigid information acquisition, for example, Hellwig and Veldkamp (2009) and (Myatt and Wallace 2009). In Hellwig and Veldkamp (2009), a continuum of agents purchase from a set of signals and then play a coordination game. In this model, information acquisition is endogenous in the sense that players decide which signals to buy according to their own interests. But information acquisition is still rigid, since signals and fundamentals are assumed to be jointly Gaussian, which renders the signals equal to (the linear

combinations of) fundamentals plus Gaussian noise. Both Hellwig and Veldkamp (2009) and our model show that if players' actions are strategic complements, so are their information choices. However, our research provides different insight about the creation of approximate common knowledge and multiplicity. While in Hellwig and Veldkamp (2009), public noise is assumed in the signals and players have the option to access such public noise to create approximate common knowledge, in our approach, the creation of approximate common knowledge comes from the flexibility of information acquisition. Nothing in the model pushes players towards common knowledge but it is just their incentive to do so. As a result, while public noise is necessary to generate multiplicity in Hellwig and Veldkamp (2009), our model admits multiple equilibria once strategic complementarity exceeds information cost, no matter whether there is public information. More importantly, players in Hellwig and Veldkamp (2009) could potentially benefit from committing to throwing away some information, but they are not able to do so due to the setup of rigid information acquisition; while flexibility in our model ensures that players never acquire such redundant and even harmful information and thus may retrieve efficient coordination. Myatt and Wallace (2011) consider another setting in which there are also multiple parametric signals available, and each signal contains a public noise component and a private noise component. Agents can get access to all signals, but have to choose the precision for the private component of each signal, subject to a constraint on the total precision of the private components of all signals. By such a setting, Myatt and Wallace (2011) offer a special interpretation of endogenous information acquisition on public signal in coordination problems, which preserve the equilibrium uniqueness in the sense that acquiring more accurate public information is actually conducted by acquiring more clear private information, and thus does not lead to the kink in marginal benefit of public information as that in Hellwig and Veldkamp (2009). Similar to the contrast between our framework and Hellwig

78

Table 2.1: payoff matrix conditional on fundamental

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

.

and Veldkamp (2009), the information acquisition in Myatt and Wallace (2011) is rigid with respect to my model, and their focus and underlying mechanisms are also different.

2.2   The Model

2.2.1   The Basic Environment

We define our game as follows. Two players[4] play a coordination game with payoffs shown by Table 2.2.1.

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

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

Note that $r > 0$ is the cost of miscoordination. It measures the degree of strategic complementarity[5]. Fundamental θ is drawn from Θ according to a common prior P,

---

[4]The "two-player" setup is not as restrictive as it seems. All our results remain valid when there is a continuum of players if we redefine the payoff for "invest" as $\theta - r \cdot (1 - m)$, where m is the fraction of the players that invest. We also discuss the "n-player" case in Subsection 2.6.3.

[5]We assume constant strategic complementarity for the sake of simplicity. The more general case with a state-dependent strategic complementarity $r(\theta) > 0$ almost surely is discussed in Subsection 2.6.4.

which is a probability measure over Θ. We assume that P is absolutely continuous with respect to Lebesgue measure over $\mathbb{R}$.[6]

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

To gain some intuition, note that when information is complete and $\theta \in [0, r]$, there exist two strict Nash equilibria: (invest, invest) and (not invest, not invest). This multiplicity results from excessive coordination built upon common knowledge of the fundamental.

Our model focuses on the case of incomplete information. Suppose player i is equipped with an information structure (S_i, q_i) that conveys information about θ. Here $S_i \subset \mathbb{R}$ is the set of realizations of player i's signal, and $q_i(s_i|\theta)$ is the conditional probability density function of her signal. We assume conditional independence between s_i and s_j given θ to represent the private nature of the players' information. Given information structure (S_i, q_i), player i's strategy can be represented by a mapping σ_i from S_i to $[0, 1]$, where $\sigma_i(s_i)$ denotes the probability of choosing 1 upon observing $s_i \in S_i$. Player i's expected payoff given (s_i, s_j, θ) becomes

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

[6]Note that public information affects the common prior, thus its effects on the equilibria can be studied through comparative static analysis with respect to the common prior, as shown in Section 2.5.

Then we can define

$$U_i \left(\left((S_i, q_i), \sigma_i \right), \left((S_j, q_j), \sigma_j \right) \right)$$
$$= \int_\theta \int_{s_i} \int_{s_j} \sigma_i (s_i) \left[\theta - r \cdot (1 - \sigma_j (s_j)) \right] q_i (s_i|\theta) q_j (s_j|\theta) \, ds_i ds_j dP (\theta), \quad (2.1)$$

as player i's expected payoff with strategy profile (σ_i, σ_j) under information struc-

ture $((S_i, q_i), (S_j, q_j))$, $i, j \in \{1, 2\}$, $i \neq j$.

Now we consider a larger game with flexible information choices: player i chooses

strategy $((S_i, q_i), \sigma_i)$ according to the preference given by (2.1), $i \in \{1, 2\}$. More pre-

cisely, player i acquires information through choosing information structure (S_i, q_i),

and then takes (mixed) action σ_i according to her signal s_i generated by (S_i, q_i).

The solution concept is Nash equilibrium. It is worth noting that the strategy profile

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

is common knowledge in equilibrium, but the players' beliefs about the fundamental

and others' actions are heterogeneous due to the private nature of their signals.

The conditional density $q_i (s_i|\theta)$ describes player i's information acquisition strat-

egy. By choosing different functional forms for $q_i (s_i|\theta)$, player i can make her signal

covary with the fundamental in any way she would like. Intuitively, if player i's

welfare is sensitive to the fluctuation of the fundamental within some range $A \subset \Theta$,

she would pay much attention to this event by making her signal s_i highly correlated

with $\theta \in A$. In this sense, choosing an information structure can be interpreted as

hiring an analyst to write a report with emphasis on your interests.

If information acquisition incurs no cost, player i would like to establish a one-to-

one mapping between s_i and θ, and thus obtain all information of the fundamental.

This makes our problem a trivial one since it is just a coordination game with

complete information. In practice, however, information acquisition is unlikely to be free. We study the more interesting case of costly information acquisition. To do so, we associate each conditional density $q_i(s_i|\theta)$ with a cost $\mu \cdot I(q_i)$, where

$$
\begin{aligned}
I(q_i) \ = \ & \int_\theta \int_{s_i} q_i(s_i|\theta) \ln q_i(s_i|\theta) \cdot ds_i \cdot dP(\theta) \\
& - \int_{s_i} \int_\theta q_i(s_i|\theta) \cdot dP(\theta) \cdot \ln\left(\int_\theta q_i(s_i|\theta) \cdot dP(\theta)\right) \cdot ds_i \quad (2.2)
\end{aligned}
$$

is the mutual information between the two random variables s_i and θ. It measures the amount[7] of information about θ conveyed by s_i.[8] The marginal cost of information acquisition is $\mu > 0$. It reflects the difficulty in acquiring information. By definition, $I(q_i)$ is uniquely determined by the functional form of $q_i(s_i|\theta)$. Indeed, a functional form of $q_i(s_i|\theta)$ defines a specific way of information acquisition. It determines the qualitative nature of information to be acquired. Different forms of $q_i(s_i|\theta)$ may generate the same value for $I(q_i)$, i.e., the same amount of information may be collected from different aspects of θ. Since information acquisition is costly, it is not wise for player i to have a signal very informative of all values of θ. She should make it sensitive to the events most relevant to her welfare.

Taking into account of information cost, player i's payoff through taking strategy $((S_i, q_i), \sigma_i)$ is her expected future cash flow minus her cost of information

---

[7]Here the unit of $I(q_i)$ is "nat". If "ln" is replaced by "$\log_2$", the unit becomes "bit". Knowing the result of a single toss of a fair coin obtains 1 bit of information. Since 1 bit equals $\ln 2$ nats, choosing bit or nat as the unit does not make any difference to our analysis.

[8]Shannon's mutual information is a natural measure of information about one random variable conveyed by another one. In Shannon's information theory, information is defined as the reduction of uncertainty, which is reflected by the difference between the two terms in formula (2.2). Moreover, it is the uniquely "right" way to measure information under some intuitive axioms. Cover and Thomas (1991) provide a detailed discussion of mutual information. However, this specific entropic functional form is not necessary for our qualitative results. The essence is the flexibility in choosing information structures. Subsection 2.4.3 extends our results under more functional general forms.

acquisition, i.e.,

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

$$= \quad U_i \left(\left(\left(S_i, q_i \right), \sigma_i \right), \left(\left(S_j, q_j \right), \sigma_j \right) \right) - \mu \cdot I \left(q_i \right), i \neq j. \qquad (2.3)$$

The following definition summarizes the game:

Definition 1 *(Costly Information Acquisition Game $G \left(r, \mu \right)$): Two players with preference (2.3) play the game through choosing $\left(\left(S_i, q_i \right), \sigma_i \right)$, where $S_i \subset \mathbb{R}$ is the set of realizations of player i's signal, $q_i \left(s_i | \theta \right)$ is the conditional probability density of her signal, and σ_i is a mapping from S_i to $[0, 1]$ that defines player i's action upon receiving signal $s_i \in S_i$, $i \in \{1, 2\}$. The equilibrium concept is Nash equilibrium.*

In principle, this problem seems hard to deal with, since players' possible strategies belong to a functional space, and even S_1 and S_2, the sets of realizations of the signals, are endogenous. Fortunately, some patterns emerge from the players' optimal information acquisition behavior. They help significantly simplify our problem.

2.2.2 Some Simple Facts About The Equilibria

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

$$S_{i,I} \quad = \quad \{s_i \in S_i : \sigma_i \left(s_i \right) = 1\},$$

$$S_{i,N} \quad = \quad \{s_i \in S_i : \sigma_i \left(s_i \right) = 0\},$$

and

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

Note that $S_{i,I}$ ($S_{i,N}$) is the set of signal realizations such that player i definitely *invests* (*not invests*). Player i is indifferent when her signal belongs to $S_{i,ind}$. Then $(S_{i,I}, S_{i,N}, S_{i,ind})$ forms a partition of S_i. Since the only use of the signal is to make a binary decision, a signal differentiating more finely among the states only conveys redundant information and wastes the players' attention. Hence player i will not discern signal realizations within any of $S_{i,I}$, $S_{i,N}$ and $S_{i,ind}$. In addition, because she is indifferent between the two actions upon event $S_{i,ind}$, she would rationally pay no attention to distinguish this event from other realizations. Hence, the players always play pure strategies upon receiving their signals. Therefore, the players always prefer binary-signal information structures. Woodford (2008) has a similar argument that the agent only needs to acquire a "yes/no" signal. To maintain the completeness of this chapter, we prove the results in our context.

Lemma 2.1 *In any equilibrium of costly information acquisition game* $G(r, \mu)$, $\#S_i = 1$ *or* 2, *and* $\Pr(S_{i,ind}) = 0$, $\forall i \in \{1, 2\}$.

Proof. See Appendix 2.7. ∎

This lemma follows the fact that the information cost given by rational inattention respects Blackwell's ordering[9] of information structures. That is, more informative information structures are more expensive. Hence, the players only choose binary information structures to save their information costs.

Suppose $((S_i, q_i), \sigma_i)$ is player i's equilibrium strategy. Then it induces a conditional probability function m_i from Θ to $[0, 1]$, such that player i invests with probability $m_i(\theta)$ when the fundamental equals θ. On the other hand, Lemma 2.1 implies that m_i also suffices to characterize player i's equilibrium strategy $((S_i, q_i), \sigma_i)$. That is, we can recover $((S_i, q_i), \sigma_i)$ from m_i. Specifically, in the trivial case

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

[9]See Blackwell (1953).

let

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

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

otherwise, let

$$S_i = \{s_{i,I}, s_{i,N}\},$$

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

$$\sigma_i(s_{i,I}) = 1, \, \sigma_i(s_{i,N}) = 0.$$

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

Without any confusion, we can abuse the notation a little to rewrite player i's expected payoff as

$$U_i(m_i, m_j) = \int m_i(\theta) \cdot [\theta - r \cdot (1 - m_j(\theta))] \cdot dP(\theta), \, i, j \in \{1, 2\} \text{ and } i \neq j. \quad (2.4)$$

This expression is derived from (2.1).

Lemma 2.1 also implies that $I(q_i)$, the amount of information acquired, is a functional of m_i. Hence we use $I(m_i)$ instead of $I(q_i)$ hereafter. By (2.2) and Lemma 2.1, $I(m_i)$ has the expression

$$I(m_i) = \int [m_i(\theta) \ln m_i(\theta) + (1 - m_i(\theta)) \ln(1 - m_i(\theta))] \, dP(\theta)$$
$$- p_{Ii} \ln p_{Ii} - (1 - p_{Ii}) \ln(1 - p_{Ii}), \quad (2.5)$$

where

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

85

is player i's unconditional probability of investing. It is worth noting that mutual information $I(m)$ measures function m's variability, which reflects the informativeness of actions to the fundamental. For example, when $m(\theta)$ is constant, the actions convey no information about θ and the corresponding mutual information is zero. This is because the integrand in the first term of (2.5) is strictly convex and thus $I(m)$ is zero if and only if $m(\theta)$ is constant. Hence, a nice property of our technology of information acquisition is that there exists information acquisition if and only if $m(\theta)$ varies over θ, if and only if information cost is positive. Also note that the "shape" (functional form) of m determines not only the quantity but also the qualitative nature of information. For instance, a player can concentrate her attention to some event through making $m(\theta)$ highly sensitive to θ within such event. In this sense, our technology of information acquisition is flexible since the players can decide both the quantity and quality of their information through freely choosing m.

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

$$V_i(m_i, m_j) = U_i(m_i, m_j) - \mu \cdot I(m_i), \ i, j \in \{1, 2\} \ \text{and} \ i \neq j. \qquad (2.6)$$

For the sake of simplicity, the rest of this chapter abstracts away from the story of costly information acquisition and treats the problem as a two-player game with preference (2.6) and strategy profile (m_1, m_2). We assume that each player's strategy space is $L^1(\Theta, P)$, i.e., the space of all P-integrable functions on Θ equipped with the norm

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

2.3 The Equilibria of the Game

A Nash equilibrium of the costly information acquisition game is a strategy profile (m_1, m_2) solving the the following problem:

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

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

where $L^1(\Theta, P)$ is the space of P-integrable functions on Θ, $i, j \in \{1, 2\}$ and $i \neq j$.

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

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

As shown in Proposition 1.1, this payoff gain function determines player i's incentive to acquire information. Player i acquires information for two reasons. First, she wants to reduce her uncertainty about the fundamental. Second, she can coordinate her investment decision with her opponent's through coordinating in acquiring information. For the second reason, player i should pay attention to the events that player j pays attention to. If player j never acquires information (e.g., she always invests), however, this second reason does not exist for player i. In addition, if the fundamental is very likely to be positive ex ante, the first reason does not hold either. Then player i may find it optimal to always invest without acquiring any information. This in turn confirms player j's non-information acquisition strategy and thus constitutes an equilibrium. Because such non-information acquisition equilibria are trivial and not interesting, we will exclude them by imposing the following assumption and focus on the equilibria with information acquisition.

Assumption: $\mathbf{E} \exp(-\mu^{-1}\theta) > 1$ *and* $\mathbf{E} \exp(\mu^{-1}\theta) > e^{\mu^{-1}r}$, *where the expecta-*

tion is taken according to common prior P.[10]

This assumption is similar to the "limit dominance" assumption in the global games approach.[11] The underlying intuition is that common prior P should not concentrate within interval $[0, r]$. Otherwise, the players are commonly confident about the event $\theta \in [0, r]$. Once a player always invests (not invests), the other one finds her payoff gain very likely to be positive (negative) ex ante and thus loses her incentive to acquire information. To gain some intuition, note that if the common prior is $N(t, \sigma^2)$, this assumption is equivalent to

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

if the common prior is a uniform distribution over interval $[-A, r + A]$, then the assumption holds when $A > 0$ is large enough. All results that follow are derived under this assumption, unless otherwise noted.

Proposition 2.1 *In equilibrium, player i's strategy is characterized by*

$$\theta - r \cdot [1 - m_j(\theta)] = \mu \cdot \left[\ln\left(\frac{m_i(\theta)}{1 - m_i(\theta)}\right) - \ln\left(\frac{p_{Ii}}{1 - p_{Ii}}\right)\right] \quad \text{almost surely,} \quad (2.7)$$

where

$$p_{Ii} = \int m_i(\theta) \, dP(\theta) \in (0, 1) \tag{2.8}$$

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

Proof. Note that

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

---

[10] In Proposition 2.8 and 2.9 of Appendix 2.8, we prove that there exists a pooling equilibrium "$m_i(\theta) = m_j(\theta) = 1$ almost surely" ("$m_i(\theta) = m_j(\theta) = 0$ almost surely") if and only if $\mathbf{E}\exp\left(-\mu^{-1}\theta\right) \leq 1$ ($\mathbf{E}\exp\left(\mu^{-1}\theta\right) \leq e^{\mu^{-1}r}$), where the expectation is taken according to common prior P. Therefore, this assumption excludes all equilibria with no information acquisition.

[11] See the survey of global game models by Morris and Shin (2001).

and

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

where the last inequality follows the assumption $\mathbf{E} \exp \left(\mu^{-1} \theta \right) > e^{\mu^{-1} r}$. *Also note that*

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

where the last inequality is just the assumption. Therefore, (2.7) and (2.8) are direct implications of case c) in ii) of Proposition 1.1. ∎

This result is a direct implication of Proposition 1.1. It reads that a pair $(m_1, m_2) \in \Omega \times \Omega$ is an equilibrium if and only if it satisfies (2.7) and (2.8). Given player j's strategy $m_j \left(\theta \right)$, the left hand side of (2.7) is player i's marginal benefit of increasing her conditional probability of "invest". Since $\mu > 0$ is the marginal cost of acquiring an extra bit of information and

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

is the "derivative" of the quantity of information with respect to $m_i \left(\theta \right)$, the right hand side of (2.7) is player i's marginal cost of increasing $m_i \left(\theta \right)$. Then (2.7) reads

that the marginal cost must equal the marginal benefit. Also note that

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

is player i's average odds ratio of "invest" relative to "not invest", while

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

is her odds ratio conditional on θ. Then (2.7) indicates that player j's strategy m_j shapes player i's marginal benefit

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

which in turn determines the deviation of player i's odds ratio from its average level.

Since payoff matrix (Table 2.2.1) is symmetric and the players' actions are strategic complements, it is natural to expect symmetric equilibria.

Proposition 2.2 *All equilibria of the costly information acquisition game are symmetric, i.e.,* $\Pr(m_1(\theta) = m_2(\theta)) = 1$.

Proof. See Appendix 2.8. ■

The strategic complementarity between the players' actions gives rise to their coordination motive in acquiring information. Due to the private nature of their information acquisition, they can achieve this coordination only through choosing the same information structure (i.e., the same $m(\cdot)$). According to this proposition, we can use a single function m to represent the equilibrium hereafter.

Corollary 2.1 (m_1, m_2) *is an equilibrium of the costly information acquisition game if and only if there exists an* $m \in \Omega$, *such that* $m_i(\theta) = m(\theta)$ *almost surely for*

$i \in \{1, 2\}$, *and*

$$\forall \theta \in \Theta, \; \theta - r \cdot [1 - m(\theta)] = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1 - m(\theta)} \right) - \ln\left(\frac{p_I}{1 - p_I} \right) \right], \qquad (2.9)$$

where

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

Proof. This corollary directly follows Proposition 2.2. ∎

This corollary is a direct implication of Propositions 2.1 and 2.2. It is worth noting that equation (2.9) summarizes all the previous derivations and is sufficient and necessary for all equilibria of the game. The following two sections conduct equilibrium analysis through analyzing this equation.

It is easy to verify that the graph

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

is central-symmetric[12] in the $\theta \sim m$ plane about the point $(\theta_0, 1/2)$, where

$$\theta_0 = r/2 - \mu \cdot \ln\left(\frac{p_I}{1 - p_I} \right). \qquad (2.10)$$

Combining (2.9) and (2.10) leads to

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

Hence any solution of (2.9) has an expression $m(\theta - \theta_0)$ and can be indexed by θ_0, i.e.,

$$\theta - \theta_0 = \mu \cdot \ln\left(\frac{m(\theta - \theta_0)}{1 - m(\theta - \theta_0)} \right) + r \cdot \left(\frac{1}{2} - m(\theta - \theta_0) \right). \qquad (2.11)$$

---

[12]This symmetry comes from the fact that the strategic complementarity r does not depend on θ.

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

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

A solution of (2.9) is jointly determined by its position θ_0 and its "shape" $m(\theta)$. It is worth pointing out that not every $\theta_0 \in \mathbb{R}$ suffices to make $m(\theta - \theta_0)$ a solution of the game. The position θ_0 is endogenously determined in equilibrium.

We first analyze the "shape" of the equilibrium. The "shape" $m(\theta)$ is determined by $\widetilde{r} \triangleq \frac{r}{4\cdot\mu}$, the ratio of strategic complementarity r to marginal cost of information acquisition μ. Figure 2.1 shows how $m(\theta)$ evolves as $\widetilde{r}$ increases.

What information is acquired in equilibrium? According to Lemma 2.1, we can recover the equilibrium information structure from $m(\theta)$. Let $S_i = \{0, 1\}$ be the set of realizations for player i's signal s_i. Player i invests if $s_i = 1$ and does not invest if otherwise. This information structure is characterized by conditional probability $\Pr(s_i = 1|\theta) = m(\theta)$. Since the probability of investing is highly sensitive to θ where slope $\left|\frac{dm(\theta)}{d\theta}\right|$ is large, $\left|\frac{dm(\theta)}{d\theta}\right|$ reflects player i's attentiveness around θ[13]. Under this interpretation, Figure 2.1 reveals that players actively collect information for intermediate values of the fundamental but are rationally inattentive to values at the tails. This result coincides with our intuition. When θ is too high (low), the players should invest (not invest) anyway. Hence the information about θ on the tails is not so relevant to their payoffs. When θ takes intermediate values, the player's payoff gain from investing over not investing depends crucially on the value of θ as well as its implication of her opponent's action. Therefore, the information about θ in the intermediate region is payoff-relevant and attracts most of their attention.

