Optimality of Debt under Flexible Information Acquisition

By Ming Yang

This paper studies the optimality of debt in a setting of endogenous and 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 with information, we assume no initial information asymmetry. The buyer has expertise in acquiring information about the fundamentals in a flexible way. 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 to induce the buyer to acquire the information least harmful to the seller's interest. Issuing debt (i.e., pooling and tranching) is uniquely optimal in raising liquidity, regardless of the stochastic interdependence of the underlying assets and the distribution of bargaining power. Fixed aggregate risk and homogeneous information cost are the key factors driving the results.

JEL: D86, D82, G32

Keywords: security design, flexible information acquisition, pooling and tranching

I. Introduction

Pooling assets and issuing asset-backed securities (ABSs) is a widely-used means of raising liquidity. Commercial banks pool large numbers of 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 follows: a risk-neutral seller owns assets that generate uncertain future cash flows. The seller 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, setting its price, 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 paper, but will be greatly enriched to capture our key ideas concerning the optimality of debt.

* Yang: Fuqua School of Business, Duke University, 100 Fuqua Drive, Durham, North Carolina, ming.yang@duke.edu. An earlier version has been circulated under the title "Optimality of Securitized Debt with Endogenous and Flexible Information Acquisition". This work is partially supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370. I am grateful to Stephen Morris for his generous guidance and considerations. I am also indebted to Christopher Sims and Hyun Shin for their constant guidance. I thank Ulf Axelson, Sudipto Bhattacharya, Richard Blundell, Paolo Colla, Martin Cripps, Tri Vi Dang, Jan Eeckhout, Emmanuel Farhi, Simon Gervais, Antonio Guarino, Oliver Hart, Barney Hartman-Glaser, Christian Hellwig, Bengt Holmstrom, Arvind Krishnamurthy, Thomas Mariotti, Guillaume Plantin, Adriano Rampini, Morten Ravn, Andrei Shleifer, Alp Simsek, Jean Tirole, Dimitri Vayanos, Vish Viswanathan, Michael Woodford, Wei Xiong, Kathy Yuan, Yao Zeng, and seminar participants at Harvard, NYU, Princeton, Stanford GSB, Duke, Toulouse School of Economics, Kellogg, LSE, the 2011 North American Summer Meeting of the Econometric Society and the 2012 Winter Meeting of American Economic Association for helpful comments and discussions. All remaining errors are mine.

1
The literature on security design (e.g., (Peter M. DeMarzo and Darrell Duffie 1999), (Peter M. DeMarzo 2005)) has provided insight into this securitization process. Much of this literature models sellers as “insiders” who have private information about the assets, a source of adverse selection that makes potential buyers reluctant to provide liquidity. This literature considers the possibility of signaling by sellers, where buyers are passive because they cannot acquire information about the assets. The conclusions are also sensitive to various assumptions on information, assets and feasible security designs.

This paper explores adverse selection from another perspective. Rather than assume exogenous private information, we consider adverse selection resulting from information acquisition. A justification for this approach is that in real world people acquire different kinds of information from different sources, so that the observed information asymmetry stems from differing ability to acquire information in the first place. While other models posit the information asymmetry as exogenous, we model it as endogenous. The main finding of this paper is that when one party to the trade, seller or buyer, designs the contract and the other party acquires information about the fundamentals, then issuing debt is the uniquely optimal means of liquidity provision. Without loss of generality, we first consider the case in which the buyer can acquire information about the assets according to the security proposed by the seller in the benchmark model, which has endogenous adverse selection, unlike the classic security design literature. We follow (Tri Vi Dang, Gary Gorton and Bengt Holmstrom 2010) in treating the buyer as an “expert” who acquires information accordingly. In reality, the buyers of ABSs are highly sophisticated. Their expertise in assessing investment opportunities is naturally modeled by endogenous information acquisition, not exogenous endowment. Here endogeneity means that the agents can choose from a set of information structures according to their investment opportunities. Given this endogeneity, sellers design securities so as to generate the least incentive for buyers to acquire information. (Dang, Gorton and Holmstrom 2010) model such information acquisition by an all-or-not technology, in which buyers either acquire a specific signal on the future cash flow of the asset or get no information whatever. In other words, the buyer can only choose between two specific information structures. This is essentially the same as the costly state verification (CSV) approach of (Robert M. Townsend 1979). We refer to this type of information acquisition technology as CSV. Given this rigid information acquisition process, (Dang, Gorton and Holmstrom 2010) show that debt is the least information-sensitive and thus optimal for liquidity provision. However, there also exist infinitely many other securities, called “quasi-debts”, which are just as information-sensitive as the standard debt contract. Further, some restrictive conditions are required to ensure the optimality of these quasi-debts in the case of pooling. As noted below, their non-uniqueness result stems from the rigidity of information acquisition under the CSV approach.

This paper differs from (Dang, Gorton and Holmstrom 2010) in allowing for flexi-

---

1See (U. L. F. Axelsson 2007) for a justification for the case in which the buyer can acquire information and thus has an information advantage over the seller. Here, again, however our result of unique optimality of debt does not depend on this assumption.
ble information acquisition, which helps achieve the unique optimality of debt securities even when various assets are pooled. Like (Dang, Gorton and Holmstrom 2010), we assume no information asymmetry at the beginning in order to focus on the adverse selection resulting from endogenous information acquisition. Moreover, in our version of flexibility, the buyer can acquire information about the underlying assets according to the shape of the security. This flexibility makes intuitive sense. For example, a debt holder would pay attention mainly to bad states, as the only consideration is default risk, while an equity holder will pay more attention to good states, given the greater benefit from upside payments. For an arbitrary security, the investor’s incentive to acquire information, i.e., allocation of attention, would be determined accordingly, in turn affecting the seller’s incentives in designing the security. In other words, the security design interacts with the investor’s incentive to allocate attention to different aspects of the underlying cash flows. Capturing the potential variety of attention allocation thus calls for sufficiently flexible characterization of information acquisition.

We model flexible information acquisition through the paradigm of rational inattention, building upon (Christopher A. Sims 2003), where any information structure\(^2\) can be acquired at a cost proportional to the reduction of entropy. This cost could result from the time or resources required to run models, do statistical tests, or write reports. Flexibility enables the buyer to acquire payoff-relevant information accordingly, while the information cost requires her to optimally acquire such information in both quantitative and qualitative aspects.