How does information acquisition affect coordination? First, the equilibrium strategy becomes flatter as μ increases. Higher informational cost directly weakens

---

[13] $\left|\frac{dm(\theta)}{d\theta}\right| \triangleq \infty$ when $m(\theta)$ is discontinuous at θ.
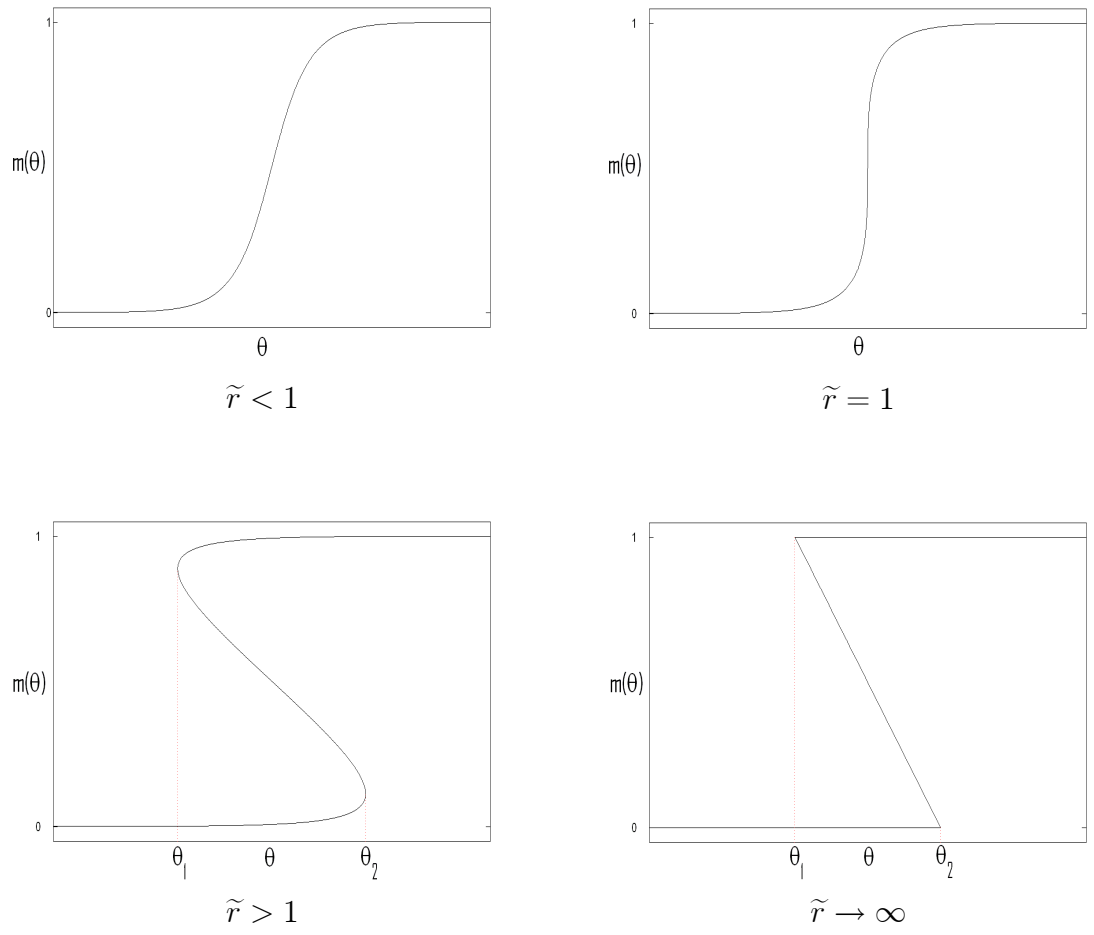
Figure 2.1: Evolution of the shape of equilibrium

players' ability to acquire information. Hence more idiosyncratic errors enter players' responses. Moreover, expecting that her opponent reacts in a noisier fashion, the player no longer has as much incentive to coordinate as before. Therefore, the equilibrium strategy becomes even less decisive.

Second, multiple equilibria might emerge[14] as $\widetilde{r} = \frac{r}{4 \cdot \mu}$ exceeds unity. As shown in the low-left subgraph of Figure 2.1, there exist $\theta_1 < \theta_2$ such that multiple values of $m(\theta)$ satisfy (2.11) for all θ within $[\theta_1, \theta_2]$. Note that while the strategic complementarity between players' actions measures their coordination motive, the marginal cost of information acquisition reflects the cost of coordination, since acquiring information is a prerequisite to coordinating investment decisions. Hence, the condition $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$ reads that when coordination motive dominates coordination cost, the players have multiple ways to coordinate their information acquisition, which leads to approximate common knowledge and thus multiplicity.

Third, Monotonic Likelihood Ratio Property (MLRP), an assumption often made in applied models with incomplete information, could be violated by our rationally inattentive players when $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$.[15] When coordination motive exceeds coordination cost, a player has enough incentive and ability to coordinate with her opponent's weird non-MLRP strategy. Therefore, our approach provides a condition to assess the fitness of MLRP.

Finally, as shown in the low-right subgraph of Figure 2.1, the equilibrium approximates the switching strategy when information cost vanishes. This result coincides with the equilibria of coordination games with complete information.

When $\widetilde{r} = \frac{r}{4 \cdot \mu} \leq 1$, there is a unique shape of $m(\theta)$ to satisfy (2.11). However, infinitely many shapes of $m(\theta)$ satisfy (2.11) when $\widetilde{r}$ exceeds unity. Figure 2.2 shows

---

[14]We prove this multiplicity later.

[15]We say a player's strategy satisfies MLRP if her conditional probability of investing increases in the fundamental. In other words, the information structure is more likely to suggest the players to invest for higher fundamental.
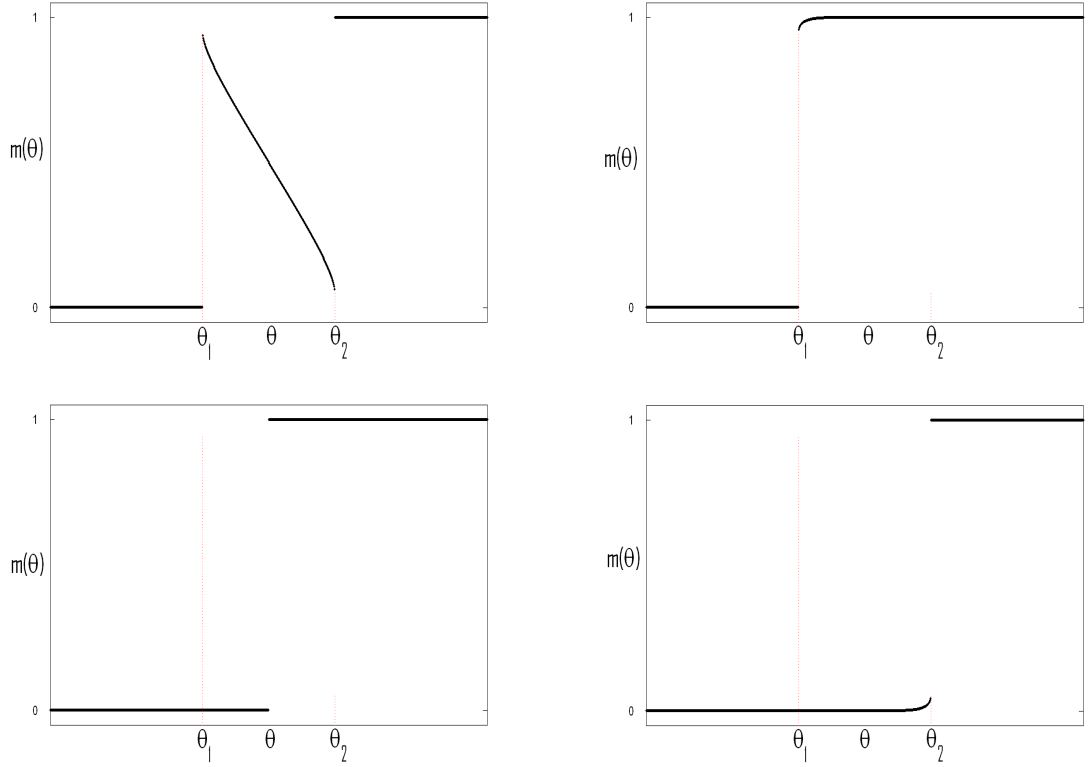
Figure 2.2: benchmark shapes of equilibria

four benchmark shapes.

Define the set of possible shapes of equilibria as

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

Note that

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

Given r and μ, an equilibrium $m\left(\theta - \theta_0\right)$ is determined by its shape $m \in \mathbf{M}\left(r,\mu\right)$ as well as its position θ_0. According to (2.10), the equilibrium condition for θ_0 is

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

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

$$g(\theta_0, m) \triangleq r/2 - \mu \cdot \ln\left(\frac{\int m(\theta - \theta_0) \cdot dP(\theta)}{1 - \int m(\theta - \theta_0) \cdot dP(\theta)}\right). \tag{2.12}$$

It is worth noting that since public information is summarized in common prior P, equation (2.12) also shows that public information affects the equilibrium only through changing its position θ_0 but leaves its shape unaffected.

As $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$ allows multiple shapes, a natural question is whether this multiplicity of possible shapes leads to multiple equilibria. We answer this question in Section 2.4.

2.4 Private Information Acquisition: Rigidity versus Flexibility

This section conveys the main results of this chapter. We first show that multiple equilibria emerge when strategic complementarity dominates information cost. We then contrast this result to that of an extended global game model to illustrate why rigid and flexible information acquisition play so different roles. We then show how and why more efficient coordination could be achieved through flexible rather than rigid information acquisition. Finally, we go beyond the entropic information cost to explore the essence of flexibility and its key aspect that drives our results.

Lemma 2.2 *For any possible shape $m \in \mathbf{M}(r, \mu)$, there exists $\theta_0 \in \mathbb{R}$ such that $m(\theta - \theta_0)$ is an equilibrium.*

Proof. See Appendix 2.9. ∎

This lemma proves the existence of the equilibrium. Moreover, it also provides a sufficient condition for multiple equilibria, as shown in the following proposition.

Proposition 2.3 *If $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$, then the costly information acquisition game has infinitely many equilibria.*

Proof. As shown in Section 2.3, $\#\mathbf{M}(r, \mu) = \infty$ when $\widetilde{r} = \frac{r}{4 \cdot \mu} > 1$. Therefore, this proposition is a direct implication of Lemma 2.2. ∎

This result is consistent with our previous intuition. Since strategic complementarity exceeds informational cost, multiple ways of information acquisition can be supported in equilibrium. Flexibility together with relatively low information cost enables players to achieve approximate common knowledge of payoffs and actions, which leads to multiplicity.

Besides the condition for multiplicity, we also enrich our results in Subsection 2.6.1 from the angle of supermodularity. This very property enables us to calculate the extreme equilibria and show the emergence of multiplicity as informational cost evolves.

In the rest of this section, we focus on the comparison between flexible and rigid information acquisition.

2.4.1 An Extended Global Game Model

In order to highlight the indispensable mechanism of flexibility, it is instructive to contrast the role of information acquisition in our benchmark model from that in an extended global game model. In this extended model, the players are allowed to purchase more accurate signals but cannot change any other aspect of the information structure. Specifically, let two players play the game with payoff matrix (Table 2.2.1). The common prior about fundamental θ is P. Player $i \in \{1, 2\}$ takes action $a_i \in \{0, 1\}$ after observing her private signal $x_i = \theta + \beta_i^{-1/2} \cdot \varepsilon_i$, where

ε_i is distributed according to a density function f with full support, $\mathbf{E}\varepsilon_i = 0$ and $Var(\varepsilon_i) < \infty$. Here β_i represents the precision of player i's private information. The cost of acquiring information of precision β is $c \cdot h(\beta)$, where $c > 0$ is an exogenous parameter controlling the difficulty of information acquisition and h is continuous and non-decreasing with $h(0) = 0$. The information structure is rigid in the sense that the additive nature of the signal generating process is not adjustable.

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

Proposition 2.4 *Let $(\beta_1(c), \beta_2(c))$ be the precision pair chosen in an equilibrium of $G(c)$. Then for any $\beta > 0$, there exists $\bar{c} > 0$, such that for all $c < \bar{c}$, $\beta_i(c) > \beta$, $i \in \{1, 2\}$.*

Proof. See Appendix 2.9. ∎

This proposition reads that 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 uniqueness is guaranteed if private information is sufficiently accurate relative to public information (e.g., Morris and Shin (2004)). Proposition 2.4 allows us to retrieve the standard global game result in this extended model with information acquisition.

Corollary 2.2 *For any $\delta > 0$, there exists $\bar{c} > 0$, such that for all $c < \bar{c}$, if strategy $s : \mathbb{R} \to \{0, 1\}$ survives iterated deletion of strictly dominated strategies in game $G(c)$, then $s(x) = 0$ for all $x \leq r/2 - \delta$ and $s(x) = 1$ for all $x \geq r/2 + \delta$.*

Proof. The proof is a direct application of Proposition 2.2 in Morris and Shin (2001) together with Proposition 2.4. According to Proposition 2.2 of Morris and

Shin (2001), $\forall \delta > 0$, $\exists \overline{\beta} > 0$, such that the above statement holds for all $\beta > \overline{\beta}$. Then Proposition 2.4 shows the existence of $\overline{c} > 0$ such that the players acquire information of precision at least $\overline{\beta}$. ∎

According to Corollary 2.2, all equilibria become approximately the unique switching strategy

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

when informational cost vanishes. This result is consistent with the standard global game arguments. That is, lowering informational cost induces more accurate private signals, which undermines common knowledge and thus facilitates the uniqueness of the equilibrium. According to Proposition 2.3, however, our model with flexible information acquisition has an opposite prediction: lowering informational cost enhances approximate common knowledge and facilitates multiplicity.

How should we understand the sharp discrepancy between these two approaches? The strategic complementarity between actions generates the coordination motive in acquiring information. This motive evolves to coordination in information acquisition in our benchmark model with flexibility, especially when information cost becomes lower. Hence, we recreate approximate common knowledge of the payoffs and actions, which leads multiplicity.

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

2.4.2   Welfare Implications: Rigidity versus Flexibility

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

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

However, this strategy cannot be supported in equilibrium. A player with a signal just above zero will rationally assign a large probability (close to $1/2$) to the event that her opponent's signal is negative, and thus be reluctant to invest due to the corresponding fear of miscoordination. This problem could be solved if they could commit to forgetting the exact values of their signals but only remembering the signs. For example, if there exists a third party who observes the signals and only tells the players the signs of their own signals, each player would find it optimal to invest if and only if her signal is positive. Hence, the players achieve the efficient coordination through throwing away information with a commitment device. It is worth noting that, in a trivial sense, the players always throw away some information since they map the continuous signals to binary actions. However, the key point here is that, although some information is not reflected in players' actions, it is still employed in making inference about others' beliefs and resulting actions. This is the way information matters, which differs from any other economic resource. Therefore, "throwing away information" means committing to forgetting such information, i.e., refraining from making inference upon such information.

In our benchmark model with flexible information acquisition, however, $\forall \widehat{\theta} \in$

$[0, r]$,

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

is an equilibrium when informational cost μ vanishes. Hence the most efficient strategy with cutoff $\widehat{\theta} = 0$ can be supported in equilibrium. Here, flexible information acquisition together with the fact that information is costly helps players only acquire information valuable for efficient coordination and refrain themselves from information harmful to coordination even if information cost goes to zero, and thus could become strictly better off than in the case of rigid information acquisition. In contrast to the extended global game model where players can only throw away harmful information with an explicit commitment device, the flexibility in our benchmark model helps players choose the quantitative and qualitative natures of their information, through which way they act as if they commit to throwing away information harmful to coordination. In other words, we can interpret that an implicit commitment device inhabits the flexibility of information acquisition. It should be highlighted that this contrast could not be seen in single person decision problems with information acquisition. This is not only because, literally, single person decision problems do not involve coordination, but also due to the fact that more information is always more desirable disregarding its cost, no matter whether information acquisition is flexible or rigid.

2.4.3 Flexibility: General Information Cost

This subsection goes beyond the entropic information cost to explore the essence of flexibility and its key aspect that drives our results. To this end, it is important to first focus on the fact that different forms of information acquisition could be exclusively captured by different schemes of information cost. For example, if information

structure is endowed exogenously, it could be viewed that only this endowed information structure can be acquired at finite cost while all other information structures incur an infinite cost. If information acquisition is endogenous but rigid as in the extended global game model described above, only information structures following the form of $x = \theta + \beta^{-1/2} \cdot \varepsilon$ are associated with finite cost. In our benchmark model where information acquisition follows rational inattention, any information structure is associated with a cost proportional to the resulted reduction of entropy. In all, analysis on schemes of information cost covers any consideration on the forms of information acquisition.

It is worth noting that the information cost given by rational inattention respects Blackwell's ordering[16], due to which it suffices for players to only consider binary information structures; and further features uniform boundedness for those binary information structures, which ensures the availability of any such essential information. According to these two aspects, we generalize the concept of flexible information acquisition.

Definition 2 *Information acquisition is flexible, if the information cost respects Blackwell's ordering and is uniformly bounded over* Ω.

The condition regarding Blackwell's ordering precludes the information structures that contain information of no potential value[17], and the uniform boundedness guarantees the availability of all potentially valuable information structures. Having abstracted the two aspects from rational inattention, we can go beyond the entropic information cost to show the validity of our main results in more general settings.

---

[16]An information cost respects Blackwell's ordering if it assigns lower cost to less informative information structures. An information structure is less informative than the other if it can be obtained from the other by adding garbling noise in the sense of Blackwell (1953). See Blackwell (1953) for detailed discussion.

[17]I.e., non-binary information structures, which contain redundant information.

Consider a coordination game with payoffs given by Table 2.2.1. Information acquisition is flexible, and thus we can focus on Ω, the set of binary information structures. Consider a general information cost $\mu \cdot c(m)$, where $\mu > 0$, $c(\cdot)$ is a non-negative functional defined on Ω and m has the same interpretation as before.

Proposition 2.5 *If information acquisition is flexible under $c(\cdot)$ and $c(\cdot)$ is submodular, then multiple equilibria exist when*

$$\frac{r}{4\mu} > \frac{2K}{\min\left(\Pr\left(\theta \leq 0\right), \Pr\left(\theta \geq r\right)\right)} \, , \tag{2.13}$$

where K is any uniform bound of $c(\cdot)$ over Ω.

Proof. See Appendix 2.10. ∎

This proposition is a generalization of Proposition 2.3 and has the same interpretation. The strategies in Ω can be partially ordered by the pointwise ordering, under which Ω becomes a complete lattice. Hence, we can prove Proposition 2.5 by constructing two equilibria close to the largest and smallest equilibria of the game with complete information (i.e., $\mu = 0$), respectively. The submodularity of $c(\cdot)$ is assumed to guarantee the existence of such equilibria but is not essential[18]. Here we find that multiple equilibria emerge under a condition similar to that in Proposition 2.3 regardless of the specific functional form of $c(\cdot)$. This proposition supports our intuition that the flexibility rather than the entropic functional form drives our result. The entropic information cost is employed in our benchmark model not only for its meaningful interpretation in information theory, but also because it allows

---

[18]Submodularity is not a strange assumption, as the information cost given by rational inattention is submodular. Adding this assumption relieves us from the technicalities of searching fixed points in a non-compact space of functions of θ. However, this assumption may not be necessary, since the original problem can always be approximated by discretizing the domain of θ, which leads to a compact space of functions of θ and thus guarantees the existence of fixed points.

us to obtain a clear condition $r > 4\mu$, which intuitively inspires our thinking of the current problem.

We define that information acquisition is rigid if it is not flexible. The standard global game models and the extended global game model discussed in previous subsections are two typical setups with rigid information acquisition. They violate both conditions in our definition of flexibility.

Moreover, it is worth highlighting that the second rather than the first condition in our definition of flexibility is essential. Consider the information costs respecting Blackwell's ordering, which allows us to focus on Ω, the set of binary information structures. We show that almost any strategy $m^* \in \Omega$ can be supported in equilibrium through properly choosing such an information cost featuring some rigidity. Specifically, choose a subset $S \subset \Omega$ such that $m^* \in S$ is the unique equilibrium of our coordination game when players can choose any member of S without any cost but have no choice outside S. Define an information cost such that $c(m) = \infty$ for all $m \in \Omega \backslash S$ and $c(m^*) \leq c(m) < \infty$ for all $m \subset S$. Then for all $\mu \in \mathbb{R}_+$, m^* is the unique equilibrium of the coordination game with information cost $\mu \cdot c(m)$. The key point here is that restricting the players within subset S physically precludes some information that is potentially valuable. Therefore, the rigidity caused by the failure of the second condition, the uniform boundedness, leads to the uniqueness.

2.5   Impacts of Public Information

In our benchmark model, players acquire private information at some cost. We assume that public information is directly observable without incurring any cost. Hence public information is common knowledge. It affects players' decisions through changing the common prior about the fundamental. This section conducts a comparative static analysis with respect to common prior P to study the impacts of

public information. Roughly speaking, if the common prior concentrates within the intermediate region $[0, r]$, both players are confident that the event $\{\theta \in [0, r]\}$ happens with high probability. As a result, their coordination motive dominates their concern of fluctuating fundamental and multiple equilibria emerge regardless of the informational cost.

Proposition 2.6 *For any $r > 0$ and $\mu > 0$, the costly information acquisition game has multiple equilibria if $\mathbf{E}e^{-\mu^{-1}\theta} \leq 1$ and $\mathbf{E}e^{\mu^{-1}(\theta-r)} \leq 1$, where the expectation is taken according to common prior P.*

Proof. See Appendix 2.10. ∎

Under the condition of this proposition, at least both "always invest" (i.e., $m(\theta) = 1$ almost surely) and "never invest" (i.e., $m(\theta) = 0$ almost surely) are equilibria[19]. The players find it optimal to not acquire any information and coordinate their investment decisions perfectly. We can gain some intuition from a Gaussian common prior $N(t, \sigma^2)$. In this case, it is easy to verify that condition

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

is equivalent to

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

That is, the common prior should have a small dispersion and its probability peak should be close to $r/2$. Proposition 2.6 is strong in the sense that the criterion

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

---

[19]We prove this result in Proposition 2.8 and 2.9 in Appendix 2.8.

is uniform for all common priors. To make our result comparable to the standard global game results, we establish the following corollary:

Corollary 2.3 *Let $p(\theta)$ be a probability density function. Then, for any $r > 0$, $\mu > 0$ and $y \in (0, r)$, there exists $\overline{\beta} > 0$ such that for all $\beta > \overline{\beta}$, the costly information acquisition game with common prior $\beta^{1/2} p\left(\beta^{1/2}(\theta - y)\right)$ (density function) has multiple equilibria.*

Proof. First note that

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

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

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

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

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

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

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

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

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

Let $\overline{\beta} = \max\left(\overline{\beta}_1, \overline{\beta}_2\right)$, then according to Proposition 2.6, the game has multiple equilibria for all $\beta > \overline{\beta}$. ∎

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

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

where ε is distributed according to a density function p. This public signal results in an updated common prior[20] with density function $\beta^{1/2} p\left(\beta^{1/2}(\theta - y)\right)$. Therefore, Corollary 2.3 reads that providing public information of high precision leads to multiplicity. This is consistent with the well known result in the literature of global games.

Another famous result is that uniqueness is guaranteed if private signals are sufficiently accurate relative to public signals (e.g., Morris and Shin (2004)). Hence, regarding the uniqueness, the effects of increasing precision of public signals can be offset by increasing the precision of private signals. In the context of our extended global game model, Corollary 2.2 implies that the effect of increasing precision of public signals can be offset by lowering cost of acquiring private information. In our benchmark model with flexible information acquisition, however, Proposition 2.3 states that there are always infinitely many equilibria when $\widetilde{r} = \frac{r}{4\cdot\mu} > 1$, regardless of the precision of public information. That is, the effects of public information and private information acquisition are disentangled. The reason is that when

---

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

information cost is small, players have enough freedom in coordinating their private information acquisition. This freedom has nothing to do with public information. Therefore, the entangled effects in global game models also result from the rigidity implicitly imposed on the information structure.

2.6 Discussion

In this section, we first enrich our results through the angle of supermodularity. We then discuss a related game in which players are endowed with a capacity to acquire a fixed amount of information at no cost. Finally, we address three extensions to the benchmark model, i.e., a) $n \geq 2$ players; b) state-dependent strategic complementarity $r = r(\theta)$; and c) discontinuous payoff gain (from choosing one action over the other) with respect to the fundamental and to the proportion of players taking a specific action.

2.6.1 Supermodularity and Extreme Equilibria

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

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

Note that the mutual information $I(\cdot)$ is submodular over strategy space Ω, thus our game is a supermodular one. The set of equilibria of a supermodular game forms a complete lattice, the supremum and infimum of which correspond to the largest and smallest Nash equilibria of the game. We refer the interested readers to Topkis (1979), Milgrom and Roberts (1990) and Vives (1990) for more details of supermodular games. Here we utilize this property to calculate the extreme
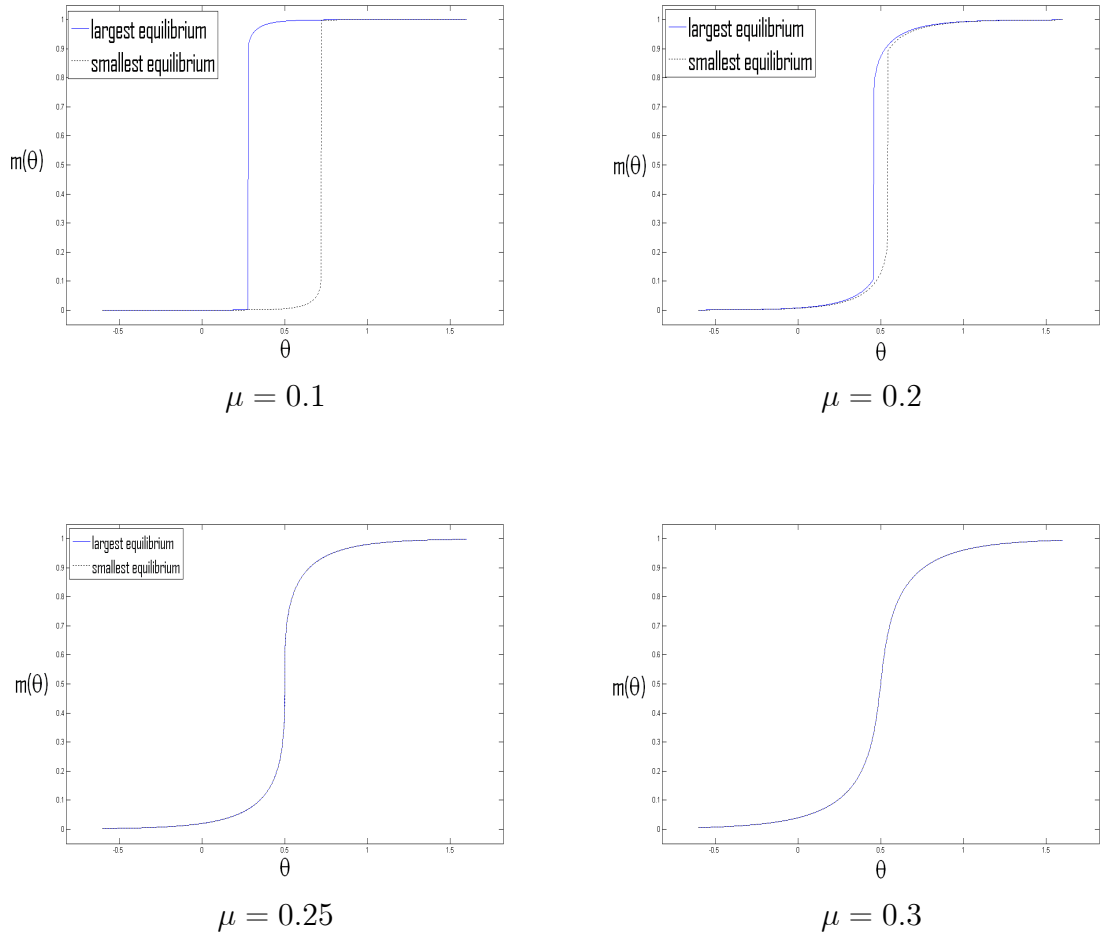
Figure 2.3: extreme equilibria

equilibria of our game. Figure 2.3 shows how multiple equilibria emerge as μ, the marginal cost of information acquisition shrinks.

These graphs show numerical solutions of the game with strategic complementarity $r = 1$ and uniform common prior over interval $[-0.6, 1.6]$. We vary the value of μ, the marginal cost of information acquisition to see how the equilibria evolve. In the two graphs of the first row, r is larger than 4μ. Hence we have multiple shapes for the equilibria. It is clear that multiple equilibria emerge as the largest and smallest equilibria differ from each other. When μ increases from 0.1 to 0.2, the two extreme equilibria get closer, suggesting a tendency to uniqueness. In the

southwest graph, r equals 4μ and thus we have a unique shape. We still have multiple equilibria, although it is hard to distinguish them from the graph. In the southeast graph, r is smaller than 4μ and we have a unique equilibrium.

2.6.2 Constrained Information Acquisition Game

In the conventional setup of rational inattention models, the decision maker is capacity constrained. She is able to acquire information up to a given amount without incurring any cost. This subsection discusses such case. Let $\kappa > 0$ denote the maximal amount of information that players can acquire. By an argument similar to that of Lemma 2.1, player i's equilibrium strategy is characterized by a function $m_i \in L^1(\Theta, P)$, and an equilibrium is a pair (m_1, m_2) solving the the following problem:

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

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

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

Since payoffs are symmetric and players have the same capacity of information acquisition, all equilibria of this game are symmetric.[21] When solving for the equilibrium, the multiplier for the capacity constraint plays a role similar to μ, the marginal cost of information acquisition in our benchmark model. However, it is worth highlighting two differences. First, the multiplier is an endogenous variable. Its value may vary for different equilibria. Hence, it is difficult to conduct comparative static analysis in this setup. Second, switching strategies could be supported in equilibrium when κ is large enough (e.g., $\kappa > \ln 2$ nats). Since a binary decision problem requires at most $\ln 2$ nats of information, the capacity constraint does not bind for large κ. This case corresponds to our benchmark model with zero marginal

---

[21]This symmetry is proved in an earlier version of this paper and is omitted here.

information cost, and thus has multiple equilibria.

2.6.3 The Game with Multiple Players

We have focused on a 2-player game, but our arguments also work for games with multiple players. Suppose now we have $n \geq 2$ players. Let $N \triangleq \{1, 2, \cdots, n\}$ denote the set of players. All the other assumptions remain the same, except that player $i \in N$ enjoys a payoff

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

from choosing "*invest*" if the fundamental is θ and n' other players choose "*invest*". Obviously, Lemma 2.1 is still applicable here, thus again player i's strategy can be characterized by

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

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

$$\theta - r \cdot \left(1 - \frac{\sum_{j \neq i} m_i(\theta)}{n-1}\right). \tag{2.14}$$

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

$$
\begin{aligned}
m_i \quad \in \quad & \arg\max_{\widetilde{m}_i \in L^1(\Theta, P)} V_i(\widetilde{m}_i, m_{-i}) = U_i(\widetilde{m}_i, m_{-i}) - \mu \cdot I(\widetilde{m}_i) \\
\text{s.t.} \quad \widetilde{m}_i(\theta) \quad \in \quad & [0, 1] \text{ for all } \theta \in \Theta.
\end{aligned}
$$

Similar to the argument in Proposition 2.2, we can show that all equilibria are symmetric. Hence any equilibrium can be represented by a single function m, and all the remaining arguments in the benchmark model still work here.

2.6.4   State-dependent Strategic Complementarity

A natural extension of the benchmark model is to allow players' coordination motive to vary with the fundamental, i.e., $r = r(\theta)$. Under the condition $r(\theta) > 0$ almost surely, the proof of Lemma **2.2** is still valid if we simply replace r with $r(\theta)$. Hence the game only admits symmetric equilibria, and most remaining analysis proceeds as does in the benchmark model.

2.6.5   Discontinuous Payoff Gain Function

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

Follow the notation in Morris and Shin (2001), let $\pi(m, \theta)$ denote the payoff gain of choosing "*invest*" when the fundamental is θ and the opponent chooses "*invest*" with probability m. For example, in our benchmark model $\pi(m, \theta) = \theta - r \cdot (1 - m)$, and $\frac{\partial \pi}{\partial m} = r$ represents the strategic complementarity. When $\pi(m, \theta)$ is discontinuous with respect to m and θ, $\frac{\partial \pi}{\partial m} = \infty$ for some (m, θ), which implies an infinite strategic complementarity. The intuition developed in our benchmark model suggests multiple equilibria no matter how large is μ, as shown in the following example. The underlying story in our mind is the currency attack model of Morris and Shin (1998).

There is a continuum of players playing a costly information acquisition game.

Their payoff gain from choosing "attack" is defined as

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

The interpretation of this payoff gain is as following. When the fundamental of the currency is weak and too many speculators are attacking, government has to drop the currency peg and each attacker enjoys one dollar. When the opposite happens, the currency attack fails and each attacker loses one dollar. A speculator receive zero if she does not attack.

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

$$\frac{A}{1+2A}e^{\mu^{-1}} + \frac{1+A}{1+2A}e^{-\mu^{-1}} > 1. \tag{2.15}$$

Then according to Proposition 1.1, any equilibrium satisfies

$$\forall \theta \in [-A, 1+A], \pi(m, \theta) = \mu \cdot \left[\ln\left(\frac{m(\theta)}{1-m(\theta)}\right) - \ln\left(\frac{p_I}{1-p_I}\right) \right], \tag{2.16}$$

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

Since $\pi(m, \theta)$ takes only two possible values, so does $m(\theta)$ according to (2.16). Hence the equilibrium strategy can be represented by two numbers $\underline{m}, \overline{m} \in [0, 1]$. Let $S_I \triangleq \{\theta \in [-A, 1+A] : m(\theta) = \overline{m}\}$ and $S_N \triangleq \{\theta \in [-A, 1+A] : m(\theta) = \underline{m}\}$ denote the region of "attack" with "high" probability and "low" probability, respectively. By definition, we have $S_I \subset [-A, \overline{m}]$ and $S_N \subset (\underline{m}, 1+A]$. Then a symmetric

113

equilibrium is characterized by $\underline{m}$, $\overline{m}$, $S_I \subset [-A, \overline{m}]$ and $S_N \subset (\underline{m}, 1 + A]$ such that

$$1 = \mu \cdot \left[\ln \left(\frac{\overline{m}}{1 - \overline{m}} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right], \tag{2.17}$$

$$-1 = \mu \cdot \left[\ln \left(\frac{\underline{m}}{1 - \underline{m}} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right], \tag{2.18}$$

and

$$p_I = \frac{1}{2 \cdot A + 1} \cdot \left[\Pr(S_I) \cdot \overline{m} + \Pr(S_N) \cdot \underline{m} \right]. \tag{2.19}$$

Proposition 2.7 *This game with discontinuous payoff gain has infinitely many equilibria for all $\mu > 0$.*

Proof. See Appendix 2.10. ∎

This proposition confirms our intuition learnt from the benchmark model. Discontinuous payoff gain generates infinite strategic complementarity. Therefore, coordination motive always dominates information cost and infinitely many ways of coordinating information acquisition can be supported in equilibrium.

2.7 Appendix

Proof of Lemma 2.1.

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

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

and

$$\widetilde{\sigma}_i\left(s_{i,I}\right) = 1,$$

$$\widetilde{\sigma}_i\left(s_{i,N}\right) = 0,$$

$$\widetilde{\sigma}_i\left(s_{i,ind}\right) \in [0,1].$$

Since player i is indifferent between *invest* and *not invest* upon $S_{i,ind}$, we have

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

However, if $\#(S_i) > 3$, $\left(\widetilde{S}_i,\widetilde{q}_i\right)$ incurs strictly less mutual information than does (S_i,q_i), i.e., $I\left(\widetilde{q}_i\right) - I\left(q_i\right) < 0$. The reason is that $\left(\widetilde{S}_i,\widetilde{q}_i\right)$ does not require player i to discern signal realizations within any of $S_{i,I}$, $S_{i,N}$, and $S_{i,ind}$. Then

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

i.e., $\left(\left(S_i,q_i\right),\sigma_i\right)$ is suboptimal and cannot be an equilibrium strategy. Hence we proved that in any equilibrium, $\forall i \in \{1,2\}$, $\#(S_i) \le 3$, $S_{i,I} = \{s_{i,I}\}$, $S_{i,N} = \{s_{i,N}\}$, and $S_{i,ind} = \{s_{i,ind}\}$.

Now we prove $\forall i \in \{1,2\}$, $\Pr\left(S_{i,ind}\right) = 0$ (i.e., $\Pr\left(s_{i,ind}\right) = 0$).

Suppose $\exists i \in \{1,2\}$, $\Pr\left(S_{i,ind}\right) > 0$ (thus, $\Pr\left(s_{i,ind}\right) > 0$). 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 $\Pr\left(s_{i,I}\right) > 0$.

Construct a new strategy $\left(\left(\overline{S}_i, \overline{q}_i\right), \overline{\sigma}_i\right)$ with $\overline{S}_i = \{\overline{s}_{i,I}, \overline{s}_{i,N}\}$ such that $\forall \theta \in \Theta$,

$$\overline{q}_i\left(\overline{s}_{i,I}|\theta\right) = q_i\left(s_{i,I}|\theta\right) + q_i\left(s_{i,ind}|\theta\right),$$

$$\overline{q}_i\left(\overline{s}_{i,N}|\theta\right) = q_i\left(s_{i,N}|\theta\right),$$

$$\overline{\sigma}_i\left(\overline{s}_{i,I}, a_i = 1\right) = 1$$

and

$$\overline{\sigma}_i\left(\overline{s}_{i,N}, a_i = 0\right) = 1 .$$

Since player i is indifferent between *invest* and *not invest* upon receiving $s_{i,ind}$, we have

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

Since $\Pr\left(s_{i,I}\right) > 0$ and $\Pr\left(s_{i,ind}\right) > 0$, this new strategy $\left(\left(\overline{S}_i, \overline{q}_i\right), \overline{\sigma}_i\right)$ incurs strictly less mutual information then does $\left(\left(S_i, q_i\right), \sigma_i\right)$. Then by the same argument, we know that $\left(\left(S_i, q_i\right), \sigma_i\right)$ is suboptimal and cannot be an equilibrium strategy.

Now the only possibility is $\Pr\left(s_{i,ind}\right) = 1$, i.e., player i does not acquire any information in equilibrium since she is always indifferent between *invest* and *not invest*. Let $m_i\left(\theta\right) \triangleq \Pr\left(a_i = 1|\theta\right)$ be the probability that player i invests when fundamental equals θ. Then m_i is totally determined by player i's strategy $\left(\left(S_i, q_i\right), \sigma_i\right)$. $\Pr\left(s_{i,ind}\right) = 1$ implies that in equilibrium, $\Pr\left(\theta - r \cdot \left(1 - m_j\left(\theta\right)\right) = 0\right) = 1$, and $\exists m \in [0, 1]$, s.t. player i always invests with probability m, i.e., $\Pr\left(m_i\left(\theta\right) = m\right) = 1$. Let

$$F_+ = \{\theta \in \Theta | \theta - r \cdot (1 - m) > 0\},$$

$$F_- = \{\theta \in \Theta | \theta - r \cdot (1 - m) < 0\}$$

and

$$B = \{\theta \in \Theta | \theta - r \cdot (1 - m) = 0\} .$$

Since the common prior P is absolutely continuous with respect to Lebesgue measure, we have $\Pr(B) = 0$, i.e., $\Pr(F_+) + \Pr(F_-) = 1$. Without loss of generality, assume $\Pr(F_+) > 0$. Construct a new strategy $\left(\left(\widetilde{S}_j, \widetilde{q}_j \right), \widetilde{\sigma}_j \right)$ for player j, s.t.

$$\widetilde{S}_j = \{0, 1\} ,$$

$$\widetilde{q}_j (s_j = 1 | \theta) = 1 \text{ if } \theta \in F_+ \cup B,$$

$$\widetilde{q}_j (s_j = 0 | \theta) = 1 \text{ if } \theta \in F_-$$

and

$$\widetilde{\sigma}_j (s_j = 1) = 1, \ \widetilde{\sigma}_j (s_j = 0) = 0 .$$

Note that $\forall \theta \in F_+$,

$$\theta - r \cdot (1 - m) > 0 = \theta - r \cdot (1 - m_j (\theta)) ,$$

i.e.,

$$m_j (\theta) < m_i \leq 1 .$$

Then we have

$$
U_j \left(\left(\left(\widetilde{S}_j, \widetilde{q}_j \right), \widetilde{\sigma}_j \right), \left((S_i, q_i), \sigma_i \right) \right) - U_j \left(\left((S_j, q_j), \sigma_j \right), \left((S_i, q_i), \sigma_i \right) \right)
$$

$$
= \int_{F_+ \cup B} 1 \cdot [\theta - r \cdot (1 - m)] \cdot dP(\theta) + \int_{F_-} 0 \cdot [\theta - r \cdot (1 - m)] \cdot dP(\theta)
$$

$$
- \int m_j(\theta) \cdot [\theta - r \cdot (1 - m)] \cdot dP(\theta)
$$

$$
\geq \int_{F_+} [1 - m_j(\theta)] \cdot [\theta - r \cdot (1 - m)] \cdot dP(\theta)
$$

$$
> \quad 0
$$

On the other hand, it is obvious that $I(\widetilde{q}_j) \leq I(q_j)$. Therefore, $\left(\left(\widetilde{S}_j, \widetilde{q}_j \right), \widetilde{\sigma}_j \right)$ strictly dominates $((S_j, q_j), \sigma_j)$ and thus $((S_j, q_j), \sigma_j)$ cannot be player j's equilibrium strategy.

Now we proved $\forall i \in \{1, 2\}$, $\Pr(S_{i,ind}) = 0$ by contradiction. Together with the previous result that $\forall i \in \{1, 2\}$, $\#(S_i) \leq 3$, it also implies that $\#(S_i) = 1$ or 2. ∎

2.8   Appendix

Proof of Proposition 2.2.

Proof. According to (2.7),

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

$$
\theta - r \cdot (1 - m_1(\theta)) \ = \ \mu \cdot \left[\ln \left(\frac{m_2(\theta)}{1 - m_2(\theta)} \right) - \ln \left(\frac{p_{I2}}{1 - p_{I2}} \right) \right], \qquad (2.20)
$$

$$
\theta - r \cdot (1 - m_2(\theta)) \ = \ \mu \cdot \left[\ln \left(\frac{m_1(\theta)}{1 - m_1(\theta)} \right) - \ln \left(\frac{p_{I1}}{1 - p_{I1}} \right) \right]. \qquad (2.21)
$$

(2.20) and (2.21) imply

$$\forall \theta \in \Theta,$$
$$\left[\ln \left(\frac{p_{I2}}{1 - p_{I2}} \right) - \ln \left(\frac{p_{I1}}{1 - p_{I1}} \right) \right]$$
$$= \left[\ln \left(\frac{m_2 \left(\theta \right)}{1 - m_2 \left(\theta \right)} \right) - \ln \left(\frac{m_1 \left(\theta \right)}{1 - m_1 \left(\theta \right)} \right) \right] + \frac{r}{\mu} \left(m_2 \left(\theta \right) - m_1 \left(\theta \right) \right). \quad (2.22)$$

If $p_{I2} = p_{I1}$, (2.22) becomes

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

and we must have $m_2 \left(\theta \right) = m_1 \left(\theta \right)$ a.s. since $\frac{r}{\mu} > 0$. Now suppose $p_{I2} \neq p_{I1}$. Without loss of generality, let $p_{I2} > p_{I1}$. Denote $z = \ln \left(\frac{p_{I2}}{1 - p_{I2}} \right) - \ln \left(\frac{p_{I1}}{1 - p_{I1}} \right) > 0$. Then (2.22) becomes

$$\forall \theta \in \Theta,$$
$$0 < z = \left[\ln \left(\frac{m_2 \left(\theta \right)}{1 - m_2 \left(\theta \right)} \right) - \ln \left(\frac{m_1 \left(\theta \right)}{1 - m_1 \left(\theta \right)} \right) \right] + \frac{r}{\mu} \left(m_2 \left(\theta \right) - m_1 \left(\theta \right) \right) \quad (2.23)$$

which suggests that $\Pr \left(m_2 \left(\theta \right) > m_1 \left(\theta \right) \right) = 1$. Let $\ln \left(\frac{m_2(\theta)}{1 - m_2(\theta)} \right) = x \left(\theta \right)$ and $\ln \left(\frac{m_1(\theta)}{1 - m_1(\theta)} \right) = y \left(\theta \right)$. (2.23) implies

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

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

119

$m_1\left(\theta\right)=\frac{\exp(y(\theta))}{1+\exp(y(\theta))}$, thus

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

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.,

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

i.e.,

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

i.e.,

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

$$
(2.24)
$$

where

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

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

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

and

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

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

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

where the last inequality follows the fact that $z > 0$. Therefore, the left hand side of (2.24) is strictly negative, which is a contradiction. Therefore, $\Pr\left(m_1\left(\theta\right) = m_2\left(\theta\right)\right) = 1$. ∎

Lemma 2.3 *The costly information acquisition game has an equilibrium with at least one player always investing if and only if* $\mathbf{E}\exp\left(-\mu^{-1}\theta\right) \leq 1$, *where the expectation is taken according to common prior* P.