As in (Dang, Gorton and Holmstrom 2010), in our model standard debt is optimal for liquidity provision. But our result is sharper in the sense that debt is the sole optimal means, and pooling can be taken into account naturally when there exist multiple underlying assets. In (Dang, Gorton and Holmstrom 2010), given the CSV framework, only two extreme information structures are available, but infinite forms of securities can be designed, so that inevitably some are indistinguishable. In our framework, thanks to flexibility, the variety of information structures available matches that of potential securities to be designed. So the uniqueness of the standard debt can be guaranteed in our model, and quasi-debts are no longer optimal. By reshaping as flat the uneven tail above the price of a quasi-debt the information cost to the buyer can be saved and the potential loss of trading due to adverse selection mitigated. The resulting surplus could be used by the seller to make both parties better off, and thus ultimately permit greater provision of liquidity. Moreover, flexible information acquisition offers a unified framework for analyzing the securitization of multiple assets. We show that pooling and issuing debt is uniquely optimal to raise liquidity, regardless of the stochastic interdependence among the underlying assets or the distribution of bargaining power.

There are two essential factors in the optimality of standard debt. One is the fixed aggregate risk implicitly specified in the benchmark trading game, in the sense that the total cash flows of seller and buyer are invariant with respect to the success or failure of the transaction. This factor—and its underlying mechanism in the unique optimality of debt—does not depend on whether the buyer or the seller acquires information, as long

\(^2\) An information structure is a conditional distribution of signals given the fundamentals.
as the roles of security design and information acquisition are allocated separately to the two parties. Specifically, in our benchmark model, since the aggregate risk is fixed, information acquisition is not socially valuable, so acquiring information is a waste of resources when both parties are considered together as a whole. Moreover, fixed aggregate risk produces conflicting interests between the two parties, so that the buyer gains by acquiring information but at the expense of the seller through adverse selection which further reduces the potential gain from trading. Since the buyer’s incentive to acquire information is shaped by the offer, the seller deliberately designs the ABS to discourage information acquisition harmful to her own interests. Given the limited liability constraint, any feasible ABS is bounded above by the sum of underlying cash flows. When information cost is not too high, flexibility allows the buyer to distinguish between all states with different payoffs. Hence the seller makes the ABS a constant whenever it is beyond the limited liability boundary in order to discourage information acquisition and mitigate adverse selection. This creates a flat tail. In states where the underlying cash flows are too small to support that constant level of payment, the ABS reaches the boundary and equals the sum of the cash flows. Hence, the flat tail and the boundary component constitute a debt, which is uniquely optimal for liquidity provision. We also provide an example with variable aggregate risk to illustrate the importance of fixed aggregate risk in our framework. Consider the seller as an entrepreneur who raises funds from the buyer to undertake a project with uncertain cash flow. If the buyer accepts the offer, they are jointly exposed to an aggregate risk. If the offer is rejected, they are not exposed to the risk. This situation resembles a production economy\(^3\), in which information acquisition could increase aggregate social welfare and the interests of the two parties could be partly aligned. Therefore, the seller could deliberately design a contract to encourage the buyer to acquire a certain type of information to avert investment where cash flows are too low. This increases the seller’s own benefit while also increasing social welfare.

Another key factor is homogeneity in information acquisition. That is, no state is characterized by special difficulty in information acquisition. This feature stems from rational inattention; it is the reason why our qualitative result does not depend on stochastic interdependence among the underlying assets. Intuitively, if information about some assets is much easier to acquire, the flat part of the debt cannot be preserved in the optimal ABS. We provide an example to illustrate this idea.

Finally, the uniqueness of the optimal contract does not derive from flexible information acquisition alone, but from the combination of flexible information acquisition and the inherent flexibility of security design. We call this feature symmetric flexibility. In principle, general flexibility of choice, not necessarily restricted to information acquisition alone, enables an economic agent to make state-contingent responses. In other words, the agent can make one best response in one state, another best response in another state. Symmetric flexibility requires that both parties are endowed with the same level of flexibility.

How this symmetric flexibility works can be seen by comparing our framework with

\(^3\) Accordingly, the situation with fixed aggregate risk resembles an exchange economy.
that of (Dang, Gorton and Holmstrom 2010) and the traditional models of CSV à la (Townsend 1979). In all three, the contract designer is endowed with flexibility; that is, she can assign state-contingent repayment by designing any form of security. What matters in shaping the different results on uniqueness of the optimal contract is the potential flexibility of the other party, who decides whether or not to accept the offer. In our framework, ex-ante symmetric information (in the form of two-sided ignorance) prevents the buyer from making a state-contingent choice via the traditional CSV approach to information acquisition. But in our framework the buyer can choose the state-contingent probability of accepting the offer. In this sense, the buyer enjoys as much flexibility as the seller. This symmetry of flexibility guarantees the uniqueness of the optimal contract, which is the standard debt. In (Dang, Gorton and Holmstrom 2010), however, the buyer can only take the traditional CSV approach, which offers only two options, namely, acquiring all the information existing or none at all. In other words, the buyer in (Dang, Gorton and Holmstrom 2010) cannot make state-contingent decisions. Hence, the desired symmetry of flexibility fails, and so, therefore, does the uniqueness of the optimal contract. Interestingly, (Townsend 1979) also employs the CSV approach with two options to model information acquisition (to audit or not), but the unique optimality of a standard debt still emerges. Why is this the case? Unlike (Dang, Gorton and Holmstrom 2010) and our framework, (Townsend 1979) gives the entrepreneur an informational advantage, in the sense that the entrepreneur, unlike the lender, knows the profit that the project will yield. Thanks to the revelation principle, the lender who acquires information in the interim stage can decide whether to audit or not in any state, based on the truth told by the entrepreneur. In other words, although in (Townsend 1979) the lender still only has two options to acquire information, like the buyer in (Dang, Gorton and Holmstrom 2010), the choice can nevertheless be state-contingent. Therefore, the symmetry of flexibility is still established in (Townsend 1979), and the uniqueness of the optimal contract, also a standard debt, is ensured as well.4

We proceed as follows. Section II 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 III derives the unique optimality of debt in various circumstances and identifies two key driving forces for this result. We conclude and discuss in Section IV.

Relation to the Literature. We model players’ information acquisition in the rational inattention framework building on (Christopher A. Sims 1998) and (Sims 2003).5 In applied work, rational inattention is studied mainly in two cases: linear-quadratic (e.g., (Bartosz Mackowiak and Mirko Wiederholt 2009)), and binary-action. A good example of the latter is (Michael Woodford 2009) , where firms acquire information and then

---

4Nevertheless, the argument of (Townsend 1979) does not apply to the cases with random audits. (Dilip Mookherjee and Ivan Png 1989) generalize the result of (Townsend 1979) and point out that debt may not be optimal when random auditing is allowed.

5For more about rational inattention, see also (Christopher A. Sims 2005),(Christopher A. Sims 2006), (Christopher A. Sims 2010), (Yulei Luo 2008), (Bartosz Mackowiak and Mirko Wiederholt 2011), (Stijn Van Nieuwerburgh and Laura Veldkamp 2009a), (Stijn Van Nieuwerburgh and Laura Veldkamp 2009b), (Luigi Paciello 2009), (Filip Matejka 2010), (Jordi Mondria 2010), (Filip Matejka and Christopher A. Sims 2011).
decide whether to revise their prices. Like (Ming Yang 2011), this paper adopts the binary-action setup in a strategic framework, which differs from the single-person decision problem studied in (Woodford 2009). Compared with (Yang 2011), where both players acquire information and move simultaneously, here we posit that players move sequentially and only one party acquires information, which results in information asymmetry. Also, here we focus on a specific security design problem, not a general coordination game as in (Yang 2011). Together with (Yang 2011), this paper is one of the first to incorporate rational inattention-based flexible information acquisition into strategic problems. It offers a number of new results.

This paper is also closely related to the security design literature, in much of which sellers are modeled as “insiders” with private information. Their information advantage over buyers results in adverse selection, hence inefficient trading. To overcome adverse selection, given that buyers cannot acquire information, sellers want to signal their private information and so partly restore efficient trading. In this process, appropriate security design is crucial. This is because signaling is costly, so designing a security that is less information-sensitive than the original asset can save the signaling cost, thus increasing sellers’ profit. This argument is plausible, and insightful results are well established in the literature, but other interesting possibilities are worth investigating. Further, this literature imposes various assumptions in order to deliver its results. In our model, buyers may actively acquire information, which could alter the interplay between the two parties and produce different results on security design; and we can get clearer results from a single assumption.

The fundamental difference with much of the security design literature emerges clearly if we examine some of their assumptions and results in detail. (Gary Gorton and George G. Pennacchi 1990) show that splitting assets into debt and equity mitigates the “lemon” problem between outsiders and insiders. They do not consider a security design problem but instead directly assume the existence of debt. In (DeMarzo and Duffie 1999), informed sellers signal the quality of assets to competitive liquidity suppliers by retaining part of the cash flows. Equity is issued when the contractible information is not particularly sensitive to sellers’ inside knowledge. If the information structure allows a uniform worst case, standard debt is optimal within the set of non-decreasing securities. (Bruno Biais and Thomas Mariotti 2005) study the effects of market power on liquidity. They derive both the optimal security and the trading mechanism through mechanism design. The debt contract turns out to be optimal under the distributional conditions of the underlying cash flows. (DeMarzo 2005) focuses on the consequences of pooling and tranching. Pooling has an information destruction effect, which destroys the seller’s ability to signal the quality of the assets separately. When tranching is possible, pooling may also have a risk diversification effect, reducing information sensitivity of the senior tranche. Under certain distributional assumptions of the noise structure, (DeMarzo 2005) shows that as the number of underlying assets goes to infinity the risk diversification effect dominates that of the information destruction. In this limiting case, pooling and tranching become optimal. These models also restrict their attention to non-decreasing securities\(^6\).

\(^6\)(Biais and Mariotti 2005) also assume dual monotonicity, i.e., both the security and the residual cash flow are non-
(Robert D. Innes 1990) provides a standard motivation for this constraint. When the security is not monotone, a seller may cheat by borrowing from a third party, reporting a large cash flow to reduce the repayment and then repaying the side loan. The validity of this argument depends on the context. In the case of publicly traded stocks or bonds, such fraud is improbable; it is difficult or even illegal for the seller to manipulate the cash flows. Moreover, when the security is written on multiple underlying assets, the concept of monotonicity itself is poorly defined. Our framework does not suffer from these limits. Last, some recent works on the relationship between security design and market liquidity, such as (Gilles Chemla and Christopher Hennessy 2011) and (Marco Pagano and Paolo Volpin 2012), argue that the design of different securities is closely related to liquidity in the primary market, the secondary market, or both, and that this relationship depends crucially on market participants’ information. But these works are still limited to cases in which the issuers of securities have ex-ante information advantages.

It is also interesting to contrast our work with (Axelson 2007). Unlike the signaling literature on security design such as (DeMarzo and Duffie 1999) and subsequent work, (Axelson 2007) considers a security design problem in which the buyer, not the seller, has private information about the asset. Axelson’s benchmark model could be placed in my own framework as the case of variable aggregate risk, with the seller designing contract and the buyer having information advantage, so that it is natural to expect equity rather than debt to be a better solution to the seller’s financing problem. Our framework differs more deeply from (Axelson 2007) in that we take the buyer’s private information as endogenous, a consequence of flexible information acquisition. Further, as noted above, our results are more general in that they are not restricted to the case where the buyer gets private information, but only require that one party designs the contract and the other acquires information.

From a partially related angle, Christopher Hennessy ((Christopher Hennessy 2009) and (Christopher Hennessy 2011)) considers the interaction between information acquisition and market microstructure. Specifically, following the noisy rational expectation framework, he simultaneously considers two information channels: endogenous acquisition by speculators and market price. When speculators account for most of the market, riskless debt proves to be the optimal security, as it mitigates speculators’ incentive for information acquisition and thus endogenous adverse selection, which fits with the present paper. By contrast, when noise traders’ liquidity demand is high, risky debt becomes optimal, since its issuance exploits the information role of market price when endogenous information acquisition is less important.