Proof. (Sufficiency.) If $m_j\left(\theta\right) = 1$ almost surely, player i's payoff gain from investing over not investing becomes

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

Then according to case a) in ii) of Proposition 1.1,

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

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

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

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

According to case b) and c) in ii) of Proposition 1.1, $\mathbf{E}\exp(-\mu^{-1}\theta) > 1$ implies $m_i(\theta) < 1$ almost surely. Then player j's payoff gain from investing over not investing becomes

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

which implies

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

Hence according to case b) and c) in ii) of Proposition 1.1, we find $m_j(\theta) < 1$ almost surely, which is a contradiction. ∎

Lemma 2.4 *In an equilibrium of the costly information acquisition game with one player always investing, the other player must also always invest.*

Proof. By the necessity part of Lemma 2.3, we know that $\mathbf{E}\exp(-\mu^{-1}\theta) \leq 1$. Then the sufficiency part of Lemma 2.3 has already proved that the other player

must also always invest. ∎

Proposition 2.8 *The costly information acquisition game has an equilibrium with both players always investing if and only if $\mathbf{E}\exp\left(-\mu^{-1}\theta\right) \leq 1$.*

Proof. This proposition is a direct implication of Lemma 2.3 and 2.4. ∎

Lemma 2.5 *The costly information acquisition game has an equilibrium with at least one player always not investing if and only if $\mathbf{E}\exp\left(\mu^{-1}\theta\right) \leq e^{\mu^{-1}r}$, where the expectation is taken according to common prior P.*

Proof. (Sufficiency.) If $m_j\left(\theta\right) = 0$ almost surely, player i's payoff gain from investing over not investing becomes

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

Then according to case b) in ii) of Proposition 1.1,

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

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

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

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

According to case a) and c) in ii) of Proposition 1.1, $\mathbf{E}\exp\left(\mu^{-1}\left(\theta - r\right)\right) > 1$ implies $m_i\left(\theta\right) > 0$ almost surely. Then player j's payoff gain from investing over not

123

investing becomes

$$\Delta u_j \left(\theta \right) \;\; = \;\; \theta - r \cdot \left[1 - m_i \left(\theta \right) \right]$$

$$> \;\; \theta - r \text{ almost surely,}$$

which implies

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

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

$$> \;\; 1 \, .$$

Hence according to case a) and c) in ii) of Proposition 1.1, we find $m_j \left(\theta \right) > 0$ almost surely, which is a contradiction. ∎

Lemma 2.6 *In an equilibrium of the costly information acquisition game with one player always not investing, the other player must always not invest either.*

Proof. By the necessity part of Lemma 2.5, we know that $\mathbf{E} \exp \left(\mu^{-1} \theta \right) \le e^{\mu^{-1} r}$. Then the sufficiency part of Lemma 2.5 has already proved that the other player must always not invest either. ∎

Proposition 2.9 *The costly information acquisition game has an equilibrium with both players always not investing if and only if* $\mathbf{E} \exp \left(\mu^{-1} \theta \right) \le e^{\mu^{-1} r}$.

Proof. This proposition is a direct implication of Lemma 2.5 and 2.6. ∎

2.9 Appendix

Proof of Lemma 2.2.

Proof. Let $m \in \mathbf{M}(r, \mu)$ be an arbitrary shape. Let $\theta_0(p_I)$ be defined by (2.10) and

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

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

$$\theta - r \cdot (1 - m(\theta, p_I)) = \mu \cdot \left[\ln \left(\frac{m(\theta, p_I)}{1 - m(\theta, p_I)} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) \right] \text{ almost surely.}$$
(2.25)

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

$$p_I = \int_\Theta m(\theta, p_I) \, dP(\theta) .$$
(2.26)

Therefore, our objective is to show the existence of $p_I \in (0, 1)$ satisfying (2.26).

Step 1. We show

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

for p_I sufficiently close to 1.

By (2.25),

$$\ln \left(\frac{m(\theta, p_I)}{1 - m(\theta, p_I)} \right) - \ln \left(\frac{p_I}{1 - p_I} \right) < \mu^{-1}\theta \text{ almost surely,}$$

i.e.,

$$m(\theta, p_I) < \frac{\frac{p_I}{1 - p_I}}{e^{-\mu^{-1}\theta} + \frac{p_I}{1 - p_I}} \text{ almost surely.}$$

Hence it suffices to show

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

Let

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

and

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

then it suffices to show

$$\int_\Theta \frac{w-1}{v(\theta) + w} dP(\theta) \leq \frac{w-1}{w} . \tag{2.27}$$

Since $w > 1$ by definition, (2.27) becomes

$$\int_\Theta \frac{1}{1 + v(\theta)/w} dP(\theta) \leq 1 . \tag{2.28}$$

By assumption,

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

i.e.,

$$\int_\Theta v(\theta) dP(\theta) > 0 . \tag{2.29}$$

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

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

Let

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

then

$$|v(\theta)| \leq B$$

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

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

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

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

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

$$w > \max\left(\overline{w}, \frac{2B^2}{\int_{\Theta \cap [-N,+\infty)} v(\theta)\, dP(\theta)}\right), \tag{2.30}$$

then

$$
\begin{aligned}
& \int_{\Theta \cap [-N,+\infty)} \frac{1}{1 + v(\theta)/w} dP(\theta) \\
< \; & \int_{\Theta \cap [-N,+\infty)} \left[1 - \frac{v(\theta)}{w} + \frac{2B^2}{w^2}\right] dP(\theta) \\
= \; & \Pr(\theta \geq -N) + \frac{2B^2}{w^2} \cdot \Pr(\theta \geq -N) - w^{-1} \int_{\Theta \cap [-N,+\infty)} v(\theta)\, dP(\theta) \\
\leq \; & \Pr(\theta \geq -N) + \frac{2B^2}{w^2} - w^{-1} \int_{\Theta \cap [-N,+\infty)} v(\theta)\, dP(\theta) \\
< \; & \Pr(\theta \geq -N), \tag{2.31}
\end{aligned}
$$

where the last inequality follows (2.30). Hence,

$$
\begin{aligned}
&\int_{\Theta} \frac{1}{1 + v\left(\theta\right)/w} dP\left(\theta\right) \\
={}& \int_{\Theta \cap [-N, +\infty)} \frac{1}{1 + v\left(\theta\right)/w} dP\left(\theta\right) + \int_{\Theta \cap (-\infty, -N)} \frac{1}{1 + v\left(\theta\right)/w} dP\left(\theta\right) \\
\leq{}& \int_{\Theta \cap [-N, +\infty)} \frac{1}{1 + v\left(\theta\right)/w} dP\left(\theta\right) + \int_{\Theta \cap (-\infty, -N)} 1 \cdot dP\left(\theta\right) \\
<{}& \Pr\left(\theta \geq -N\right) + \Pr\left(\theta < -N\right) \\
={}& 1,
\end{aligned}
$$

where the first inequality holds since $v\left(\theta\right)$ for all $\theta \in (-\infty, -N)$ and the last inequality comes from (2.31). Therefore, (2.28) holds and if we let

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

we have

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

Step 2. We show

$$
\int_{\Theta} m\left(\theta, p_I\right) dP\left(\theta\right) > p_I
$$

for p_I sufficiently close to 0.

By (2.25),

$$
\ln\left(\frac{m\left(\theta, p_I\right)}{1 - m\left(\theta, p_I\right)}\right) - \ln\left(\frac{p_I}{1 - p_I}\right) > \mu^{-1}\left(\theta - r\right) \text{ almost surely,}
$$

i.e.,

$$
1 - m\left(\theta, p_I\right) < \frac{1}{1 + e^{\mu^{-1}(\theta - r)} \frac{p_I}{1 - p_I}} \text{ almost surely.}
$$

Hence it suffices to show

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

Let

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

and

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

then it suffices to show

$$\int_\Theta \frac{w - 1}{v(\theta) + w} dP(\theta) \leq \frac{w - 1}{w} . \tag{2.32}$$

By assumption,

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

i.e.,

$$\int_\Theta v(\theta) dP(\theta) > 0 . \tag{2.33}$$

Note that (2.32) and (2.33) are the same as (2.28) and (2.29), thus (2.32) can be proved by the same argument in Step 1. Therefore, we can find a $\underline{p}_I \in (0, 1)$ s.t.

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

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

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

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

of $p_I^* \in (0, 1)$ s.t.

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

According to (2.10), let

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

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

Lemma 2.7 *Let P be any probability measure over $\mathbb{R}$. A set of functions $M \subset L^1(\mathbb{R}, P)$ is relatively compact if M is uniformly bounded and equicontinuous.*

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

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

then $\cup_{T=1}^\infty A_T$ is dense in $\mathbb{R}$. Since $\cup_{T=1}^\infty A_T$ is countable, we can list its elements as $\{\theta_1, \theta_2, \theta_3, \cdots\}$. Note that the numerical sequence $\{m_n(\theta_1)\}_{n=1}^\infty$ is bounded, so by Bolzano-Weierstrass theorem it has a convergent subsequence, which we will write using double subscripts: $\{m_{1,n}(\theta_1)\}_{n=1}^\infty$. Now the numerical sequence $\{m_{1,n}(\theta_2)\}_{n=1}^\infty$ is also bounded, so it has a convergent subsequence $\{m_{2,n}(\theta_2)\}_{n=1}^\infty$. Note that the sequence of functions $\{m_{2,n}\}_{n=1}^\infty$ converges at both θ_1 and θ_2 since it is a subsequence of $\{m_{1,n}\}_{n=1}^\infty$. Proceeding in this fashion we obtain a countable collection of

subsequences of our original sequence:

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

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

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

$$
\Pr\left([-T_0, T_0]\right) \geq 1 - \frac{\varepsilon}{5 \cdot B} , \tag{2.34}
$$

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

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

implies

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

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

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

131

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

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

thus we have

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

and

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

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

$$\left| m_{n',n'}(y) - m_{n,n}(y) \right|$$

$$\leq \ \left| m_{n',n'}(y) - m_{n',n'}(\theta) \right| + \left| m_{n',n'}(\theta) - m_{n,n}(\theta) \right| + \left| m_{n,n}(\theta) - m_{n,n}(y) \right|$$

$$< \ \frac{3 \cdot \varepsilon}{5} \ .$$

$$(2.35)$$

Then

$$\left\| m_{n',n'} - m_{n,n} \right\|_{L^1(\mathbb{R}, P)}$$

$$= \ \int_\Theta \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y)$$

$$= \ \int_{[-T_1, T_1]} \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y) + \int_{\Theta \backslash [-T_1, T_1]} \left| m_{n',n'}(y) - m_{n,n}(y) \right| dP(y)$$

$$< \ \frac{3 \cdot \varepsilon}{5} \cdot \Pr([-T_1, T_1]) + \int_{\Theta \backslash [-T_1, T_1]} 2 \cdot B \cdot dP(y)$$

$$\leq \ \frac{3 \cdot \varepsilon}{5} \cdot 1 + \frac{\varepsilon}{5 \cdot B} \cdot 2 \cdot B$$

$$= \ \varepsilon \ ,$$

where the first inequality follows (2.35) and the second inequality comes from (2.34).

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

Proof of Proposition 2.4.

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

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

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

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

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

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

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

and

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

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

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

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

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

and

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

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

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

Hence for any $\varepsilon > 0$,

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

implies

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

for all $m \in M_{\overline{\beta}_i}$, i.e., $M_{\overline{\beta}_i}$ is equicontinuous. By definition, $\forall m \in M_{\overline{\beta}_i}$, $\forall \theta \in \Theta$, $|m\left(\theta\right)| \leq 1$, i.e., $M_{\overline{\beta}_i}$ is uniformly bounded. Therefore, according to Lemma 2.7, $M_{\overline{\beta}_i}$ is relatively compact in Ω.

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

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

With an optimal choice of s_j this gives

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

$$(2.36)$$

Note that $\lim_{\beta \to \infty} \beta^{1/2} \cdot f \left(\beta^{1/2} \left(x - \theta \right) \right) = \delta \left(x - \theta \right)$, where $\delta \left(\cdot \right)$ is the Dirac delta function. Then (2.36) implies

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

$$(2.37)$$

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

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

Note that our assumptions

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

and

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

imply that

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

and

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

respectively. Hence we have

$$\Pr\left(\theta - r \cdot \left(1 - m_i\left(\theta\right)\right) > 0\right)$$

$$\geq \ \Pr\left(\theta - r > 0\right)$$

$$> \ 0$$

and

$$\Pr\left(\theta - r \cdot \left(1 - m_i\left(\theta\right)\right) < 0\right)$$

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

$$> \ 0 \ .$$

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

$$\max\left\{0, \int_\theta \left[\theta - r \cdot \left(1 - m_i\left(\theta\right)\right)\right] \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}\left(x_j - \theta\right)\right) \cdot p\left(\theta\right) d\theta\right\}$$

$$\leq \ \int_\theta \max\left\{0, \left[\theta - r \cdot \left(1 - m_i\left(\theta\right)\right)\right] \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}\left(x_j - \theta\right)\right)\right\} \cdot p\left(\theta\right) d\theta$$

$$= \ \int_\theta \max\left\{0, \left[\theta - r \cdot \left(1 - m_i\left(\theta\right)\right)\right]\right\} \cdot \beta_j^{1/2} \cdot f\left(\beta_j^{1/2}\left(x_j - \theta\right)\right) \cdot p\left(\theta\right) d\theta \ . \quad (2.38)$$

Since

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

$$= \ \Pr\left(\theta - r \cdot \left(1 - m_i\left(\theta\right)\right) > 0\right)$$

$$> \ 0$$

and

$$\Pr\left([\theta - r \cdot (1 - m_i(\theta))] \cdot f\left(\beta_j^{1/2} (x_j - \theta) \right) < 0 \right)$$

$$= \Pr\left(\theta - r \cdot (1 - m_i(\theta)) < 0 \right)$$

$$> 0$$

for all $x_j \in \mathbb{R}$, (2.38) holds strictly. Then, (2.36) implies

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

where the last equality follows (2.37). Therefore,

$$\forall m_i \in \Omega, \ \forall \beta_j > 0, \ V_j^{**}(m_i) > V_j^*\left(\beta_j, m_i\right). \tag{2.39}$$

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

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

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

$$V_j^*\left(\beta_j', m_i\right) - V_j^*\left(\beta_j, m_i\right) > \delta\left(\beta_j, \beta_j'\right).$$

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

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

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

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

and

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

since $M_{\overline{\beta}_i}$ is relatively compact and $V_j^* (\beta, m)$ is a continuous functional of m for all $\beta > 0$. However, (2.39) implies that

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

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

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

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

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

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

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

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

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

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

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

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

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

and

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

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

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

139

and

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

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

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

and

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

Then

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

where the last inequality follows (2.39). Therefore, for $k \in \mathbb{N}$ large enough,

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

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

This concludes the proof. ∎

2.10   Appendix

Proof of Proposition 2.6.

Proof. According to Proposition 2.8 and 2.9 in Appendix 2.8, both "always invest" (i.e., $m(\theta) = 1$ almost surely) and "never invest" (i.e., $m(\theta) = 0$ almost surely) are equilibria. By the way, there may also exist "intermediate" equilibria with information acquisition. This concludes the proof. ∎

Proof of Proposition 2.7.

Proof. We consider a special class of equilibria with the following form

$$m(\theta) = \begin{cases} m_1 \text{ if } \theta \leq m_\lambda \\ m_0 \text{ if } \theta > m_\lambda \end{cases} ,$$

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

Given $p_I \in (0,1)$, (2.17) and (2.18) imply that

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

and

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

Let

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

If

$$g(p_I) = p_I ,$$

according to (2.17), (2.18) and (2.19), $m_1(p_I)$, $m_0(p_I)$,

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

and

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

141

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

Step 1. We show

$$g(p_I) < p_I$$

for p_I close to 1.

Since $m_\lambda \leq 1$,

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

Hence it suffices to show

$$\frac{1+A}{1+2A} \cdot m_1(p_I) + \frac{A}{1+2A} \cdot m_0(p_I) < p_I \tag{2.40}$$

for p_I close to 1. Let

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

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

and

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

then (2.40) can be rewritten as

$$\frac{1+A}{1+2A} \cdot \frac{w-1}{w+v_0} + \frac{A}{1+2A} \cdot \frac{w-1}{w+v_1} < \frac{w-1}{w} \; . \tag{2.41}$$

Since $p_I \in (0, 1)$, $w - 1 > 0$ and (2.41) is equivalent to

$$\frac{1+A}{1+2A} \cdot \frac{1}{1 + v_0/w} + \frac{A}{1+2A} \cdot \frac{1}{1 + v_1/w} < 1 \; . \tag{2.42}$$

142

It suffices to show (2.42) for w large enough.

Choosing w large enough such that

$$1 + v_0/w > 0$$

and multiplying both sides of (2.42) with $(1 + v_0/w)(1 + v_1/w)$ lead to

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

i.e.,

$$-\frac{v_0 v_1}{w} < \frac{A}{1+2A} \cdot v_1 + \frac{1+A}{1+2A} \cdot v_0 \; . \tag{2.43}$$

Note that (2.15) implies

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

Hence, we can choose w large enough such that (2.43) holds. Therefore, let

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

we have

$$g\left(\bar{p}_I\right) < \bar{p}_I \; .$$

Step 2. We show

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

for p_I close to 0.

Since $m_\lambda \geq 0$,

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

Hence it suffices to show

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

i.e.,

$$\frac{A}{1+2A} \cdot [1 - m_1(p_I)] + \frac{1+A}{1+2A} \cdot [1 - m_0(p_I)] < 1 - p_I \ , \tag{2.44}$$

for p_I close to 0. Let

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

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

and

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

then (2.44) can be rewritten as

$$\frac{A}{1+2A} \cdot \frac{w-1}{w+v_1} + \frac{1+A}{1+2A} \cdot \frac{w-1}{w+v_0} < \frac{w-1}{w} \ . \tag{2.45}$$

It suffices to show (2.45) for w large enough. Note that (2.45) is the same as (2.41), thus by the same argument in Step 1, we can find a $\underline{p}_I$ close to 0 such that

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

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

$$g(p_I^*) = p_I^* .$$

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

Proof of Proposition 2.5.

Proof. Given strategy profile (m_i, m_j), player i's expected utility is

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

Let

$$\Omega^* \triangleq \left\{ m \in \Omega : \left| m(\theta) - 1_{\{\theta > r/4\}} \right| \leq \frac{1}{4} \right\} .$$

We first show that if player j's strategy m_j belongs to Ω^*, so does player i's best response m_i. The trick is to show that $1_{\{\theta > r/4\}} \in \Omega^*$ strictly dominates any $\widetilde{m}_i \notin \Omega^*$. Note that for all $\theta \leq r/4$, we have

$$
\begin{aligned}
& \theta - r \cdot (1 - m_j(\theta)) \\
\leq\ & r/4 - r \cdot (1 - 1/4) \\
=\ & -r/2 .
\end{aligned}
\tag{2.46}
$$

For all $\theta > r/4$, we have

$$
\begin{aligned}
& \theta - r \cdot (1 - m_j(\theta)) \\
>\ & r/4 - r \cdot (1 - 3/4) \\
=\ & 0 .
\end{aligned}
\tag{2.47}
$$

Now we can calculate player i's expected gain of deviating from $1_{\{\theta>r/4\}}$ to $\widetilde{m}_i \notin \Omega^*$:

$$V_i\left(1_{\{\theta>r/4\}}, m_j\right) - V_i\left(\widetilde{m}_i, m_j\right)$$

$$= \int [\theta - r \cdot (1 - m_j(\theta))] \cdot \left[1_{\{\theta>r/4\}} - \widetilde{m}_i(\theta)\right] dP(\theta)$$

$$+ \mu \cdot \left[c\left(\widetilde{m}_i\right) - c\left(1_{\{\theta>r/4\}}\right)\right]$$

$$\geq \int [\theta - r \cdot (1 - m_j(\theta))] \cdot \left[1_{\{\theta>r/4\}} - \widetilde{m}_i(\theta)\right] dP(\theta) - \mu \cdot K$$

$$= \int_{\theta \leq r/4} [\theta - r \cdot (1 - m_j(\theta))] \cdot [0 - \widetilde{m}_i(\theta)] dP(\theta)$$

$$+ \int_{\theta > r/4} [\theta - r \cdot (1 - m_j(\theta))] \cdot [1 - \widetilde{m}_i(\theta)] dP(\theta) - \mu \cdot K$$

$$\geq \int_{\theta \leq r/4} (-r/2) \cdot (-1/4) \, dP(\theta) + \int_{\theta > r/4} 0 \cdot [1 - \widetilde{m}_i(\theta)] \, dP(\theta) - \mu \cdot K$$

$$= \frac{r}{8} \cdot \Pr\left(\theta \leq r/4\right) - \mu \cdot K$$

$$> 0 . \tag{2.48}$$

Here the first inequality holds since $c(\cdot)$ is non-negative and uniformly bounded above by K, the second inequality follows (2.46) and (2.47), and the last inequality comes from condition (2.13) of this proposition. Inequality (2.48) implies that player i's best response to any $m_j \in \Omega^*$ must also belong to Ω^*. Now consider a "smaller" game with the same payoffs but a smaller strategy space Ω^*. Since Ω^* is a complete lattice under the natural order and $c(\cdot)$ is assumed to be submodular, this new game is a supermodular game and must have an equilibrium. Let $(\overline{m}_i, \overline{m}_j)$ be such an equilibrium. Then $(\overline{m}_i, \overline{m}_j)$ is also an equilibrium of the original game, since Ω^* is closed under the best response dynamics.

Similarly, we can construct

$$\Omega_* \triangleq \left\{m \in \Omega : \left|m(\theta) - 1_{\{\theta \geq 3r/4\}}\right| \leq \frac{1}{4}\right\}$$

and find an equilibrium

$$\left(\underline{m}_i, \underline{m}_j\right) \in \Omega_* \times \Omega_* \ .$$

Since

$$\Omega_* \cap \Omega^* = \phi \ ,$$

$(\overline{m}_i, \overline{m}_j)$ and $\left(\underline{m}_i, \underline{m}_j\right)$ are two different equilibria of the game. This completes the proof. ∎

Chapter 3

Linear Best-response Games with Flexible Information Acquisition

3.1 Introduction

This chapter examines flexible information acquisition in the linear best-response games. This class of games have been widely employed in modeling nominal price adjustment, asset pricing and financial crisis, etc.. In most of these models, information structures are exogenously given and the agents passively respond to their endowed signals (e.g., Morris and Shin (2002), Angeletos and Pavan (2007)). Since people always make decisions according to their information at hand, it makes sense to investigate what information people acquire and how it affects the equilibrium outcomes. The current model takes this into account by allowing people to collect information according to their own interests and the qualities of their information sources.

The economy consists of a continuum of agents. Each agent enjoys a utility quadratic in her own action, the economy's aggregate/average action and a fluctuating fundamental. The agents do not know the exact value of the fundamental

but share a common prior. Each agent has access to a public information source and her private information source. She can flexibly acquire information from her information sources in the manner of rational inattention[1] subject to her capacity constraint. This setup of preference is similar to Angeletos and Pavan (2007). However, in order to get general results, we use an extended version that accommodates heterogeneous preferences, information sources and information processing capacities. In this general setup, we examine the role played by information acquisition.

Information acquisition plays a crucial role in shaping the equilibrium outcomes. Firstly, the agents face a trade-off between two objectives. One is to match the fundamental and the other is to match the aggregate/average action of the economy. The second objective induces a strategic complementarity among agents' actions, which leads the agents to pay more attention to the public information source and be less attentive to the private ones. In other words, the strategic complementarity among actions makes agents' information choices also strategic complements. Secondly, it is worth noting that if there is no capacity constraint, our game essentially reduces to the game with exogenous information structures where agents just treat their information sources as the signals in Morris and Shin (2002), Angeletos and Pavan (2007). Introducing capacity constraints directly dampens the agents' responses to the information content conveyed by their information sources, since their responses now contain some information processing errors. In other words, the capacity constraints weaken the agents' ability to employ information. Moreover, rationally expecting others' dampened (and thus less effective) responses, an agent finds it less attractive to match the economy's aggregate/average action, since it makes no sense to match others' information processing errors. Upon these two effects, we can establish an equivalence between games with and without capacity constraints, which is the main result of this chapter.

---

[1]Chapter 1 introduces rational inattention and the related literature.

The most related paper is Hellwig and Veldkamp (2009). They introduce endogenous information acquisition into a version of Morris and Shin (2002) static beauty contest game where the agents can buy information through choosing from a set of jointly normal distributed signals. Each of these signals is a linear combination of the fundamentals plus some public and/or private noises, and differs from each other in the precision. Both Hellwig and Veldkamp (2009) and our model show that the strategic externality among agents' actions induces the same type of strategic externalities among agents' information choices. However, in Hellwig and Veldkamp (2009) adding information choices to the game with strategic complementarity leads to multiple equilibria even though the corresponding game with exogenous information has a unique equilibrium. As a comparison, our model always predicts a unique equilibrium. This uniqueness comes from the continuous nature of rational inattention as a modeling technique, while the multiplicity in Hellwig and Veldkamp (2009) comes from the agents' essentially discrete (binary) decision of whether to buy the signals that have already been purchased by others.

This chapter proceeds as following. Section 3.2 sets up the model. In Section 3.3, we explore the basic properties of the equilibria. Section 3.4 derives the agents' information acquisition and establishes the equivalence theorems. It also characterizes the equilibria. Section 3.5 concludes the chapter. Most proofs are relegated to the appendix.

3.2 The Model

The economy consists of a continuum of agents indexed by $i \in [0, 1]$. Agent i enjoys a utility

$$u_i = U^i \left(k_i, K, \sigma, \overrightarrow{\theta} \right) \ ,$$

which is quadratic in its arguments. Here $k_i \in \mathbb{R}$ is agent i's action and

$$K = \int_0^1 k_j \cdot dj$$

is the aggregate/average action of the economy. The fundamental

$$\overrightarrow{\theta} \sim N\left(\overrightarrow{0}, \Sigma_{\theta\theta}\right)$$

is an exogenous n_θ-dimension payoff-relevant vector [2] and

$$\sigma^2 = \int_0^1 (k_j - K)^2 \cdot dj$$

denotes the dispersion of the agents' actions. Following the setup of Angeletos and Pavan (2007), we assume each agent i's utility function to satisfy

$$U_{k\sigma}^i = U_{K\sigma}^i = U_{\theta\sigma}^i = 0$$

and

$$U_\sigma^i\left(k, K, 0, \overrightarrow{\theta}\right) = 0$$

for all $\left(k, K, \overrightarrow{\theta}\right)$. The last equation means that dispersion has only a second-order and non-strategic external effect. We also assume $U_{kk}^i < 0$ to guarantee a well defined best response and matrix $U_{k\theta}^i$ non-singular to exclude the trivial case where some entries of fundamental $\overrightarrow{\theta}$ are irrelevant for equilibrium behavior. Note that $r_i \triangleq -U_{kK}^i / U_{kk}^i$ measures the strategic externality between agent i's action and the aggregate action. As in Angeletos and Pavan (2007), to guarantee the uniqueness

---

[2] The assumption of zero-mean here is not essential since we can always redefine the fundamentals by subtracting its mean.

of the equilibrium, we assume

$$\int |r_i| \cdot di < 1 \ .$$

Here we allow the utility function $U^i(\cdot)$ to vary over $i \in [0,1]$ to represent the possible heterogeneity in the agents' preferences.

Information Sources. An information source is a random vector correlated with the fundamental and thus contains payoff relevant information. All the agents have access to a public information source $\overrightarrow{f} \in \mathbb{R}^{n_f}$, which is jointly normal distributed with the fundamental $\overrightarrow{\theta}$, i.e.

$$\begin{pmatrix} \overrightarrow{\theta} \\ \overrightarrow{f} \end{pmatrix} \sim N \left(\overrightarrow{0}, \begin{bmatrix} \Sigma_{\theta\theta} & \Sigma_{\theta f} \\ \Sigma_{f\theta} & \Sigma_{ff} \end{bmatrix} \right).$$

Besides the public information source $\overrightarrow{f}$, each agent i also has a private information source $\overrightarrow{x}_i \in \mathbb{R}^{n_{x_i}}$ satisfying

$$\begin{pmatrix} \overrightarrow{\theta} \\ \overrightarrow{x}_i \end{pmatrix} \sim N \left(\overrightarrow{0}, \begin{bmatrix} \Sigma_{\theta\theta} & \Sigma_{\theta x_i} \\ \Sigma_{x_i\theta} & \Sigma_{x_i x_i} \end{bmatrix} \right).$$

Information source $\overrightarrow{x}_i$ is private in the sense that only agent i has access to it. We assume that $\forall i, j \in [0,1]$, $i \neq j$, $\overrightarrow{x}_i$, $\overrightarrow{x}_j$ and $\overrightarrow{f}$ are conditionally independent with respect to the fundamental $\overrightarrow{\theta}$. Then the joint distribution of $\left(\overrightarrow{\theta}, \overrightarrow{f}, \overrightarrow{x}_i \right)$ becomes

$$\begin{pmatrix} \overrightarrow{\theta} \\ \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix} \sim N \left(\overrightarrow{0}, \begin{bmatrix} \Sigma_{\theta\theta} & \Sigma_{\theta f} & \Sigma_{\theta x_i} \\ \Sigma_{f\theta} & \Sigma_{ff} & \Sigma_{f x_i} \\ \Sigma_{x_i\theta} & \Sigma_{x_i f} & \Sigma_{x_i x_i} \end{bmatrix} \right),$$

where

$$\Sigma_{fx_i} \triangleq Cov\left(\overrightarrow{f}, \overrightarrow{x}'_i\right) = \Sigma_{f\theta}\Sigma_{\theta\theta}^{-1}\Sigma_{\theta x_i} \ .$$

Here $\Sigma_{\theta x_i}$ and Σ_{fx_i} vary over $i \in [0,1]$ to represents the heterogeneity in agents' information sources.

Information Acquisition. Agent i's information acquisition is modeled as choosing a signal $\overrightarrow{s}_i$ correlated with her information sources $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$. To make the model tractable, signal $\overrightarrow{s}_i$ is also assumed to be jointly normal distributed with agent i's information sources. Agent i can flexibly acquire different pieces of information through varying the covariance matrix between $\overrightarrow{s}_i$ and $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$, but her capacity constraint requires that the mutual information between $\overrightarrow{s}_i$ and $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$ can not exceed some upper bound $\kappa_i \geq 0$. Hence, agent i has to rationally allocate her attention to acquire the information most relevant to her welfare. Intuitively, she should pay more attention to the sources with higher quality (i.e. higher correlation with the fundamentals) and rationally ignore others. Moreover, if her action and the aggregate action are strategic complements (i.e., $r_i > 0$), she should pay more attention to the information that others pay more attention to. In other words, there is also some complementarity in information acquisition. Since $\overrightarrow{s}_i$ and $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$ are jointly Gaussian, $\overrightarrow{s}_i$ consists of some linear combinations of $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$ and an information processing error independent with $\begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix}$. It is natural to assume that each agent's information processing error is independent with the fundamental, all public and private information sources as well as all others' information processing errors. This assumption reflects the private nature of agents' information acquisition.

Simplification of Information Acquisition. Without loss of generality, we can always standardize $\vec{s}_i$ to $N\left(\vec{0}, I\right)$. Thus agent i's information choice is characterized by the covariance matrix between $\vec{s}_i$ and $\left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)$. Moreover, since

$$Var\left(\left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)\right) = \left[\begin{array}{cc} \Sigma_{ff} & \Sigma_{fx_i} \\ \Sigma_{x_if} & \Sigma_{x_ix_i} \end{array}\right]$$

is a symmetric real matrix, there exists some orthogonal matrix P_i such that

$$Var\left(\left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)\right) = P_i'\Lambda_i P_i \ ,$$

where $\Lambda_i = diag\left(\lambda_{i,1}, \lambda_{i,2}, \cdots, \lambda_{i,n_f+n_{x_i}}\right)$. Let $\vec{z}_i = \Lambda_i^{-1/2} P_i \left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)$, then $\vec{z}_i$ $\sim N\left(\vec{0}, I\right)$. Note that $\vec{z}_i$ and $\left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)$ have the same information content because one is just an invertible transform of the other. Therefore, $I\left(\vec{s}_i; \vec{z}_i\right) = I\left(\vec{s}_i; \left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)\right)$. Later, we will see that this change-of-variable simplifies our derivation since we can just focus on $\Sigma_{z_is_i} \triangleq Cov\left(\vec{z}_i, \vec{s}_i\right)$ rather than the covariance matrix between $\left(\begin{array}{c} \vec{f} \\ \vec{x}_i \end{array}\right)$ and $\vec{s}_i$. The joint distribution of $\vec{z}_i$ and $\vec{s}_i$ now becomes

$$\left(\begin{array}{c} \vec{z}_i \\ \vec{s}_i \end{array}\right) \sim N\left(\vec{0}, \left[\begin{array}{cc} I & \Sigma_{z_is_i} \\ \Sigma_{s_iz_i} & I \end{array}\right]\right),$$

and the capacity constraint becomes

$$I\left(\overrightarrow{s}_i; \overrightarrow{z}_i\right) = -\frac{1}{2} \cdot \ln\left(\det\left(I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}\right)\right) \leq \kappa_i ,$$

i.e.,

$$\det\left(I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}\right) \geq \exp\left(-2 \cdot \kappa_i\right) .$$

In the current model, we call the resulted joint distribution of $\left(\overrightarrow{\theta}, \{\overrightarrow{s}_i\}_{i \in [0,1]}\right)$ the information structure, which is endogenously determined by the agents' information choices $\{\Sigma_{z_i s_i}\}_{i \in [0,1]}$. It is worth noting that in the extreme case where capacity $\kappa_i = \infty$, the capacity constraint disappears and the agent can directly observe her information sources. Hence the problem reduces to the classical one with exogenous information structure, which is a special case of the current model.

Timing of the game. First, each agent i simultaneously chooses $\Sigma_{z_i s_i}$ subject to her capacity constraint, i.e.,

$$\det\left(I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}\right) \geq \exp\left(-2 \cdot \kappa_i\right) ,$$

where $I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}$ should be positive semi-definite[3]; second, each agent i's private signal $\overrightarrow{s}_i$ realizes; third, each agent i takes action k_i to maximize her expected utility according to the realization of $\overrightarrow{s}_i$.

This game is solved through backward induction. We first analyze the agents' actions at the third stage. The first order condition implies agent i's unique best response

$$k_i = E_i\left(r_i \cdot K + \overrightarrow{\eta}_i' \overrightarrow{\theta}\right) + \delta_i, \tag{3.1}$$

---

[3]Note that $I - \Sigma_{z_i s_i} \Sigma_{s_i z_i} = Var\left(\overrightarrow{z}_i | \overrightarrow{s}_i\right)$ is positive semi-definite.

where

$$E_i\left(\cdot\right) \triangleq E\left(\cdot \,|\, \overrightarrow{s}_i\right), r_i = -\frac{U^i_{kK}}{U^i_{kk}}, \overrightarrow{\eta}_i = -\frac{U^i_{k\theta}}{U^i_{kk}} \in \mathbb{R}^{n_\theta}, \delta_i = -\frac{U^i_k\left(0,0,0,0\right)}{U^i_{kk}}.$$

We omit the derivation here since it is almost the same as Proposition 1 of Angeletos and Pavan (2007). Here we also normalize δ_i to zero for all agents, since this term just changes the level of best responses but has on strategic effect.

Now agent i's utility function can be rewritten as

$$
\begin{aligned}
U^i\left(k_i, K, \sigma, \overrightarrow{\theta}\right) &= \frac{1}{2} \cdot U^i_{kk} \cdot \left[k_i - \left(r_i \cdot K + \overrightarrow{\eta}'_i \overrightarrow{\theta}\right)\right]^2 + G^i\left(K, \sigma, \overrightarrow{\theta}\right) \\
&\triangleq \frac{1}{2} \cdot \left(-U^i_{kk}\right) \cdot \widetilde{U}^i\left(k_i, K, \overrightarrow{\theta}\right) + G^i\left(K, \sigma, \overrightarrow{\theta}\right),
\end{aligned}
$$

where $G^i\left(K, \sigma, \overrightarrow{\theta}\right)$ is quadratic in its arguments. Since $G^i\left(K, \sigma, \overrightarrow{\theta}\right)$ does not affect any individual agent's best response, $\left\{\widetilde{U}^i\right\}_{i\in[0,1]}$ (i.e., the parameters $\{r_i, \overrightarrow{\eta}_i\}_{i\in[0,1]}$) suffices to describe the agents' preferences. Therefore, we can just focus on $\{r_i, \overrightarrow{\eta}_i\}_{i\in[0,1]}$ when analyzing the equilibria in the rest of the paper.

Let $\overrightarrow{w} = \left(\overrightarrow{\theta}', \overrightarrow{f}'\right)' \in \mathbb{R}^{n_w}$ and $\overrightarrow{b}_i = \left(\overrightarrow{\eta}'_i, \overrightarrow{0}'\right)' \in \mathbb{R}^{n_w}$, where $n_w = n_\theta + n_f$. Then Equation (3.1) can be rewritten as

$$k_i = E_i\left(r_i \cdot K + \overrightarrow{b}'_i \overrightarrow{w}\right). \tag{3.2}$$

Taking integral for both sides of the above equation leads to

$$K = \widetilde{E}\left(K\right) + \widehat{E}\left(\overrightarrow{w}\right), \tag{3.3}$$

where

$$\widetilde{E}\left(\cdot\right) \triangleq \int_0^1 r_i \cdot E_i\left(\cdot\right) \cdot d_i$$

is an average expectation operator, a linear mapping from the space of random

variables to itself, and

$$\widehat{E}\left(\cdot\right) \triangleq \int_0^1 \overrightarrow{b}'_i E_i\left(\cdot\right) \cdot d_i$$

is another average expectation operator that is a linear mapping from the space of n_w-dimensional random vectors to the space of random variables. We call $\overrightarrow{w} = \left(\overrightarrow{\theta}', \overrightarrow{f}'\right)'$ the state (random vector) of the economy.

Note that any equilibrium aggregate action K should satisfy Equation (3.3). Once K is given, the individual action k_i is determined by Equation (3.2). Hence, we first focus on Equation (3.3). Since $\overrightarrow{s}_i$ and $\overrightarrow{w}$ are jointly normal distributed, $\overrightarrow{s}_i$ can be expressed as

$$\overrightarrow{s}_i = \Sigma_i \overrightarrow{w} + \overrightarrow{\mu}_i + \overrightarrow{\epsilon}_i,$$

where $\Sigma_i \in \mathbb{R}^{n_{s_i} \times n_w}$, $\overrightarrow{\mu}_i \in \mathbb{R}^{n_{s_i}}$, and $\overrightarrow{\epsilon}_i \in \mathbb{R}^{n_{s_i}}$ is an idiosyncratic Gaussian noise. Actually, $\overrightarrow{\epsilon}_i$ consists of agent i's information processing error and the noise in her private information source. By the definition of $\widetilde{E}\left(\cdot\right)$, for any random variable K, $\widetilde{E}\left(K\right)$ is a function of $\overrightarrow{w}$ and $\{\overrightarrow{\epsilon}_i\}_{i \in [0,1]}$. However, since each single agent only has a zero measure, $\widetilde{E}\left(K\right)$ does not depend on any specific $\overrightarrow{\epsilon}_i$. This implies that $\widetilde{E}\left(K\right)$ is a function of the state $\overrightarrow{w}$. Similarly, $\widehat{E}\left(\overrightarrow{w}\right)$ is also a function of $\overrightarrow{w}$. Thus Equation (3.3) implies that the equilibrium aggregate action K must be a function of the state $\overrightarrow{w}$.

The equilibrium is characterized by the aggregate action $K\left(\overrightarrow{w}\right)$, the agents' information choices $\{\Sigma_{z_i s_i}\}_{i \in [0,1]}$ and their actions $\{k_i\left(\overrightarrow{s}_i\right)\}_{i \in [0,1]}$.

Definition: $\left(K\left(\overrightarrow{w}\right), \{\Sigma_{z_i s_i}\}_{i \in [0,1]}, \{k_i\left(\overrightarrow{s}_i\right)\}_{i \in [0,1]}\right)$ *is an equilibrium iff 1)* $\forall i \in [0,1]$,

$$k_i\left(\overrightarrow{s}_i\right) = E_i\left(r_i \cdot K + \overrightarrow{b}'_i \overrightarrow{w}\right) ;$$

2)

$$K\left(\overrightarrow{w}\right) = \widetilde{E}\left(K\left(\overrightarrow{w}\right)\right) + \widehat{E}\left(\overrightarrow{w}\right) ;$$

3) $\forall i \in [0, 1]$,

$$\Sigma_{z_i s_i} = \arg \max_{\substack{\det\left(I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}\right) \geq \exp(-2 \cdot \kappa_i) \\ I - \Sigma_{z_i s_i} \Sigma_{s_i z_i} \text{ positive semi-definite}}} E\left[\widetilde{U^i}\left(k_i\left(\overrightarrow{s}_i\right), K\left(\overrightarrow{w}\right), \overrightarrow{\theta}\right)\right] . \qquad (3.4)$$

In our equilibrium, the functional form $K\left(\cdot\right)$ is common knowledge. Each agent cares about the state $\overrightarrow{w}$ and the implied aggregate action $K\left(\overrightarrow{w}\right)$ rather than others' utility functions and strategies. If the utility parameters $\{r_i, \overrightarrow{\eta}_i\}_{i \in [0,1]}$ are common knowledge, we can also define the Bayesian Nash equilibrium. These two definitions generate the same equilibrium strategies. We use the first one in the rest of this paper since the assumptions are weaker (since it does not require $\{r_i, \overrightarrow{\eta}_i\}_{i \in [0,1]}$ to be common knowledge).

3.3 Basic Properties of the Equilibrium Aggregate Action $K\left(\overrightarrow{w}\right)$

The definition shows that any equilibrium aggregate action $K\left(\overrightarrow{w}\right)$ is a solution of functional equation (3.3). To proceed, we first study operator $\widetilde{E}\left(\cdot\right)$. Let $\mathbf{P}$ be the probability space on which the random vectors $\overrightarrow{\theta}$, $\overrightarrow{f}$ and $\{\overrightarrow{x}_i\}_{i \in [0,1]}$ are defined. Let Ω be the space of all random variables defined on $\mathbf{P}$. Hence, Ω is a vector space and by definition $\widetilde{E}\left(\cdot\right)$ is a linear mapping from Ω to itself. For any pair of random variables u and v in Ω, define their inner product as $\langle u, v \rangle = E\left[u \cdot v\right]$. Hence the induced norm of v is $||v||_{rv} = \left[E\left(v^2\right)\right]^{1/2}$, where the subscript "$rv$" stands for "random variable". Then the norm of the operator $\widetilde{E}\left(\cdot\right)$ can be defined as $||\widetilde{E}\left(\cdot\right)||_{op} = \sup_{||v||_{rv}=1} ||\widetilde{E}\left(v\right)||_{rv} = \sup_{||v||_{rv}=1} |\langle \widetilde{E}\left(v\right), v \rangle|$, where the subscript "$op$" stands for "operator". Generally speaking, the property of $\widetilde{E}\left(\cdot\right)$ and the resulted solution of Equation (3.3) depend on the underlying information

structure, i.e., the agents' information choices $\{\Sigma_{z_i s_i}\}_{i \in [0,1]}$. We first summarize some important properties of Equation (3.3) that are independent of the information choices in the following proposition.

Proposition 3.1 *Equation (3.3) has a unique solution, which is linear in $\overrightarrow{w}$, i.e.,*
$\exists \overrightarrow{u} \in \mathbb{R}^{n_w}$, *s.t.* $K(\overrightarrow{w}) = \overrightarrow{u}'\overrightarrow{w}$.