II. Binary Decision under Flexible Information Acquisition

Before introducing the economic environment of the security design problem, let us review the logic of binary choice under flexible information acquisition, which will play a crucial role in the analysis. Readers interested strictly in the security design problem may want to skip this section and go back to it when needed.
In our central example, the buyer faces a take-it-or-leave-it offer, and must make a binary decision to accept it or not after acquiring information about some payoff-relevant state, where information acquisition is modeled as choosing an information structure. First we posit information structures with binary signals, then showing that this is sufficient.

A. The Decision Problem

Consider an agent who has to choose an action \( a \in \{0, 1\} \) and will receive a payoff \( u(a, \theta) \), where \( \theta \in \Theta \subseteq \mathbb{R} \) is an unknown state distributed according to probability measure \( P \) over \( \Theta \). We assume that \( P \) is continuous with respect to Lebesgue measure over \( \mathbb{R} \).

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 \( \theta \) (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. Acquisition is flexible in that no restriction is imposed on \( m \). By choosing different functional forms for \( m(\theta) \), the agent can make the signal covary with the state in any way she likes. Intuitively, if the payoff is sensitive to fluctuations of the state within some range \( A \subseteq \Theta \), she would pay attention to the states 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 classified by 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 (Claude E. Shannon 1948). Information conveyed by an information structure \( m(\cdot) \) is given by the mutual information between the signal and the state, defined as the expected reduction in uncertainty due to observation of the signals generated according to \( m(\cdot) \), where the uncertainty is measured by Shannon’s entropy.

Specifically, given information structure \( m(\cdot) \), mutual information \( I(m) \) is given by

\[
I(m) = \mathbb{E}g(m(\theta)) - g(\mathbb{E}m(\theta)),
\]

where the expectation operator \( \mathbb{E} \cdot \) is with respect to \( \theta \) under prior \( P \), and

\[
g(x) = x \cdot \ln x + (1 - x) \cdot \ln (1 - x).
\]

Without loss of generality, any binary-signal information structure can be represented by a Lebesgue measurable function that takes value in \([0, 1]\). Let \( M \) be the set of all such

---

7 At first glance, positing binary information structures would seem to be a stringent restriction that makes information acquisition inflexible. In Subsection II.B, however, we will see that this restriction is unnecessary.

8 Appendix A provides a derivation of mutual information. We refer readers to (Thomas M. Cover and Joy A. Thomas 1991) for further details.
functions. We assume that the agent can choose any information structure in \( M \). In this sense, information acquisition is flexible.

Information acquisition is also costly. Let \( c : M \rightarrow \mathbb{R}_+ \) be the cost (in terms of utility) of information. We assume that this cost is proportional to the quantity of information acquired, i.e.,

\[
(2) \quad c ( m ) = \mu \cdot \mathcal{I} ( m ) ,
\]

where \( \mu > 0 \) is the marginal cost of information acquisition, i.e., it measures the difficulty in acquiring information. When \( \mu = 0 \), information acquisition is cost-free and the agent observes the true state directly. When \( \mu \to \infty \), the agent cannot acquire any information at all.

An intuitive way to understand mutual information \( \mathcal{I} ( m ) \) is that it measures the variability of function \( m \), the informativeness of the signal with respect to the state. For example, when \( m ( \theta ) \) is constant, the signal conveys no information about \( \theta \) and the corresponding mutual information is nil. This is because function \( g \) is strictly convex, hence \( \mathcal{I} ( m ) \) is zero if and only if \( m ( \theta ) \) is constant. A nice property of our technology of information acquisition, therefore, is that there exists information acquisition if and only if \( m ( \theta ) \) varies over \( \theta \), if and only if the 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 by making \( m ( \theta ) \) highly sensitive to \( \theta \) within the event. In this sense our technology of information acquisition is flexible, since the agent can decide the pointwise quality of their information by choosing any information structure from \( M \). Note further that \( c ( \cdot ) \) is convex, i.e.,

\[
c ( t \cdot m_1 + (1 - t) \cdot m_2 ) \leq t \cdot c ( m_1 ) + (1 - t) \cdot c ( m_2 )
\]

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 constant in \( \theta \).

**Solving the Binary Decision Problem with Information Acquisition**

We proceed to 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 expected utility

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

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

\[
m^* ( \theta ) = m ( \theta ) f (1) + (1 - m ( \theta )) f (0) ,
\]

...
then $V(m^*, f^*) \geq V(m, f)$. This is because the first term of (3) remains the same, while the information cost becomes smaller due to the convexity of $c(\cdot)$.9

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

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

$$V^*(m) = \mathbb{E} [m(\theta) \cdot u(1, \theta) + (1 - m(\theta)) \cdot u(0, \theta)] - c(m)$$

$$= \mathbb{E} [m(\theta) \cdot [u(1, \theta) - u(0, \theta)]] - c(m) + \mathbb{E}u(0, \theta) .$$

Since $\mathbb{E}u(0, \theta)$ does not depend on $m$, we can redefine the agent’s objective as

$$\max_{m \in M} V^*(m) = \mathbb{E} [\Delta u(\theta) \cdot m(\theta)] - c(m) ,$$

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 for information acquisition.

The following lemma characterizes the optimal strategy $m$ for the agent.10

**PROPOSITION 1:** 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

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

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

i) the optimal strategy is unique;

ii) there are three possibilities for the optimal strategy:

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

$$\mathbb{E}\exp(-\mu^{-1} \Delta u(\theta)) \leq 1 ;$$

9A simple proof: the convexity of $c(\cdot)$ implies

$$c(\alpha \cdot m) < \alpha \cdot c(m)$$

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

$$c(m^*) = c(f(1) \cdot [\alpha \cdot m + 1 - \alpha])$$

$$\leq f(1) \cdot c(\alpha \cdot (a \cdot m + 1 - \alpha))$$

$$\leq f(1) \cdot c((\alpha \cdot c(m) + 0))$$

$$\leq \Delta f \cdot c(m) < c(m) .$$

10We became aware of the related work of Michael Woodford while working on this paper. Here we use Lemma 2 from (Michael Woodford 2008) to characterize the optimal strategy. To maintain the completeness of our paper, we give a proof in our context.

11We do not have to require $\Theta \subset \mathbb{R}$. This proposition holds for any probability space $\Theta$. 
b) \( p_1 = 0 \) (i.e., \( m(\theta) = 0 \) almost surely) if and only if
\[
E \exp (\mu^{-1} \Delta u (\theta)) \leq 1 ;
\]

c) \( p_1 \in (0, 1) \) if and only if
\[
E \exp (\mu^{-1} \Delta u (\theta)) > 1 \quad \text{and} \quad E \exp (-\mu^{-1} \Delta u (\theta)) > 1 ;
\]
in this case, the optimal strategy \( m \) is characterized by
\[
\Delta u (\theta) = \mu \cdot \left[ g' (m(\theta)) - g'(p_1) \right]
\]
for all \( \theta \in \Theta \), where
\[
g'(x) = \ln \left( \frac{x}{1-x} \right).
\]

**Proof:** See Appendix A.

Proposition 1 fully characterizes the agent’s possible optimal decisions on information acquisition. Case a) and Case b) correspond to a scenario of an extreme prior, in the sense that there exists an ex-ante optimal action, 1 or 0. These two cases do not involve information acquisition and thus also correspond to the scenario in which information acquisition is sufficiently costly. Case c), the more interesting one, does involve information acquisition. Specifically, the optimal decision rule \( m(\theta) \) is not constant, and neither action 1 nor action 0 is ex-ante optimal. Intuitively, this corresponds to the scenario where the prior is not extreme, or the cost of information acquisition is sufficiently low. In case c), where information acquisition is involved in the optimal decision rule, the agent equates the marginal benefit of information with its marginal cost. In doing so, the agent chooses the “shape” of the optimal decision rule \( m(\theta) \) according to the “shape” of the payoff gain \( \Delta u(\theta) \) and her prior \( P \), which process is consistent with the essence of flexible information acquisition.

For an intuitive appreciation, let us consider an extreme case where action 1 is dominant, i.e., almost certainly the payoff gain \( \Delta u (\theta) > 0 \). It is obvious that the agent will always take action 1 regardless of the marginal cost of information acquisition.

In a less extreme case where neither action is dominant, i.e.,
\[
\Pr (\Delta u (\theta) > 0) > 0 \quad \text{and} \quad \Pr (\Delta u (\theta) < 0) > 0 ,
\]
the marginal cost of information acquisition \( \mu \) plays a role.

On the one hand,
\[
\lim_{\mu \to \infty} E \exp (\pm \mu^{-1} \Delta u (\theta)) = 1 .
\]
Hence Proposition 1 predicts that if \( \mu \) is high enough, no information is acquired.
On the other hand, since
\[
\lim_{\mu \to 0} \frac{d}{d\mu^{-1}} \mathbb{E} \exp \left( \mu^{-1} \Delta u (\theta) \right)
= \mathbb{E} \left[ \exp \left( \mu^{-1} \Delta u (\theta) \right) \Delta u (\theta) \right]
= \lim_{\mu \to 0} \int_{\Delta u(\theta) > 0} \exp \left( \mu^{-1} \Delta u (\theta) \right) \Delta u (\theta) \, dP(\theta)
+ \Pr (\Delta u (\theta) = 0) + \lim_{\mu \to 0} \int_{\Delta u(\theta) < 0} \exp \left( \mu^{-1} \Delta u (\theta) \right) \Delta u (\theta) \, dP(\theta)
= +\infty,
\]
we have
\[
\lim_{\mu \to 0} \mathbb{E} \exp \left( \mu^{-1} \Delta u (\theta) \right) > 1.
\]
A similar argument leads to
\[
\lim_{\mu \to 0} \mathbb{E} \exp \left( -\mu^{-1} \Delta u (\theta) \right) > 1.
\]
Therefore, Proposition 1 implies that if the marginal cost of information is sufficiently low, information acquisition must exist. This coincides with our intuition that the agent rationally decides whether to acquire information by cost-benefit analysis.

When neither action is dominant and the marginal cost of information acquisition takes intermediate values, agents find it optimal to acquire some information to make their action (partially, in a random manner) contingent on \( \theta \). This is the case specified by condition (6). Since \( g' \) is strictly increasing, (7) implies that \( m(\theta) \), the conditional probability of choosing action 1, is increasing with respect to the payoff gain \( \Delta u (\theta) \). This too is intuitive. The left hand side of (7) represents the marginal benefit of increasing \( m(\theta) \), while the right hand side is the marginal cost. Therefore, if the agent decides to acquire information, she will equate the marginal benefit with the marginal cost pointwise.

**AN EXAMPLE**

The following example shows how our framework of flexible information acquisition captures the qualitative nature of information.

Let \( \theta \) 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 \( t \leq -\mu^{-1}/2 \). In this case, action 1 (action 0) is superior to action 0 (action 1) ex ante (i.e., \( |t| \) is large) and the cost of acquiring information is relatively high (i.e., \( \mu \) is large). Hence, it is not worth acquiring any information at all. If we let \( t = 0 \), then the agent finds it optimal to acquire some information. According to (7), the optimal
information acquisition strategy \( m (\theta) \) satisfies

\[
\theta / \mu = g' (m (\theta)) - g' (E m (\theta)) ,
\]

where

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

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

\[
E m (\theta) = 1/2 .
\]

Hence

\[
g' (E m (\theta)) = 0
\]

and (8) becomes

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

Therefore,

\[
m (\theta) = \frac{1}{1 + \exp (-\theta / \mu)} .
\]

First note that

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

Step function \( a (\theta) \) captures the agent’s choice under complete information. In this case, the agent can observe the exact value of \( \theta \) since the information cost vanishes.

When \( \mu > 0 \), the best response is characterized by (9). Since information is no longer cost-free, the agent finds it optimal to allow for some mistake in her decision. The probability of a mistake conditional on \( \theta \) is given by

\[
|m (\theta) - a (\theta)| ,
\]

which is decreasing in \(|\theta|\), the “price” of the mistake. Therefore, the agent deliberately acquires information, but trades off the price of mistakes against the cost of information.