Proof. For any $v \in \Omega$, s.t. $||v||_{rv} = 1$, we have

$$
\begin{aligned}
E\left[E_i(v)\right]^2 &= Var\left[E_i(v)\right] + (Ev)^2 \\
&= Var(v) - E\left[Var_i(v)\right] + (Ev)^2 \\
&= E(v^2) - E\left[Var_i(v)\right] \\
&\leq E(v^2) = ||v||_{rv}^2 = 1 \ ,
\end{aligned}
$$

where the second equality comes form the law of total variance. Thus

$$
\begin{aligned}
||E_i(\cdot)||_{op} &= \sup_{||v||_{rv}=1} ||E_i(v)||_{rv} \\
&= \sup_{||v||_{rv}=1} \left(E\left[E_i(v)\right]^2\right)^{1/2} \\
&\leq 1
\end{aligned}
$$

and

$$
\begin{aligned}
||\widetilde{E}(\cdot)||_{op} &= ||\int_0^1 r_i \cdot E_i(\cdot) \cdot di||_{op} \\
&\leq \int_0^1 |r_i| \cdot ||E_i(\cdot)||_{op} \cdot di \\
&\leq \int_0^1 |r_i| \cdot di < 1 \ .
\end{aligned}
$$

Therefore the operator $\left(I - \widetilde{E}\right)(\cdot)$ is invertible, and the conclusion directly follows Lemma 3.1 in Appendix 3.6. ■

According to this proposition, there exists a unique equilibrium such that the aggregate action $K(\overrightarrow{w})$ is linear in $\overrightarrow{w}$ regardless of the agents' information choices. Since $K(\cdot)$ is continuous in its argument, the measures induced by $K(\overrightarrow{w})$ and $\overrightarrow{w}$ are mutually continuous. In other words, $K(\overrightarrow{w})$ is informationally efficient in the sense that any non-zero probability event about the state $\overrightarrow{w}$ corresponds to a non-zero probability event about the aggregate action $K(\overrightarrow{w})$.

3.4   Equilibrium with Flexible Information Acquisition

For any $i \in [0,1]$, let

$$\beta_i = \left[1 - \exp\left(-2 \cdot \kappa_i\right)\right]^{1/2}$$

where κ_i is agent i's capacity of information acquisition. Then, given the set of information sources, the game is characterized by $\{r_i, \overrightarrow{\eta}_i, \beta_i\}_{i \in [0,1]}$.

3.4.1   Strategic Incentives in Information Choices

This subsection examines the agents' incentives in acquiring information. In order to make the expressions simple, we focus on the case of homogeneous preference, i.e., $\forall i \in [0,1]$, $(r_i, \overrightarrow{\eta}_i, \delta_i) = (r, \overrightarrow{\eta}, \delta)$. The following argument actually restates the main result of Hellwig and Veldkamp (2009) in our context.

Let $\{\Sigma_{1,z_i s_i}\}_{i \in [0,1]}$ and $\{\Sigma_{2,z_i s_i}\}_{i \in [0,1]}$ be two profiles of information choices and $\{\overrightarrow{s}_{1,i}\}_{i \in [0,1]}$ and $\{\overrightarrow{s}_{2,i}\}_{i \in [0,1]}$ be the two sets of resulted signals, respectively. Let

$$\overline{Var_l}\left(\overrightarrow{w}\right) = \int_0^1 Var\left(\overrightarrow{w} \,|\, \overrightarrow{s}_{l,j}\right) \cdot dj, l = 1, 2 \ .$$

Let

$$EU^i\left(\Sigma_{l,z_is_i}, \overline{Var}_{l'}\left(\overrightarrow{w}\right)\right)$$

denote agent i's expected utility when her information choice is Σ_{l,z_is_i} and others' are

$$\left\{\Sigma_{l',z_js_j}\right\}_{j\in[0,1]\backslash\{i\}}, l, l' = 1, 2 .$$

Proposition 3.2 *Suppose* $\overline{Var}_2\left(\overrightarrow{w}\right) - \overline{Var}_1\left(\overrightarrow{w}\right)$ *is positive semi-definite and* $Var\left(\overrightarrow{w}|\overrightarrow{s}_{2,i}\right) - Var\left(\overrightarrow{w}|\overrightarrow{s}_{1,i}\right)$ *is positive semi-definite for some* $i \in [0,1]$. *Then*

1) if there is no strategic externality (i.e. $r = 0$*), the value of additional information is independent of others' information choices, i.e.,*

$$EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right)$$
$$= EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) ;$$

2) if there is strategic complementarity (i.e. $r > 0$*), the value of additional information is increasing in others' information acquisition, i.e.,*

$$EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right)$$
$$> EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) ;$$

3) if there is strategic substitutability (i.e. $r < 0$*), the value of additional information is decreasing in others' information acquisition, i.e.,*

$$EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_1\left(\overrightarrow{w}\right)\right)$$
$$< EU^i\left(\Sigma_{1,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) - EU^i\left(\Sigma_{2,z_is_i}, \overline{Var}_2\left(\overrightarrow{w}\right)\right) .$$

The proof is omitted here since it is almost the same as that in Hellwig and Veldkamp (2009).

Remark 3.1 *A positive semi-definite $Var\left(\overrightarrow{w}|\overrightarrow{s}_{2,i}\right) - Var\left(\overrightarrow{w}|\overrightarrow{s}_{1,i}\right)$ means that agent i's information choice $\Sigma_{1,z_i s_i}$ is more informative than $\Sigma_{2,z_i s_i}$. Similarly, a positive semi-definite $\overline{Var}_2\left(\overrightarrow{w}\right) - \overline{Var}_1\left(\overrightarrow{w}\right)$ means that the information choice profile $\{\Sigma_{1,z_i s_i}\}_{i\in[0,1]}$ generates more aggregate information than $\{\Sigma_{2,z_i s_i}\}_{i\in[0,1]}$. When actions are complementary, so is the information acquisition, regardless of whether acquiring public or private information. For more detailed discussion, please refer to Proposition 1 of Hellwig and Veldkamp (2009).*

3.4.2 Optimal Information Acquisition

This game is essentially static, but it might be better understood in a pseudo-dynamic way. Let $K_m\left(\overrightarrow{w}\right)$ denote the aggregate action in period m. Suppose all agents are aware of it and naively believe that the functional form $K_m\left(\cdot\right)$ remains the same in period $m+1$. Then in period $m+1$ each agent i chooses her optimal channel of information acquisition $\Sigma_{z_i s_i}$ according to this $K_m\left(\cdot\right)$. Once the information choice profile $\{\Sigma_{z_i s_i}\}_{i\in[0,1]}$ is determined, the two average expectation operators $\widetilde{E}\left(\cdot\right)$ and $\widehat{E}\left(\cdot\right)$ are determined and hence Equation (3.3) generates $K_{m+1}\left(\overrightarrow{w}\right)$, the "real" aggregate action in period $m+1$. Note that the optimal information choice profile $\{\Sigma_{z_i s_i}\}_{i\in[0,1]}$ is changing as the common belief of aggregate action $K_m\left(\cdot\right)$ changes, and so do the two average expectation operators. Therefore, the information structure of this pseudo-dynamic game endogenously evolves. The equilibrium defined in Section 3.2 is achieved if $K_{m+1}\left(\cdot\right) = K_m\left(\cdot\right) \triangleq K\left(\cdot\right)$ is a fixed point of the above process.

According to Proposition 3.1, the aggregate action in period m is linear in $\overrightarrow{w}$, i.e.,

$$K_m\left(\overrightarrow{w}\right) = \overrightarrow{u}'_m \overrightarrow{w}$$

for some $\overrightarrow{u}_m \in \mathbb{R}^{n_w}$. Let $\Sigma_{ww} = Var\left(\overrightarrow{w}\right)$. Since Σ_{ww} is real and symmetric, there

always exists some orthogonal matrix $\widetilde{P}$ such that

$$\Sigma_{ww} = \widetilde{P}'\widetilde{\Lambda}\widetilde{P} \ ,$$

where $\widetilde{\Lambda} = diag\left(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \cdots, \widetilde{\lambda}_{n_w}\right)$. Let

$$\overrightarrow{\xi} = \widetilde{\Lambda}^{-1/2}\widetilde{P}\overrightarrow{w} \ ,$$

then $\overrightarrow{\xi} \sim N\left(\overrightarrow{0}, I\right)$ also represents the state. To facilitate the derivations, we change the variables by letting

$$\overrightarrow{c}_i = \widetilde{\Lambda}^{1/2}\widetilde{P}\overrightarrow{b}_i$$

and

$$\overrightarrow{t}_m = \widetilde{\Lambda}^{1/2}\widetilde{P}\overrightarrow{u}_m$$

for all $i \in [0, 1]$. Then Equation (3.3) can be rewritten as

$$K_{m+1}\left(\overrightarrow{\xi}\right) = \widetilde{E}\left(K_m\left(\overrightarrow{\xi}\right)\right) + \widehat{E}\left(\overrightarrow{\xi}\right) \ ,$$

where

$$K_m\left(\overrightarrow{\xi}\right) = \overrightarrow{t}'_m\overrightarrow{\xi} \ ,$$

$$\widetilde{E}\left(\cdot\right) \triangleq \int_0^1 r_i \cdot E_i\left(\cdot\right) \cdot d_i$$

and

$$\widehat{E}\left(\cdot\right) \triangleq \int_0^1 \overrightarrow{c}'_i E_i\left(\cdot\right) \cdot d_i \ .$$

Now we are ready to derive agent i's expected utility for any given information choice $\Sigma_{z_i s_i}$.

By Equation (3.2), agent i's best response conditional on his signal $\overrightarrow{s}_i$ is

$$k_i\left(\overrightarrow{s}_i\right) = E_i\left(r_i \cdot K_m\left(\overrightarrow{\xi}\right) + \overrightarrow{c}_i'\overrightarrow{\xi}\right), \tag{3.5}$$

thus her expected utility becomes

$$
\begin{aligned}
& E\left[\widetilde{U^i}\left(k_i\left(\overrightarrow{s}_i\right), K, \overrightarrow{\theta}\right)\right] \\
=\; & -E\left[k_i\left(\overrightarrow{s}_i\right) - \left(r_i \cdot K_m\left(\overrightarrow{\xi}\right) + \overrightarrow{c}_i'\overrightarrow{\xi}\right)\right]^2 \\
=\; & -Var_i\left(r_i \cdot \overrightarrow{t}_m'\overrightarrow{\xi} + \overrightarrow{c}_i'\overrightarrow{\xi}\right) \\
=\; & -\left(r_i \cdot \overrightarrow{t}_m + \overrightarrow{c}_i\right)'Var_i\left(\overrightarrow{\xi}\right)\left(r_i \cdot \overrightarrow{t}_m + \overrightarrow{c}_i\right) \\
\triangleq\; & -\overrightarrow{h}_{m,i}'Var_i\left(\overrightarrow{\xi}\right)\overrightarrow{h}_{m,i}. \tag{3.6}
\end{aligned}
$$

In order to calculate $Var_i\left(\overrightarrow{\xi}\right)$, we first calculate $\Sigma_{s_i\xi}$, i.e.,

$$
\begin{aligned}
\Sigma_{s_i\xi} \;\triangleq\; & Cov\left(\overrightarrow{s}_i, \overrightarrow{\xi}'\right) \\
=\; & Cov\left(\Sigma_{s_iz_i}\Lambda_i^{-1/2}P_i\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}, \left(\overrightarrow{\theta}' \quad \overrightarrow{f}'\right)\widetilde{P}'\widetilde{\Lambda}^{-1/2}\right) \\
=\; & \Sigma_{s_iz_i}\Lambda_i^{-1/2}P_i\begin{pmatrix}\Sigma_{f\theta} & \Sigma_{ff}\\\Sigma_{x_i\theta} & \Sigma_{x_if}\end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}, \tag{3.7}
\end{aligned}
$$

where the second equality holds since agent i's information processing error is independent of the state. Hence,

$$
\begin{aligned}
Var_i\left(\overrightarrow{\xi}\right) \;=\; & I - \Sigma_{\xi s_i}\Sigma_{s_i\xi} \\
=\; & I - \widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P_i'\Lambda_i^{-1/2}\Sigma_{z_is_i}\Sigma_{s_iz_i}\Lambda_i^{-1/2}P_i\begin{pmatrix}\Sigma_{f\theta} & \Sigma_{ff}\\\Sigma_{x_i\theta} & \Sigma_{x_if}\end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}.
\end{aligned}
$$

$$\tag{3.8}$$

Substituting (3.8) into (3.6) leads

$$
\begin{aligned}
& E\left[\widetilde{U^i}\left(k_i, K, \overrightarrow{\theta}\right)\right] \\
= & -\overrightarrow{h}'_{m,i} Var_i\left(\overrightarrow{\xi}\right) \overrightarrow{h}_{m,i} \\
= & -\overrightarrow{h}'_{m,i} \overrightarrow{h}_{m,i} \\
& + \overrightarrow{h}'_{m,i} \widetilde{\Lambda}^{-1/2} \widetilde{P}
\begin{pmatrix} \Sigma_{\theta f} & \Sigma_{\theta x_i} \\ \Sigma_{ff} & \Sigma_{fx_i} \end{pmatrix}
P'_i \Lambda_i^{-1/2} \Sigma_{z_i s_i} \Sigma_{s_i z_i} \Lambda_i^{-1/2} P_i
\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i \theta} & \Sigma_{x_i f} \end{pmatrix}
\widetilde{P}' \widetilde{\Lambda}^{-1/2} \overrightarrow{h}_{m,i} .
\end{aligned}
$$

$$(3.9)$$

Therefore, according to Equation (3.4), agent i's problem of information acquisition can be stated as

$$
\max_{\Sigma_{z_i s_i}} E\left[\widetilde{U^i}\left(k_i\left(\overrightarrow{s}_i\right), K\left(\overrightarrow{w}\right), \overrightarrow{\theta}\right)\right]
$$

s.t. $\det\left(I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}\right) \geq 1 - \beta_i^2$ and $I - \Sigma_{z_i s_i} \Sigma_{s_i z_i}$ is positive semi-definite.

Proposition 3.3 *Given the current "belief" $K_m\left(\overrightarrow{\xi}\right) = \overrightarrow{t}'_m \overrightarrow{\xi}$, agent i's optimal information choice is*

$$
\Sigma_{z_i s_i} = \beta_i \cdot \frac{\Lambda_i^{-1/2} P_i \begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i \theta} & \Sigma_{x_i f} \end{pmatrix} \widetilde{P}' \widetilde{\Lambda}^{-1/2} \overrightarrow{h}_{m,i}}{\left|\left| \Lambda_i^{-1/2} P_i \begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i \theta} & \Sigma_{x_i f} \end{pmatrix} \widetilde{P}' \widetilde{\Lambda}^{-1/2} \overrightarrow{h}_{m,i} \right|\right|}.
$$

$$(3.10)$$

Proof. This is a direct application of Lemma 3.2 in Appendix 3.7. ∎

Remark 3.2 *Proposition 3.3 shows that it is optimal to collect information in a single dimension, and thus the agent's signal $\overrightarrow{s}_i$ is actually a scalar s_i. This is intuitive. Since the agents can only take actions in one dimension, collecting informa-*

tion other than the optimal direction just wastes their precious capacity. Moreover,
agent i acquires information in the direction specified by Equation (3.10), since this
information is the most relevant to her welfare.

Now we examine how capacity endowment affects agents' welfare. Since $\beta_i^2 = 1 - \exp\left(-2 \cdot \kappa_i\right)$ is strictly increasing with respect to capacity κ_i, we can conduct the welfare analysis with respect to β_i^2 .

Substituting (3.10) into (3.9) leads to an expression of agent i's expected utility in β_i^2:

$$
\begin{aligned}
&E\left[\widetilde{U^i}\left(k_i, K, \overrightarrow{\theta}\right)\right]\\
={}& -\overrightarrow{h}\,'_{m,i}\,\overrightarrow{h}_{m,i}\\
&+\overrightarrow{h}\,'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\ \Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P_i'\Lambda_i^{-1/2}\Sigma_{z_i s_i}\Sigma_{s_i z_i}\Lambda_i^{-1/2}P_i\begin{pmatrix}\Sigma_{f\theta} & \Sigma_{ff}\\ \Sigma_{x_i\theta} & \Sigma_{x_i f}\end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}\\
={}& -\overrightarrow{h}\,'_{m,i}\,\overrightarrow{h}_{m,i} + \beta_i^2\,\overrightarrow{h}\,'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\left[Var\begin{pmatrix}\overrightarrow{\theta}\\ \overrightarrow{f}\end{pmatrix} - Var\left(\begin{pmatrix}\overrightarrow{\theta}\\ \overrightarrow{f}\end{pmatrix}\Bigg|\begin{pmatrix}\overrightarrow{f}\\ \overrightarrow{x}_i\end{pmatrix}\right)\right]\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}\\
={}& -\left(r_i\cdot\overrightarrow{u}_m + \overrightarrow{b}_i\right)'\left[Var\left(\overrightarrow{w}\right) - \beta_i^2\cdot\left[Var\left(\overrightarrow{w}\right) - Var\left(\overrightarrow{w}\Bigg|\begin{pmatrix}\overrightarrow{f}\\ \overrightarrow{x}_i\end{pmatrix}\right)\right]\right]\left(r_i\cdot\overrightarrow{u}_m + \overrightarrow{b}_i\right)\\
\triangleq{}& -\left(r_i\cdot\overrightarrow{u}_m + \overrightarrow{b}_i\right)'\left[Var\left(\overrightarrow{w}\right) - \beta_i^2\cdot H_i\right]\left(r_i\cdot\overrightarrow{u}_m + \overrightarrow{b}_i\right)\,, \qquad (3.11)
\end{aligned}
$$

where

$$
H_i = Var\left(\overrightarrow{w}\right) - Var\left(\overrightarrow{w}\Bigg|\begin{pmatrix}\overrightarrow{f}\\ \overrightarrow{x}_i\end{pmatrix}\right)
$$

represents the information content of agent i's information sources. This information content is the maximal information that agent i can extract from her information sources and it can be achieved only if agent i has infinite capacity, i.e. $\beta_i^2 = 1$. When her capacity is limited ($\beta_i^2 < 1$), agent i can at most acquire β_i^2 fraction of

this information content, i.e. $\beta_i^2 \cdot H_i$. The vector $\left(r_i \cdot \overrightarrow{u}_m + \overrightarrow{b}_i \right)$ represents the direction of information most relevant to agent i's welfare and $\left[Var\left(\overrightarrow{w} \right) - \beta_i^2 \cdot H_i \right]$ is her remaining uncertainty about the state vector after acquiring information in the optimal way. We can also rewrite Equation (3.11) as

$$
\begin{aligned}
& E\left[\widetilde{U^i}\left(k_i, K, \overrightarrow{\theta} \right) \right] \\
= & -\left(r_i \cdot \overrightarrow{u}_m + \overrightarrow{b}_i \right)' \left[\left(1 - \beta_i^2 \right) \cdot Var\left(\overrightarrow{w} \right) + \beta_i^2 \cdot Var\left(\overrightarrow{w} \Big| \begin{pmatrix} \overrightarrow{f} \\ \overrightarrow{x}_i \end{pmatrix} \right) \right] \left(r_i \cdot \overrightarrow{u}_m + \overrightarrow{b}_i \right),
\end{aligned}
$$

which suggests that agent i's maximal expected utility is just a weighted average of her expected utility with unlimited capacity and that with no capacity, where the weight $\beta_i^2 \in [0, 1]$ is determined by her capacity κ_i.

3.4.3 Two Equivalence Theorems

This subsection establishes two equivalence theorems among the games. These theorems improve our understanding of the "quadratic-normal" games with endogenous information acquisition.

Definition: *Given the same state* $\left(\overrightarrow{\theta}', \overrightarrow{f}' \right)'$*, two games* $\left\{ r_{l,i}, \overrightarrow{\eta}_{l,i}, \beta_{l,i}, \overrightarrow{x}_{l,i} \right\}_{i \in [0,1]}$*,* $l = 1, 2$ *are macro-equivalent if* $\forall K_m\left(\overrightarrow{w} \right) = \overrightarrow{u}'_m \overrightarrow{w}$*, i.e., the two games generate the same aggregate action* $K_{m+1}\left(\overrightarrow{w} \right)$*.*

Two macro-equivalent games are observational equivalent at the macro-level. Since the equilibrium aggregate action is a fixed point of the above iteration, macro-equivalence implies the same equilibrium aggregate action of the two games. Here we derive the necessary and sufficient condition for macro-equivalence.

By Equation (3.5) in Subsection 3.4.2, agent i's best response to her signal s_i is

$$
\begin{aligned}
k_i\left(\overrightarrow{s}_i\right) &= E_i\left(r_i \cdot K_m\left(\overrightarrow{\xi}\right) + \overrightarrow{c}_i'\overrightarrow{\xi}\right) \\
&= E_i\left[r_i \cdot \overrightarrow{t}_m'\overrightarrow{\xi} + \overrightarrow{c}_i'\overrightarrow{\xi}\right] \\
&= \left(r_i \cdot \overrightarrow{t}_m + \overrightarrow{c}_i\right)' E_i\left(\overrightarrow{\xi}\right) \\
&= \overrightarrow{h}_{m,i}'\Sigma_{\xi s_i}s_i \\
&= \overrightarrow{h}_{m,i}'\Sigma_{\xi s_i}\Sigma_{s_i\xi}\overrightarrow{\xi} + \overrightarrow{h}_{m,i}'\Sigma_{\xi s_i}\epsilon_i \ ,
\end{aligned}
\tag{3.12}
$$

where the fourth equality comes from the definition $\overrightarrow{h}_{m,i} = r_i \cdot \overrightarrow{t}_m + \overrightarrow{c}_i$, and the idiosyncratic noise ϵ_i consists of both agent i's information processing error and the idiosyncratic noise in her private information source $\overrightarrow{x}_i$. Combining Equation (3.7) and (3.10) implies

$$
\Sigma_{s_i\xi} = \Sigma_{s_i z_i}\Lambda_i^{-1/2}P_i\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}
$$

$$
= \beta_i \cdot \frac{\overrightarrow{h}_{m,i}'\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix} \Sigma_{\theta f} & \Sigma_{\theta x_i} \\ \Sigma_{ff} & \Sigma_{fx_i} \end{pmatrix}P_i'\Lambda_i^{-1/2}\Lambda_i^{-1/2}P_i\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}}{\left\|\Lambda_i^{-1/2}P_i\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}\right\|}
$$

$$
= \beta_i \cdot \frac{\overrightarrow{h}_{m,i}'\widetilde{\Lambda}^{-1/2}\widetilde{P}H_i\widetilde{P}'\widetilde{\Lambda}^{-1/2}}{\left\|\Lambda_i^{-1/2}P_i\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}\right\|}
$$

$$
\triangleq \beta_i \cdot \frac{\overrightarrow{h}_{m,i}'A_i}{\left\|\Lambda_i^{-1/2}P_i\begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}\right\|}
\tag{3.13}
$$

where we define $A_i = \widetilde{\Lambda}^{-1/2} \widetilde{P} H_i \widetilde{P}' \widetilde{\Lambda}^{-1/2}$, which is a symmetric matrix. Substituting (3.13) into (3.12) leads to

$$
\begin{aligned}
k_i\left(\overrightarrow{s}_i\right) &= \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \Sigma_{s_i \xi} \overrightarrow{\xi} + \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \epsilon_i \\
&= \beta_i^2 \cdot \overrightarrow{h}'_{m,i} \frac{\left(A_i \overrightarrow{h}_{m,i}\right)\left(A_i \overrightarrow{h}_{m,i}\right)' \overrightarrow{\xi}}{\left|\left| \Lambda_i^{-1/2} P_i \begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix} \widetilde{P}' \widetilde{\Lambda}^{-1/2} \overrightarrow{h}_{m,i} \right|\right|^2} + \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \epsilon_{ii} \\
&= \beta_i^2 \cdot \frac{\left(\overrightarrow{h}'_{m,i} A_i \overrightarrow{h}_{m,i}\right)\left(A_i \overrightarrow{h}_{m,i}\right)' \overrightarrow{\xi}}{\left(\overrightarrow{h}'_{m,i} A_i \overrightarrow{h}_{m,i}\right)} + \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \epsilon_i \\
&= \beta_i^2 \cdot \overrightarrow{h}'_{m,i} A_i \overrightarrow{\xi} + \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \epsilon_i .
\end{aligned}
$$

Taking integral of both sides of the above equation generates the new aggregate action

$$
\begin{aligned}
K_{m+1}\left(\overrightarrow{\xi}\right) &= \left(\int_0^1 \beta_i^2 \cdot \overrightarrow{h}'_{m,i} A_i \cdot di\right) \overrightarrow{\xi} + \int_0^1 \overrightarrow{h}'_{m,i} \Sigma_{\xi s_i} \epsilon_i \cdot di \\
&= \left(\int_0^1 \beta_i^2 \cdot \overrightarrow{h}'_{m,i} A_i \cdot di\right) \overrightarrow{\xi} .
\end{aligned} \tag{3.14}
$$

Expressing K_{m+1} in the original state $\overrightarrow{w} = \left(\overrightarrow{\theta}', \overrightarrow{f}'\right)'$, we have

$$
\begin{aligned}
K_{m+1}\left(\overrightarrow{w}\right) &= \left(\int_0^1 \beta_i^2 \cdot \left(r_i \cdot \overrightarrow{u}_m + \overrightarrow{b}_i\right)' H_i \cdot di\right) \overrightarrow{w} \\
&= \overrightarrow{u}'_m \left(\int_0^1 r_i \cdot \beta_i^2 \cdot H_i \cdot di\right) \overrightarrow{w} + \left(\int_0^1 \beta_i^2 \cdot H_i \overrightarrow{b}_i \cdot di\right)' \overrightarrow{w} \tag{3.15}
\end{aligned}
$$

Let $\Gamma_l = \left\{r_{l,i}, \overrightarrow{\eta}_{l,i}, \beta_{l,i}, \overrightarrow{x}_{l,i}\right\}_{i \in [0,1]}$, $l = 1, 2$ be two games with the same fundamental $\overrightarrow{\theta}$ and public information source $\overrightarrow{f}$. Then Equation (3.15) implies that Γ_1

and Γ_2 are macro-equivalent iff

$$\forall \overrightarrow{u}_m \in \mathbb{R}^{n_w}, \; \overrightarrow{u}'_m \left(\int_0^1 r_{1,i} \cdot \beta_{1,i}^2 \cdot H_{1,i} \cdot di \right) \overrightarrow{w} + \left(\int_0^1 \beta_{1,i}^2 \cdot H_{1,i} \overrightarrow{b}_{1,i} \cdot di \right)' \overrightarrow{w}$$

$$= \; \overrightarrow{u}'_m \left(\int_0^1 r_{2,i} \cdot \beta_{2,i}^2 \cdot H_{2,i} \cdot di \right) \overrightarrow{w} + \left(\int_0^1 \beta_{2,i}^2 \cdot H_{2,i} \overrightarrow{b}_{2,i} \cdot di \right)' \overrightarrow{w} \; .$$

The result is summarized in the proposition below.

Proposition 3.4 *Given the same state* $\left(\overrightarrow{\theta}', \overrightarrow{f}' \right)'$, *two games* $\Gamma_l = \left\{ r_{l,i}, \overrightarrow{\eta}_{l,i}, \beta_{l,i}, \overrightarrow{x}_{l,i} \right\}_{i \in [0,1]}$, $l = 1, 2$ *are macro-equivalent iff: 1)*

$$\int_0^1 r_{1,i} \cdot \beta_{1,i}^2 \cdot H_{1,i} \cdot di = \int_0^1 r_{2,i} \cdot \beta_{2,i}^2 \cdot H_{2,i} \cdot di \; ;$$

2)

$$\int_0^1 \beta_{1,i}^2 \cdot H_{1,i} \overrightarrow{b}_{1,i} \cdot di = \int_0^1 \beta_{2,i}^2 \cdot H_{2,i} \overrightarrow{b}_{2,i} \cdot di \; .$$

Here, $H_{l,i}$ *reflects the information content of agent* i*'s information sources in game* Γ_l*, and* $\overrightarrow{b}_{l,i} = \left(\overrightarrow{\eta}'_{l,i}, \overrightarrow{0} \right)$*,* $l = 1, 2$.

Proposition 3.4 shows that our economy's macro behavior is detail-free. The equilibrium does not depend on the specific structure of private information sources, capacity constraint, or utility functions. As a special case, game $\Gamma_1 = \{ r_i, \overrightarrow{\eta}_i, \beta_i, \overrightarrow{x}_i \}_{i \in [0,1]}$ is macro-equivalent to $\Gamma_2 = \left\{ r_i \cdot \beta_i^2, \beta_i^2 \cdot \overrightarrow{\eta}_i, 1, \overrightarrow{x}_i \right\}_{i \in [0,1]}$. Therefore, adding capacity constraint is equivalent to reducing the magnitude of strategic externality (i.e., $|r_i| \geq |r_i \cdot \beta_i^2|$) and agents' sensitivity to the fundamentals (i.e., $||\overrightarrow{\eta}_i|| \geq ||\beta_i^2 \cdot \overrightarrow{\eta}_i||$). This result can be better understood from the concept of micro-equivalence below.

Definition: *Given the same state* $\left(\overrightarrow{\theta}', \overrightarrow{f}' \right)'$ *and the same profile of private information sources* $\{ \overrightarrow{x}_i \}_{i \in [0,1]}$*, two games* $\Gamma_l = \left\{ r_{l,i}, \overrightarrow{\eta}_{l,i}, \beta_{l,i} \right\}_{i \in [0,1]}$*,* $l = 1, 2$ *are*

micro-equivalent iff $\forall K_m \left(\overrightarrow{w} \right) = \overrightarrow{u}'_m \overrightarrow{w}, \ \forall i \in [0, 1],$

$$E \left[k_{1,i,m+1} \Big| \left(\begin{array}{c} \overrightarrow{f} \\ \overrightarrow{x}_i \end{array} \right) \right] = E \left[k_{2,i,m+1} \Big| \left(\begin{array}{c} \overrightarrow{f} \\ \overrightarrow{x}_i \end{array} \right) \right] .$$

Two games are micro-equivalent if their corresponding agents' (i.e. the two agents with the same index) responses to the shocks are effectively identical. Here two actions are "effectively identical" if they only differ in their idiosyncratic noise terms. Now we derive the necessary and sufficient condition for the micro-equivalence between two games.

For any $K_m\left(\overrightarrow{w}\right) = \overrightarrow{u}'_m \overrightarrow{w} = \overrightarrow{t}'_m \overrightarrow{\xi}$, by Equation (3.12) we have

$$E\left[k_{i,m+1}\Big|\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right]$$

$$= E\left[\overrightarrow{h}'_{m,i}\Sigma_{\xi s_i}s_i\Big|\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right]$$

$$= \overrightarrow{h}'_{m,i}\Sigma_{\xi s_i}E\left[s_i|\overrightarrow{z}_i\right]$$

$$= \overrightarrow{h}'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\Sigma_{z_i s_i}\Sigma_{s_i z_i}\overrightarrow{z}_i$$

$$= \frac{\beta_i^2}{||\Lambda_i^{-1/2}P_i\begin{pmatrix}\Sigma_{f\theta} & \Sigma_{ff}\\\Sigma_{x_i\theta} & \Sigma_{x_i f}\end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}||^2}$$

$$\cdot\overrightarrow{h}'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\Lambda_i^{-1/2}P_i\begin{pmatrix}\Sigma_{f\theta} & \Sigma_{ff}\\\Sigma_{x_i\theta} & \Sigma_{x_i f}\end{pmatrix}\widetilde{P}'\widetilde{\Lambda}^{-1/2}\overrightarrow{h}_{m,i}$$

$$\overrightarrow{h}'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\overrightarrow{z}_i$$

$$= \beta_i^2\cdot\overrightarrow{h}'_{m,i}\widetilde{\Lambda}^{-1/2}\widetilde{P}\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\overrightarrow{z}_i$$

$$= \beta_i^2\cdot\left(r_i\cdot\overrightarrow{u}_m + \overrightarrow{b}_i\right)'\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\overrightarrow{z}_i\ .$$

Note that the term $\begin{pmatrix}\Sigma_{\theta f} & \Sigma_{\theta x_i}\\\Sigma_{ff} & \Sigma_{fx_i}\end{pmatrix}P'_i\Lambda_i^{-1/2}\overrightarrow{z}_i$ only depends on the information sources, which are identical for the two games. Hence, two games are micro-equivalent iff $\forall i \in [0,1]$ and $\forall\overrightarrow{u}_m \in \mathbb{R}^{n_w}$, $\beta_{1,i}^2\cdot\left(r_{1,i}\cdot\overrightarrow{u}_m + \overrightarrow{b}_{1,i}\right) = \beta_{2,i}^2\cdot\left(r_{2,i}\cdot\overrightarrow{u}_m + \overrightarrow{b}_{2,i}\right)$, where $\overrightarrow{b}_{l,i} = \left(\overrightarrow{\eta}'_{l,i}, \overrightarrow{0}'\right)'$. The result is summarized in the following proposition.

172

Proposition 3.5 *Given the same state* $\left(\overrightarrow{\theta}{}', \overrightarrow{f}{}'\right)'$ *and the same profile of private information sources* $\{\overrightarrow{x}_i\}_{i\in[0,1]}$, *two games* $\Gamma_l = \left\{r_{l,i}, \overrightarrow{\eta}_{l,i}, \beta_{l,i}\right\}_{i\in[0,1]}$, $l = 1, 2$ *are micro-equivalent iff 1)*

$$r_{1,i} \cdot \beta_{1,i}^2 = r_{2,i} \cdot \beta_{2,i}^2 \ ;$$

2)

$$\beta_{1,i}^2 \cdot \overrightarrow{\eta}_{1,i} = \beta_{2,i}^2 \cdot \overrightarrow{\eta}_{2,i} \ .$$

Particularly, $\Gamma_1 = \left\{r_i, \overrightarrow{\eta}_i, \beta_i\right\}_{i\in[0,1]}$, *a game with capacity constraints (thus also endogenous information acquisition), is micro-equivalent to* $\Gamma_2 = \left\{r_i \cdot \beta_i^2, \beta_i^2 \cdot \overrightarrow{\eta}_i, 1\right\}_{i\in[0,1]}$, *which is a game with exogenous information structure.*

Proposition 3.5 establishes an equivalence among the games that are very different at first glance. It clearly shows how capacity constraints (hence, endogenous information acquisition) affect the outcomes of the game. There are two effects here. The first one is a "dampening effect". Because of the capacity constraint, agent i can not effectively respond to all the information content contained in her information sources. Part of her response is inefficient in that it comes on her information processing errors. Parameter $\overrightarrow{\eta}_i$ in game Γ_1 measures agent i's sensitivity to the fundamentals and $\beta_i \in [0,1]$ characterizes her capacity constraint. If instead, we believe that agent i has infinite capacity, then the equivalence theorem requires us to also believe that she has a smaller sensitivity $\beta_i^2 \cdot \overrightarrow{\eta}_i$ to correctly explain her expected actions. $\beta_i^2 \cdot \overrightarrow{\eta}_i$ in game Γ_2 reflects this dampening effect.

The second effect is called "(less) coordination effect", which is "derived" from the dampening effect. Parameter r_i measures the strategic externality between agent i's action and the aggregate action. Compared to the benchmark case $r_i = 0$, agent i responds more (less) to the public information source when her action is strategic complementary (substitutable) to others'. Thus the public information source serves as a coordination device. However, adding the capacity constraint weakens agent

i's ability to coordinate by reducing the effectiveness of her response. Furthermore, rationally expecting others' dampened (and thus less effective) responses, agent i finds it less attractive to coordinate with others, since it makes no sense to match others' information processing errors. Therefore, adding capacity constraint not only reduces the agents' ability to coordinate, but also reduces their incentive to do so. This (less) coordination effect is best reflected by the comparison between r_i in game Γ_1 and $r_i \cdot \beta_i^2$ in game Γ_2.

3.4.4   Equilibrium Analysis

This subsection characterizes the equilibrium. Once the aggregate action is determined, all the information choices and individual actions are also determined. Thus we focus on deriving the equilibrium aggregate action in this subsection.

As discussed before, the equilibrium aggregate action $K\left(\overrightarrow{\xi}\right) = \overrightarrow{t}'\overrightarrow{\xi}$ is a fixed point of Equation (3.14), i.e.,

$$
\begin{aligned}
\overrightarrow{t}'\overrightarrow{\xi} &= \left(\int_0^1 \beta_i^2 \cdot \overrightarrow{h}'_{m,i} A_i \cdot di\right)\overrightarrow{\xi} \\
&= \left(\int_0^1 \beta_i^2 \cdot \left(r_i \cdot \overrightarrow{t} + \overrightarrow{c}_i\right)' A_i \cdot di\right)\overrightarrow{\xi} \\
&= \overrightarrow{t}'\left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right)\overrightarrow{\xi} + \left(\int_0^1 \beta_i^2 \cdot A_i \overrightarrow{c}_i \cdot di\right)'\overrightarrow{\xi} .
\end{aligned}
$$

This equation should hold for any realization of $\overrightarrow{\xi}$, thus we must have

$$
\left[I - \left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right)\right]\overrightarrow{t} = \left(\int_0^1 \beta_i^2 \cdot A_i \overrightarrow{c}_i \cdot di\right). \tag{3.16}
$$

The property of Equation (3.16) mainly depends on matrix

$$
\left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right) .
$$

174

By Lemma 3.3 in Appendix 3.8, $\forall i \in [0, 1]$, all the eigenvalues of A_i belong to $[0, 1]$ and the maximal one is 1. Thus $||A_i|| = 1$ and

$$
\begin{aligned}
\left|\left|\left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right)\right|\right| &\leq \int_0^1 ||r_i \cdot \beta_i^2 \cdot A_i|| \cdot di \\
&= \int_0^1 |r_i| \cdot \beta_i^2 \cdot ||A_i|| \cdot di \\
&= \int_0^1 |r_i| \cdot \beta_i^2 \cdot di \\
&< \int_0^1 |r_i| \cdot di < 1 \ .
\end{aligned}
$$

Hence, matrix

$$
\left[I - \left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right)\right]
$$

is invertible and in equilibrium $\overrightarrow{t}$ is uniquely determined by

$$
\overrightarrow{t} = \left[I - \left(\int_0^1 r_i \cdot \beta_i^2 \cdot A_i \cdot di\right)\right]^{-1} \left(\int_0^1 \beta_i^2 \cdot A_i \overrightarrow{c}_i \cdot di\right) \tag{3.17}
$$

According to the definition that $\overrightarrow{c}_i = \widetilde{\Lambda}^{1/2} \widetilde{P} \overrightarrow{b}_i$, $\overrightarrow{t} = \widetilde{\Lambda}^{1/2} \widetilde{P} \overrightarrow{u}$, $\overrightarrow{\xi} = \widetilde{\Lambda}^{-1/2} \widetilde{P} \overrightarrow{w}$ and $A_i = \widetilde{\Lambda}^{-1/2} \widetilde{P} H_i \widetilde{P}' \widetilde{\Lambda}^{-1/2}$, we can express Equation (3.17) in the original state vector $\overrightarrow{w} = \left(\overrightarrow{\theta}', \overrightarrow{f}'\right)'$ as

$$
\overrightarrow{u} = \left[\Sigma_{ww} - \left(\int_0^1 r_i \cdot \beta_i^2 \cdot H_i \cdot di\right)\right]^{-1} \left(\int_0^1 \beta_i^2 \cdot H_i \overrightarrow{b}_i \cdot di\right) \tag{3.18}
$$

Now we summarize the results in the proposition below.

Proposition 3.6 *Let Γ be a game characterized by $\{r_i, \overrightarrow{\eta}_i, \beta_i, \overrightarrow{x}_i\}_{i \in [0,1]}$ and $\overrightarrow{w} = \left(\overrightarrow{\theta}', \overrightarrow{f}'\right)'$, then Γ has a unique equilibrium characterized by 1) (aggregate action) $K(\overrightarrow{w}) = \overrightarrow{u}' \overrightarrow{w}$, where*

$$
\overrightarrow{u} = \left[\Sigma_{ww} - \left(\int_0^1 r_i \cdot \beta_i^2 \cdot H_i \cdot di\right)\right]^{-1} \left(\int_0^1 \beta_i^2 \cdot H_i \overrightarrow{b}_i \cdot di\right) \tag{3.18}
$$

175

$$\overrightarrow{b}_i = \left(\overrightarrow{\eta}'_i, \overrightarrow{0}\right)';$$

2) (information choice)

$$\forall i \in [0,1], \ \Sigma_{z_i s_i} = \beta_i \cdot \frac{\Lambda_i^{-1/2} P_i \begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix} \left(r_i \cdot \overrightarrow{u} + \overrightarrow{b}_i\right)}{||\Lambda_i^{-1/2} P_i \begin{pmatrix} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{pmatrix} \left(r_i \cdot \overrightarrow{u} + \overrightarrow{b}_i\right) ||} \tag{3.19}$$

where Equation (3.19) comes from substituting $\overrightarrow{h}_i = \widetilde{\Lambda}^{1/2} \widetilde{P} \left(r_i \cdot \overrightarrow{u} + \overrightarrow{b}_i\right)$ into Equation (3.10) of Proposition 3.3 ;

3) (individual action)

$$\forall i \in [0,1], \ k_i(s_i) = E_i \left(r_i \cdot K(\overrightarrow{w}) + \overrightarrow{b}'_i \overrightarrow{w}\right) ,$$

where $E_i(\cdot) \triangleq E(\cdot|s_i)$.

Compared to the multi-equilibria result of Hellwig and Veldkamp (2009), under the similar condition $\int_0^1 |r_i| \cdot di < 1$, our framework admits a unique equilibrium in a general setup, regardless of the heterogeneity in preferences, information sources and capacities. Two reasons explain this difference. The first one is the continuous nature of information choices in our model. The agents can continuously allocate their attention to different information sources. In Hellwig and Veldkamp (2009), however, even though after eliminating the discreteness in the information choices, the agents still face a "yes or no" problem of either paying the additional attention to the information others paid attention to, or to the information that others ignored. This very discontinuity leads to the multiplicity in their model. The second reason is that the information processing errors are in nature private in our rational inattention approach. This private noise facilitates the uniqueness of the

equilibrium.

3.5   Conclusion

We examine information acquisition in a class of quadratic-normal games through the approach of rational inattention. We show that adding capacity constraint essentially reduces not only the agents' ability to coordinate but also their incentive to do so. Through two equivalence theorems, we show that the games with capacity constraints are observational equivalent to the games with higher level of strategic externalities but having no capacity constraints. For future research, it might be interesting to compare the game in this chapter where the action space is continuous to the game with discrete action space in Chapter 2. This comparison may further our understanding of the formation of (approximate) common knowledge.

3.6   Appendix

Lemma 3.1 *Let* $\mathbf{W}$ *be a space of* n_w*-dimensional random vectors defined on probability space* $\mathbf{P}$*. A mapping* $K : \mathbb{R}^{n_w} \to \mathbb{R}^{n_k}$ *satisfies*

$$\forall \overrightarrow{w} \in \mathbf{W}, \ K\left(\overrightarrow{w}\right) = \widetilde{E}\left(K\left(\overrightarrow{w}\right)\right) + \widehat{E}\left(\overrightarrow{w}\right), \tag{3.20}$$

where $\widetilde{E}\left(\cdot\right) = \int R_i E_i\left(\cdot\right) di$, $\widehat{E}\left(\cdot\right) = \int B_i E_i\left(\cdot\right) di$ *are two average expectation operators,* $E_i\left(\cdot\right)$ *is the expectation operator of agent* i*, and* $R_i \in \mathbb{R}^{n_k \times n_k}$*,* $B_i \in \mathbb{R}^{n_k \times n_w}$ *are two matrices. If*

$$\forall \overrightarrow{w} \in \mathbf{W}, K_0\left(\overrightarrow{w}\right) = \widetilde{E}\left(K_0\left(\overrightarrow{w}\right)\right) \ \text{implies} \ \forall \overrightarrow{w} \in \mathbf{W}, K_0\left(\overrightarrow{w}\right) = \overrightarrow{0}, \tag{3.21}$$

then the solution $K(\cdot)$ is unique and linear in its argument, i.e. $\exists$ a unique $U \in$
$\mathbb{R}^{n_k \times n_w}$, *s.t.* $K(\overrightarrow{w}) = U\overrightarrow{w}$.