Second, parameter \( \mu \) measures the difficulty of acquiring information. Figure II.A shows how \( m (\theta) \) varies with this parameter.
When $\mu = 0$, there is no cost for information acquisition, and the agent’s response is the step function $a(\theta)$. She never makes a mistake. When $\mu$ becomes larger, she starts to compromise on the accuracy of her decision in order to save information cost. Larger $\mu$ leads to flatter $m(\theta)$. Finally, when $\mu$ is extremely large, $m(\theta)$ is almost constant and the agent practically ceases to acquire information.

Third, since the agent’s action is highly sensitive to $\theta$ where slope $|\frac{dm(\theta)}{d\theta}|$ is large, $|\frac{dm(\theta)}{d\theta}|$ reflects her attentiveness around $\theta$. Under this interpretation, Figure II.A reveals that the agent actively acquires information for intermediate values of the fundamental but is rationally inattentive to tail values. This result coincides with our intuition. When $\theta$ is too high (low), the agent should take action 1 (action 0) anyway. Hence the information about $\theta$ on the tails is not so relevant to her payoff. For intermediate values of $\theta$, the payoff gain from action 1 vis-à-vis action 0 depends crucially on the sign of $\theta$. Therefore, the variation of $\theta$ around zero attracts most of the agent’s attention.

B. Justification of the Binary-signal Information Structure

We have posited binary-signal information structures. This subsection justifies this setup on the assumption of flexible information acquisition.

Assumption: i) The decision maker can acquire information (about $\theta$) flexibly by purchasing any information structure $((X, \sigma), \pi)$, where $X$ is the set of realizations of the signal, $\sigma$ is a $\sigma$-algebra on $X$, and $\pi : \Theta \rightarrow \Delta(X)$ specifies a probability measure over $X$ for each state $\theta \in \Theta$; ii) choosing information structure $((X, \sigma), \pi)$ incurs cost $\mu \cdot I(\pi)$, where $I(\pi)$ is the mutual information between the signal and the state, and $\mu > 0$ is the marginal cost of information acquisition.

The first part of Assumption II.B reads that any information desired about the state is available. The second part specifies that the information cost is proportional to the
amount of information conveyed by the signal. The binary-signal information structure is a special case with $X = \{0, 1\}$ and $\pi (1|\theta) = m (\theta)$ (and so $\pi (0|\theta) = 1 - m (\theta)$). For the binary-decision problem with flexible information acquisition, this special case is sufficient. To see this, let $((X, \sigma), \pi)$ be any information structure chosen by the agent. Given $((X, \sigma), \pi)$, the agent optimally chooses an action rule as $a : X \rightarrow [0, 1]$, where $a (x)$ is the probability of taking action 1 upon receiving signal $x$. Let

$$X_1 = \{ x \in X : a (x) = 1 \},$$

$$X_0 = \{ x \in X : a (x) = 0 \},$$

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 $X_{ind}$. Thus $(X_1, X_0, X_{ind})$ partitions $X$. Since the only use of the signal is to make a binary decision, a signal that differentiates 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, being indifferent between action 0 and 1 upon event $X_{ind}$, the agent would rationally expend no effort to distinguish this from other realizations. Hence, the agent always play pure strategies upon receiving the signal. Therefore, the agent always prefers binary-signal information structures.\(^{12}\)

### III. Security Design under Flexible Information Acquisition

#### A. The Basic Setup

We consider a two-period game with two players. One is a seller who owns $N$ assets at period 0. These assets generate verifiable random cash flows $\overrightarrow{\theta} \in \Theta \subset \mathbb{R}^N_+$ in period 1\(^{13}\). The other player is a potential buyer holding consumption goods (money) at period 0. Player $i$’s utility function is given by

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

\(^{12}\) (Woodford 2009) makes a similar argument that the agent only needs to acquire a "yes/no" signal. Here we omit the formula of $I (x)$ but point out two of its properties that guarantee the sufficiency of binary-signal information structures for binary-decision problems. First, coarsening information structures, i.e., pooling some realizations of the signal together to form a new realization, reduces information costs. This property is intuitive, since coarsening reduces informativeness. It implies that the agent acquires at most three signal realizations representing $X_1, X_0$ and $X_{ind}$, respectively. The second property is that the information cost is convex. This is because mixing information structures adds randomness and so reduces informativeness. This property precludes the realization representing $X_{ind}$, i.e., no mixing upon observing the signal. We refer readers to (Cover and Thomas 1991) for details of the general formula of mutual information.

\(^{13}\) 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.
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 $\delta_b > \delta_s$, i.e. we posit that the seller has a better investment opportunity than the buyer. This assumption creates the trading demand. Both parties may benefit from transferring some goods to the seller at date 0 and compensating the buyer with repayment backed by the random cash flows $\tilde{\theta}$ at date 1.

Like (Dang, Gorton and Holmstrom 2010), we assume no information asymmetry at period 0 to focus on the adverse selection resulted from information acquisition. Hence the two agents start with identical information about $\tilde{\theta}$, which is represented by a full support common prior $P$ over $\Theta$. 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 $\tilde{\theta}$, the cash flows of the $N$ assets, is a mapping $s : \Theta \rightarrow \mathbb{R}_+$ such that $\forall \tilde{\theta} \in \Theta, s\left(\tilde{\theta}\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right]$. A contract $(s(\cdot), q)$ is a security $s(\cdot)$ associated with a price $q > 0$. Throughout this paper, we assume one player proposes a take-it-or-leave-it contract $(s(\cdot), q)$ to the counterparty, who then acquires information and decides whether or not to accept. This captures the idea that some agents in the markets for securitized assets are less sophisticated than others and cannot produce private information about the underlying cash flows. This separation between bargaining power and the ability to acquire information also makes our problem tractable.\footnote{If the issuer could produce private information before making the proposal, the signaling game would be much more complicated. The set of possible signals would consist in all contracts, which is a functional space. To the best of our knowledge, this kind of signaling game has been studied only rarely. (Peter M. DeMarzo, Ilan Kremer and Andrzej Skrzypacz 2005) do consider a security design problem in which 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 (Stephen A. Ross 1977), or the fraction of equity retained in (Hayne E Leland and David H Pyle 1977), etc.), or else the issuer designs the security before obtaining information (e.g., (DeMarzo and Duffie 1999), (Biais and Mariotti 2005)).}

First we study the case in which the seller designs the contract and the buyer acquires information. We then highlight two factors driving the unique optimality of issuing debt and finally show that debt remains uniquely optimal when the roles of the buyer and the seller are reversed.