Proof. Let $\alpha_l \in \mathbb{R}$, $\overrightarrow{w}_l \in \mathbf{W}$, $l \in \{1, 2\}$. Since $K(\cdot)$ satisfies (3.20), we have

$$K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) = \widetilde{E}(K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2)) + \widehat{E}(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) \quad (3.22)$$

and

$$\alpha_l \cdot K(\overrightarrow{w}_l) = \alpha_l \cdot \widetilde{E}(K(\overrightarrow{w}_l)) + \alpha_l \cdot \widehat{E}(\overrightarrow{w}_l), l \in \{1, 2\}.$$

By definition, both $\widetilde{E}(\cdot)$ and $\widehat{E}(\cdot)$ are linear operators, thus

$$
\begin{aligned}
\alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2) &= \alpha_1 \cdot \widetilde{E}(K(\overrightarrow{w}_1)) + \alpha_2 \cdot \widetilde{E}(K(\overrightarrow{w}_2)) + \alpha_1 \cdot \widehat{E}(\overrightarrow{w}_1) + \alpha_2 \cdot \widehat{E}(\overrightarrow{w}_2) \\
&= \widetilde{E}(\alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2)) + \widehat{E}(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2).
\end{aligned}
$$

Combining the above equation and (3.22) leads to

$$
\begin{aligned}
&K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) - [\alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2)] \\
&= \widetilde{E}(K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) - [\alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2)]),
\end{aligned}
$$

then by condition (3.21), we have $K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) - [\alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2)] = 0$, i.e. $K(\alpha_1 \cdot \overrightarrow{w}_1 + \alpha_2 \cdot \overrightarrow{w}_2) = \alpha_1 \cdot K(\overrightarrow{w}_1) + \alpha_2 \cdot K(\overrightarrow{w}_2)$. Therefore, $K(\cdot)$ is a linear mapping, i.e. $\exists U \in \mathbb{R}^{n_k \times n_w}$, s.t. $K(\overrightarrow{w}) = U\overrightarrow{w}$.

Suppose $K_1(\cdot)$ and $K_2(\cdot)$ are two solutions of Equation (3.20), then by the linearity of $\widetilde{E}(\cdot)$, $\forall \overrightarrow{w} \in \mathbf{W}, K_2(\overrightarrow{w}) - K_1(\overrightarrow{w}) = \widetilde{E}(K_2(\overrightarrow{w}) - K_1(\overrightarrow{w}))$. Therefore, condition (3.21) implies that $K_2(\overrightarrow{w}) - K_1(\overrightarrow{w}) = \overrightarrow{0}$, i.e. the solution of Equation (3.20) is unique. ∎

Remark 3.3 *Condition (3.21) holds iff the operator $\left(I - \widetilde{E}\right)(\cdot)$ is invertible, where*

I is the identity operator. If $\left(I - \widetilde{E}\right)(\cdot)$ is not invertible and Equation (3.20) has a solution, then it has infinitely many solutions.

3.7   Appendix

Lemma 3.2 *Let $\overrightarrow{b} \in \mathbb{R}^{n_z}$ and $\overrightarrow{b} \neq \overrightarrow{0}$. Let M be the space of all real matrices with n_z rows, then $\beta \cdot \overrightarrow{b} / \| \overrightarrow{b} \|$ solves the problem below:*

$$\max_{\Sigma_{zs} \in M} \overrightarrow{b}' \Sigma_{zs} \Sigma_{sz} \overrightarrow{b}$$

$$s.t. \ \det\left(I - \Sigma_{zs}\Sigma_{sz}\right) \geq 1 - \beta^2, \tag{3.23}$$

and $I - \Sigma_{zs}\Sigma_{sz}$ is positive semi-definite,

where $\beta \in [0, 1]$.

Proof. $\because I - \Sigma_{zs}\Sigma_{sz}$ is a real and symmetric matrix, $\therefore \exists$ orthogonal matrix P s.t. $I - \Sigma_{zs}\Sigma_{sz} = Q'DQ$, where D is a $n_z \times n_z$ diagonal matrix with diagonal entries $d_1 \ \ d_2 \ \ \cdots \ \ d_{n_z}$. Actually, $d_1 \ \ d_2 \ \ \cdots \ \ d_{n_z}$ are the eigenvalues of $I - \Sigma_{zs}\Sigma_{sz}$ and $\prod_{j=1}^{n_z} d_j = \det(D) = \det(I - \Sigma_{zs}\Sigma_{sz})$, thus Inequality (3.23) becomes $\prod_{j=1}^{n_z} d_j \geq 1 - \beta^2$. Also note that $d_j \geq 0$, $j = 1, 2, \cdots, n_z$, since $I - \Sigma_{zs}\Sigma_{sz}$ is positive definite.

On the other hand, $I - Q'DQ = \Sigma_{zs}\Sigma_{sz}$ is positive semi-definite, i.e. all of its eigenvalues μ_j , $j = 1, 2, \cdots, n_z$ are non-negative. Let μ be an arbitrary eigenvalue of $I - Q'DQ$, then we have

$$
\begin{aligned}
0 &= \det(\mu \cdot I - I + Q'DQ) = \det((\mu - 1) \cdot I + Q'DQ) \\
&= \det((\mu - 1) \cdot QQ' + D) = \det((\mu - 1) \cdot I + D)
\end{aligned}
$$

179

i.e. the eigenvalues of $I - Q'DQ$ can be expressed as $\mu_j = 1 - d_j$, $j = 1, 2, \cdots, n_z$. Since $\mu_j \geq 0$, we have $d_j \leq 1$, $j = 1, 2, \cdots, n_z$.

Thus

$$\vec{b}' \Sigma_{zs} \Sigma_{sz} \vec{b} = \vec{b}' (I - Q'DQ) \vec{b} = ||\vec{b}|| - \left(Q\vec{b} \right)' D \left(Q\vec{b} \right)$$

and $\max_{\Sigma_{zs} \in M} \vec{b}' \Sigma_{zs} \Sigma_{sz} \vec{b}$ is equivalent to $\min_{D,Q} \left(Q\vec{b} \right)' D \left(Q\vec{b} \right)$.

Now we restate the optimization problem as

$$\min_{D,Q} \left(Q\vec{b} \right)' D \left(Q\vec{b} \right)$$

$$\text{s.t. } Q'Q = I , \prod_{j=1}^{n_z} d_j \geq 1 - \beta^2 \text{ and } d_j \in [0, 1] , j = 1, 2, \cdots, n_z.$$

Let $Q' = \left(\begin{array}{cccc} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_{n_z} \end{array} \right)$, then

$$\left(Q\vec{b} \right)' D \left(Q\vec{b} \right) = \sum_{j=1}^{n_z} d_j \cdot \left(\vec{q}'_j \vec{b} \right)^2$$

Since Q is orthogonal, we have $\vec{q}'_j \vec{q}_k = \delta(j, k)$, where $\delta(j, k) = \left\{ \begin{array}{cc} 1 & \text{if } j=k \\ 0 & \text{otherwise} \end{array} \right\}$, i.e. $\vec{q}_1$ $\vec{q}_2$ $\cdots$ $\vec{q}_{n_f + n_\theta}$ form an orthonormal basis of $\mathbb{R}^{n_z}$. Thus $\sum_{j=1}^{n_z} \left(\vec{q}'_j \vec{b} \right)^2 = ||\vec{b}||^2$.

Then

$$\left(Q\vec{b} \right)' D \left(Q\vec{b} \right) = \sum_{j=1}^{n_z} d_j \cdot (\vec{q}_j \vec{e})^2 \geq \min_{j=1,2,\cdots,n_z} d_j \cdot ||\vec{b}||^2$$

$\because \prod_{j=1}^{n_z} d_j \geq 1 - \beta^2$ and $d_j \in [0, 1]$, $j = 1, 2, \cdots, n_z$, $\therefore \min_{j=1,2,\cdots,n_z} d_j \cdot ||\vec{b}||^2 \geq \left(1 - \beta^2 \right) \cdot ||\vec{b}||^2$. Therefore, $\left(Q\vec{b} \right)' D \left(Q\vec{b} \right) \geq \left(1 - \beta^2 \right) \cdot ||\vec{b}||^2$. Note that this minimum can be obtained if and only if $\exists j$ s.t. $d_j = 1 - \beta^2$, $\left(\vec{q}'_j \vec{b} \right)^2 = ||\vec{b}||^2$ and

$\forall k \neq j$, $d_k = 1$, $\overrightarrow{q}'_k \overrightarrow{b} = 0$. Without loss of any generality, let $d_1 = \left(1 - \beta^2\right)$ and $\overrightarrow{q}'_1 \overrightarrow{b} = || \overrightarrow{b} ||$. Since $\forall j, k$, $\overrightarrow{q}'_j \overrightarrow{q}_k = \delta(j, k)$, we have $\overrightarrow{q}_1 = \overrightarrow{b} / || \overrightarrow{b} ||$.

Let $\overrightarrow{q}_l = \left(\begin{array}{cccc} Q_{l1} & Q_{l2} & \cdots & Q_{l,n_z} \end{array} \right)'$, $l = 1, 2, \cdots, n_z$, then Q can be expressed by $(Q_{lj})_{n_z \times n_z}$. Thus the (l, k) entry of matrix $Q'DQ$ is

$$
\begin{aligned}
(Q'DQ)_{lk} &= \sum_{j=1}^{n_z} d_j \cdot Q_{jl} \cdot Q_{jk} \\
&= \left(1 - \beta^2\right) \cdot Q_{1l} \cdot Q_{1k} + \sum_{j=2}^{n_z} d_j \cdot Q_{jl} \cdot Q_{jk}
\end{aligned}
$$

Note that $d_j = 1$, $j = 2, 3, \cdots n_z$ and $\sum_{j=1}^{n_z} Q_{jl} \cdot Q_{jk} = \delta(l, k)$ [4] , we have

$$
\begin{aligned}
Q'DQ_{lk} &= \left(1 - \beta^2\right) \cdot Q_{1l} \cdot Q_{1k} + \sum_{j=2}^{n_z} Q_{jl} \cdot Q_{jk} \\
&= \left(1 - \beta^2\right) \cdot Q_{1l} \cdot Q_{1k} + \delta(l, k) - Q_{1l} \cdot Q_{1k} \\
&= \delta(l, k) - \beta^2 \cdot Q_{1l} \cdot Q_{1k}
\end{aligned}
$$

$\therefore$

$$
\begin{aligned}
(\Sigma_{zs}\Sigma_{sz})_{lk} &= (I - Q'DQ)_{lk} \\
&= \delta(l, k) - \delta(l, k) + \beta^2 \cdot Q_{1l} \cdot Q_{1k} \\
&= \beta^2 \cdot Q_{1l} \cdot Q_{1k}
\end{aligned}
$$

i.e.

$$
\Sigma_{zs}\Sigma_{sz} = \beta^2 \cdot \overrightarrow{b} \overrightarrow{b}' / || \overrightarrow{b} ||^2
$$

$\because rank\left(\Sigma_{zs}\Sigma_{sz}\right) = rank\left(\overrightarrow{b} \overrightarrow{b}'\right) \leq rank\left(\overrightarrow{b}\right) = 1$ and it is obvious that $\Sigma_{zs}\Sigma_{sz} \neq 0$,$\therefore rank\left(\Sigma_{zs}\Sigma_{sz}\right) = 1$.

---

[4]Because the columns of an orthogonal matrix also form an orthonormal basis of $\mathbf{R}^{n_z}$.

$$\therefore \exists \ \alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_{n_s} \ \text{ and } \overrightarrow{t} \in \mathbb{R}^{n_z}, \ ||\overrightarrow{t}|| = 1 \text{ s.t.}$$

$$\Sigma_{zs} = \overrightarrow{t} \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n_s} \end{pmatrix} = \begin{pmatrix} \alpha_1 \cdot \overrightarrow{t} & \alpha_2 \cdot \overrightarrow{t} & \cdots & \alpha_{n_s} \cdot \overrightarrow{t} \end{pmatrix}$$

$$\therefore$$

$$\sum_{j=1}^{n_s} \alpha_j^2 \cdot \overrightarrow{t}\,\overrightarrow{t}' = \Sigma_{zs}\Sigma_{sz} = \beta^2 \cdot \overrightarrow{b}\,\overrightarrow{b}'/||\overrightarrow{b}||^2$$

$$\therefore$$

$$\begin{aligned}
\sum_{j=1}^{n_s} \alpha_j^2 &= tr\left(\sum_{j=1}^{n_s} \alpha_j^2 \cdot \overrightarrow{t}\,\overrightarrow{t}'\right) \\
&= tr\left(\beta^2 \cdot \overrightarrow{b}\,\overrightarrow{b}'/||\overrightarrow{b}||^2\right) = \beta^2 \cdot \overrightarrow{b}\,\overrightarrow{b}'/||\overrightarrow{b}||^2 = \beta^2
\end{aligned}$$

$$\therefore$$

$$\overrightarrow{t}\,\overrightarrow{t}' = \overrightarrow{b}\,\overrightarrow{b}'/||\overrightarrow{b}||^2$$

Let $\overrightarrow{t} = \begin{pmatrix} t_1 & t_2 & \cdots & t_{n_z} \end{pmatrix}'$ and $\overrightarrow{b} = \begin{pmatrix} b_1 & b_2 & \cdots & b_{n_z} \end{pmatrix}'$. $t_1^2 = b_1^2/||\overrightarrow{b}||^2$ implies that $t_1 = \pm b_1/||\overrightarrow{b}||$, then $t_1 \cdot t_2 = b_1 \cdot b_2/||\overrightarrow{b}||^2$ implies $t_2 = \pm b_2/||\overrightarrow{b}||$. By the same argument we have $t_j = \pm b_j//||\overrightarrow{b}||$, $j = 1, 2, \cdots, n_z$. Thus $\overrightarrow{t} = \pm \overrightarrow{b}/||\overrightarrow{b}||$. Since the sign of α_j is changeable, we just let $\overrightarrow{t} = \overrightarrow{b}/||\overrightarrow{b}||$. Since all the columns of the optimal Σ_{zs} are parallel to each other, we can just let $\Sigma_{zs} = \beta \cdot \overrightarrow{b}/||\overrightarrow{b}||$ without any loss of generality. ∎

3.8   Appendix

Lemma 3.3 $\forall i \in [0,1]$, *let*

$$A_i = \widetilde{\Lambda}^{-1/2} \widetilde{P} H_i \widetilde{P}' \widetilde{\Lambda}^{-1/2} = \widetilde{\Lambda}^{-1/2} \widetilde{P} \left[Var\left(\overrightarrow{w}\right) - Var\left(\overrightarrow{w}| \left(\begin{array}{c} \overrightarrow{f} \\ \overrightarrow{x}_i \end{array} \right) \right) \right] \widetilde{P}' \widetilde{\Lambda}^{-1/2} \; ,$$

then all the eigenvalues of A_i belong to $[0,1]$ and the maximal one is 1.

Proof. First note that

$$H_i = \left(\begin{array}{cc} \Sigma_{\theta f} & \Sigma_{\theta x_i} \\ \Sigma_{ff} & \Sigma_{fx_i} \end{array} \right) P_i' \Lambda_i^{-1} P_i \left(\begin{array}{cc} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{array} \right) ,$$

thus

$$A_i = \widetilde{\Lambda}^{-1/2} \widetilde{P} \left(\begin{array}{cc} \Sigma_{\theta f} & \Sigma_{\theta x_i} \\ \Sigma_{ff} & \Sigma_{fx_i} \end{array} \right) P_i' \Lambda_i^{-1} P_i \left(\begin{array}{cc} \Sigma_{f\theta} & \Sigma_{ff} \\ \Sigma_{x_i\theta} & \Sigma_{x_if} \end{array} \right) \widetilde{P}' \widetilde{\Lambda}^{-1/2}$$

is positive semi-definite and all its eigenvalues are non negative. Let λ be any eigenvalue of A_i, i.e,

$$
\begin{aligned}
0 &= \det\left(\lambda \cdot I - A_i\right) \\
&= \det\left(\lambda \cdot \widetilde{P}'\widetilde{\Lambda}\widetilde{P} - H_i\right) \\
&= \det\left(\lambda \cdot Var\left(\overrightarrow{w}\right) - H_i\right) \\
&= \det\left(\lambda \cdot Var\left(\overrightarrow{w}\right) - \left[Var\left(\overrightarrow{w}\right) - Var\left(\overrightarrow{w}| \left(\begin{array}{c} \overrightarrow{f} \\ \overrightarrow{x}_i \end{array} \right) \right) \right] \right) \\
&= \det\left((\lambda - 1) \cdot Var\left(\overrightarrow{w}\right) + Var\left(\overrightarrow{w}| \left(\begin{array}{c} \overrightarrow{f} \\ \overrightarrow{x}_i \end{array} \right) \right) \right) \quad\quad (3.24)
\end{aligned}
$$

183

Since $Var\left(\overrightarrow{w}\right)$ is strictly positive definite and $Var\left(\overrightarrow{w}\,\middle|\,\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right)$ is positive semi-definite, $\lambda - 1 > 0$ implies that Equation (3.24) does not hold. Thus all the eigenvalues of A_i can not exceed 1. If $\lambda - 1 = 0$, left side of Equation (3.24) becomes

$$
\det\left(0 \cdot Var\left(\overrightarrow{w}\right) + Var\left(\overrightarrow{w}\,\middle|\,\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right)\right)
$$

$$
= \det\left(Var\left(\overrightarrow{w}\,\middle|\,\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right)\right)
$$

$$
= \det\left(Var\left(\begin{pmatrix}\overrightarrow{\theta}\\\overrightarrow{f}\end{pmatrix}\,\middle|\,\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right)\right)
$$

$$
= \det\begin{pmatrix}Var\left(\overrightarrow{\theta}\,\middle|\,\begin{pmatrix}\overrightarrow{f}\\\overrightarrow{x}_i\end{pmatrix}\right) & 0\\[2mm] 0 & 0\end{pmatrix} = 0
$$

thus $\lambda = 1$ is the maximal eigenvalue of A_i. ∎

Bibliography

ANGELETOS, G.-M., AND A. PAVAN (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

AXELSON, U. L. F. (2007): "Security Design with Investor Private Information," *The Journal of Finance*, 62(6), pp. 2587–2632.

BIAIS, B., AND T. MARIOTTI (2005): "Strategic Liquidity Supply and Security Design," *Review of Economic Studies*, 72(3), 615–649.

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

CARLSSON, H., AND E. V. DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.

COVER, T. M., AND J. A. THOMAS (1991): *Elements of information theory*. Wiley-Interscience, New York, NY, USA.

DANG, T. V., G. GORTON, AND B. HOLMSTROM (2010): "Ignorance and the Optimality of Debt for Liquidity Provision," Unpublished working paper, Yale University.

DEMARZO, P. M. (2005): "The Pooling and Tranching of Securities: A Model of Informed Intermediation," *Review of Financial Studies*, 18(1), 1–35.

DeMarzo, P. M., and D. Duffie (1999): "A Liquidity-Based Model of Security Design," *Econometrica*, 67(1), 65–100.

DeMarzo, P. M., I. Kremer, and A. Skrzypacz (2005): "Bidding with Securities: Auctions and Security Design," *American Economic Review*, 95(4), 936–959.

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

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

Gorton, G., and G. G. Pennacchi (1990): "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 45(1), 49–71.

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

Innes, R. D. (1990): "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, 52(1), 45–67.

Leland, H. E., and D. H. Pyle (1977): "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance*, 32(2), 371–87.

Luo, Y. (2008): "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 11(2), 366–385.

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

——— (2011): "Business cycle dynamics under rational inattention," Working Paper Series 1331, European Central Bank.

MATEJKA, F. (2010): "Rigid Pricing and Rationally Inattentive Consumer," CERGE-EI Working Papers wp409, The Center for Economic Research and Graduate Education - Economic Institute, Prague.

MATEJKA, F., AND C. A. SIMS (2011): "Discrete Actions in Information-Constrained Tracking Problems," *SSRN eLibrary*.

MILGROM, P., AND J. ROBERTS (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255–1277.

MONDRIA, J. (2010): "Portfolio choice, attention allocation, and price comovement," *Journal of Economic Theory*, 145(5), 1837–1864.

MORRIS, S., AND H. S. SHIN (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *The American Economic Review*, 88(3), pp. 587–597.

——— (2001): "Global Games: Theory and Applications," Cowles Foundation Discussion Papers 1275R, Cowles Foundation for Research in Economics, Yale University.

——— (2002): "Social Value of Public Information," *The American Economic Review*, 92(5), pp. 1521–1534.

——— (2004): "Coordination risk and the price of debt," *European Economic Review*, 48(1), 133 – 153.

MYATT, D. P., AND C. WALLACE (2011): "Endogenous Information Acquisition in Coordination Games," *The Review of Economic Studies*.

PACIELLO, L. (2009): "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention," MPRA Paper 16407, University Library of Munich, Germany.

ROSS, S. A. (1977): "The Determination of Financial Structure: The Incentive-Signalling Approach," *Bell Journal of Economics*, 8(1), 23–40.

SHANNON, C. E. (1948): "A mathematical theory of Communication," *The Bell system technical journal*, 27, 379–423.

SIMS, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49(1), 317–356.

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

——— (2005): "Rational inattention: a research agenda," Discussion Paper Series 1: Economic Studies 2005,34, Deutsche Bundesbank, Research Centre.

——— (2006): "Rational Inattention: Beyond the Linear-Quadratic Case," *The American Economic Review*, 96(2), pp. 158–163.

——— (2010): "Rational Inattention and Monetary Economics," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3, chap. 04, pp. 155–181. Elsevier, 1 edn.

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

TOWNSEND, R. M. (1979): "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 21(2), 265–293.

Van Nieuwerburgh, S., and L. Veldkamp (2009a): "Information Acquisition and Under-Diversification," *Review of Economic Studies, Vol. 77, Issue 2, pp. 779-805, April 2010.*

——— (2009b): "Information Immobility and the Home Bias Puzzle," *The Journal of Finance*, 64(3), 1187–1215.

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

Woodford, M. (2008): "Inattention as a source of randomized discrete adjustment," Unpublished working paper, Columbia University.

——— (2009): "Information-constrained state-dependent pricing," *Journal of Monetary Economics*, 56(Supplement 1), S100 – S124, Supplement issue: December 12-13, 2008 Research Conference on 'Monetary Policy under Imperfect Information' Sponsored by the Swiss National Bank (http://www.snb.ch) and Study Center Gerzensee (www.szgerzensee.ch).

Yang, M. (2011): "Coordination with Rational Inattention," Working paper 014 2011, Princeton Economic Theory Center.