### B. Optimal Contract when the Seller Designs

We denote the seller’s optimal contract by $\left(\bar{s}(\tilde{\theta}), q^*\right)$. The strategic situation between seller and buyer is a dynamic Bayesian game with sequential moves. Concretely, the seller first designs the contract, then the buyer acquires information according to the contract and decides whether or not to accept. We characterized the general optimal decision rule for the agent who acquires information in Proposition 1. So we can apply the results of Proposition 1 to the buyer’s decision problem given the seller’s contract, and then solve the seller’s optimal contract by backward induction. To distinguish this game from the general decision problem discussed above, we denote the buyer’s optimal...
information acquisition strategy as \( m_{s,q}(\theta) \), given payoff gain
\[
\delta_b \cdot s(\theta) - q
\]
derived from contract \((s(\theta), q)\). The buyer’s optimal information acquisition strategy given the seller’s optimal contract \((s^*(\theta), q^*)\) will be denoted by \( m_{s^*,q^*}(\theta) \).

To streamline the analysis, we summarize the problems for both seller and buyer as discussed above, and formally define the equilibrium of this model as follows.

**DEFINITION 1:** The sequential equilibrium is defined as a collection of the seller’s optimal contract \((s^*(\theta), q^*)\) and the buyer’s optimal information acquisition strategy \( m_{s^*,q^*}(\theta) \) such that

i). \[
m_{s^*,q^*} = \arg \max_{m \in M} \mathbb{E} \left[ \left( \delta_b \cdot s^*(\theta) - q^* \right) \cdot m(\theta) \right] - \mu \cdot I(m);
\]

ii). \[
(s^*, q^*) \in \arg \max_{s(\theta) \in [0, \sum_{a} \theta_a], q \geq 0} \mathbb{E} \left[ m_{s,q}(\theta) \cdot \left( q - \delta_s \cdot s(\theta) \right) \right].
\]

By Proposition 1, there are three possible cases pertaining to the buyer’s behavior, given the seller’s optimal contract. First, the buyer may optimally choose not to acquire any information and accept the seller’s contract directly. Second, the buyer may optimally acquire some information induced by the seller’s optimal contract, and then accept with positive probability (but less than one). Third, the buyer may simply reject the contract without acquiring information. This corresponds to the seller’s outside option of proposing nothing and so raising no liquidity. Hence, there is no need to consider the seller’s individual rationality condition.

In what follows, we first show that in equilibrium this last case does not occur, since both seller and buyer should have gained from the trade. Then we characterize the seller’s optimal contract for the first two types of equilibrium.

Let \( W(s, q) \) denote the seller’s expected payoff from proposing contract \((s, q)\), i.e.,

\[
W(s, q) = \mathbb{E} \left( m_{s,q}(\theta) \cdot \left[ q - \delta_s \cdot s(\theta) \right] \right),
\]

where the expectation \( \mathbb{E}(\cdot) \) is taken with respect to \( \theta \) under common prior \( P \). Facing contract \((s, q)\), the buyer’s unconditional probability of acceptance is given by
\[
\overline{p}_{s,q} = \mathbb{E} m_{s,q}(\theta).
\]

**PROPOSITION 2:** \( \overline{p}_{s^*,q^*} > 0 \), i.e., trade happens with positive probability.

**PROOF:** See Appendix A.
The key to the proof is to show that the seller always has a positive expected payoff by proposing a debt security. Hence her optimal contract too necessarily generates a positive expected payoff, which can be achieved only through a successful trade. Despite adverse selection, the seller always prefers trade, because she holds all the bargaining power, and can minimize the adverse effect of information acquisition by designing the right contract and thus obtaining the gains from trade.

According to Proposition 2, we only need to consider case i): $\overline{\theta}_{s^*, q^*} = 1$, in which the buyer does not acquire information; and case ii): $\overline{\theta}_{s^*, q^*} \in (0, 1)$, in which the buyer does acquire information. We first study the seller’s optimal contract in case i).

**Optimal Contract without Inducing Information Acquisition**

In this subsection, we consider the case in which the seller’s optimal contract is simply accepted by the buyer without information acquisition. Concretely, this means $\Pr(m_{s, q}(\theta) = 1) = 1$. We first consider the buyer’s information acquisition problem, given the seller’s contract. Then we characterize the seller’s optimal contract.

By Proposition 1 and condition (4), any contract $(s, q)$ that is accepted by the buyer without information acquisition must satisfy

$$
\mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s \left( \overline{\theta} \right) - q \right) \leq 1.
$$

In particular, if the left hand side of the inequality is strictly less than one, the seller could always raise price $q$ in order to increase the expected payoff, leaving the buyer’s response unchanged. Hence, in equilibrium this inequality must bind, i.e.,

$$
q = -\mu \ln \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s \left( \overline{\theta} \right) \right).
$$

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

$$
\mathbb{E} \left[ q - \delta_b \cdot s \left( \overline{\theta} \right) \right] = -\mu \ln \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s \left( \overline{\theta} \right) \right) - \delta_s \cdot \mathbb{E} s \left( \overline{\theta} \right).
$$

Hence the seller’s problem can be formalized as

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

subject to the limited liability constraint

$$
(13) \quad s \left( \overline{\theta} \right) \in \left[ 0, \sum_{n=1}^{N} \theta_n \right].
$$

**Proposition 3:** If the seller’s optimal contract $(s^*, q^*)$ induces the buyer to accept
it without acquiring information, it must be a debt security

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

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

\[ D^* = D \left( q^* \right) = \mu \tilde{\delta}^{-1} [\ln \delta_b - \ln \delta_i] + \delta_b^{-1} q^*, \]

and \( q^* > 0 \) is the unique fixed point of

\[ h (q) = -\mu \ln \mathbb{E} \exp \left( -\mu^{-1} \tilde{\delta} \cdot \min \left( \sum_{n=1}^{N} \theta_n, D (q) \right) \right). \]

**Proof:** See Appendix A.

Debt is the unique optimal contract. The intuition here is clear. We are looking at the asset-backed securities that make the buyer break even between acquiring and not acquiring information. Thanks to flexible information acquisition, any mean-preserving spread of the optimal security would induce the buyer to acquire information. In other words, debt is the least information-sensitive security that gives the seller the desired expected payoff. More specifically, it is the flat part of the debt that reduces the buyer’s incentive to acquire information. As a result, a debt security is the seller’s unique optimal choice.

We can also see this point from another angle. We can show that the seller’s optimal contract, as a debt characterized by Proposition 3, second order stochastic dominates any other asset-backed securities with the same expected payoff. That said, for all possible asset-backed securities that make the buyer indifferent between acquiring and not acquiring information, debt delivers the highest expected payoff to the seller. This is exactly the logic underlying the seller’s optimization problem. The face value of the debt is determined in such a way that it delivers the greatest expected payoff, without inducing the buyer to acquire information.

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

\[ D^* > \mu \tilde{\delta}^{-1} \cdot \left[ \ln \delta_b - \ln \delta_i \right]. \]

Hence if the maximal cash flow

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

then the optimal security is actually the pool of all assets. This can occur when the seller has an extremely good investment opportunity compared with the buyer (i.e., \( \ln \delta_b - \)
ln \( \delta_x \gg 1 \) or when it is too hard for the buyer to acquire information (i.e., \( \mu \gg 1 \)). In the extreme case, when the buyer cannot acquire any information (i.e., \( \mu \rightarrow \infty \)), the seller just sells the asset pool at price

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

and reaps the maximum trading surplus

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

Another interesting observation is implied by equation (12), i.e.,

\[
q^* = -\mu \ln \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s^* \left( \bar{\theta} \right) \right) < \delta_b \cdot \mathbb{E} s^* \left( \bar{\theta} \right),
\]

which follows Jensen’s inequality. Since the offer induces no information acquisition, both parties remain symmetrically informed and theoretically the seller should have charged \( \delta_b \cdot \mathbb{E} s^* \left( \bar{\theta} \right) \). Actually, however, the seller finds it optimal to charge a lower price \( q^* \) so as to "bribe" the buyer not to acquire information.

**OPTIMAL CONTRACT WITH INFORMATION ACQUISITION**

Now we characterize the optimal contract for the seller that does induce the buyer to acquire information and accept the contract with positive probability (but less than one). Concretely, this means

\[
\text{Pr}(m_{s,q}(\theta) \in (0, 1)) = 1 .
\]

Again, according to Proposition 1, any contract \((s, q)\) that induces the buyer to acquire information must satisfy

\[
\mathbb{E} \exp \left( \mu^{-1} \left[ \delta_b \cdot s \left( \bar{\theta} \right) - q \right] \right) > 1
\]

and

\[
\mathbb{E} \exp \left( -\mu^{-1} \left[ \delta_b \cdot s \left( \bar{\theta} \right) - q \right] \right) > 1 ,
\]

where the expectation \( \mathbb{E}(\cdot) \) is taken with respect to \( \theta \) under common prior \( P \). As noted above, when conditions (14) and (15) are satisfied, for the buyer neither accepting nor rejecting is optimal, ex ante. So the buyer finds it optimal to first acquire some information and then make the decision. In other words, in this case the proposed contract will induce information acquisition.
Given such a contract \((s, q)\), Proposition 1 and condition (7) also prescribe that the buyer’s optimal decision rule of information acquisition \(m_{s,q}\) is characterized by

\[
\delta_b \cdot s\left(\overrightarrow{\theta}\right) - q = \mu \cdot \left[ g'\left(m_{s,q}\left(\overrightarrow{\theta}\right)\right) - g'\left(\overrightarrow{p_{s,q}}\right)\right],
\]

where

\[
\overrightarrow{p_{s,q}} = \mathbb{E}m_{s,q}\left(\overrightarrow{\theta}\right)
\]
is the buyer’s unconditional probability of accepting the contract. Condition (16) concludes the buyer’s decision problem, given the proposed contract \((s, q)\).

We derive the seller’s optimal contract by backward induction. Taking into account of buyer’s response \(m_{s,q}\), the seller chooses \((s, q)\) to maximize the expected payoff

\[
W\left(s, q\right) = \mathbb{E}\left(m_{s,q}\left(\overrightarrow{\theta}\right) \cdot \left[q - \delta_s \cdot s\left(\overrightarrow{\theta}\right)\right]\right)
\]
subject to (14), (15), (16) and the limited liability constraint

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

We derive the seller’s optimal contract \((s^*, q^*)\) through calculus of variations. Specifically, we characterize how the seller’s expected payoff responds to the perturbation of the optimal security \(s^*\).

Let \(s\left(\overrightarrow{\theta}\right) = s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right)\) be an arbitrary perturbation of \(s^*\). Note that the buyer’s optimal decision rule \(m_{s,q}\) appears in the seller’s expected payoff \(W\left(s, q\right)\), and it is implicitly determined by the proposed security \(s\) through the functional equation (16). Hence, we need first characterize how \(m_{s,q}\) varies with respect to the perturbation of \(s^*\). This is given by the following lemma.

**LEMMA 1:** For any perturbation \(s\left(\overrightarrow{\theta}\right) = s^*\left(\overrightarrow{\theta}\right) + \alpha \cdot \varepsilon\left(\overrightarrow{\theta}\right)\), the buyer’s decision rule \(m_{s,q}^*\) changes according to:

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

\[
+ \frac{\left[ g''\left(m_{s,q}^*\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \mu^{-1} \delta_b \mathbb{E}\left[ g''\left(m_{s,q}^*\left(\overrightarrow{\theta}\right)\right)\right]^{-1} \varepsilon\left(\overrightarrow{\theta}\right)}{\left[ g''\left(\overrightarrow{p_{s,q}^*}\right)\right]^{-1} - \mathbb{E}\left[ g''\left(m_{s,q}^*\left(\overrightarrow{\theta}\right)\right)\right]^{-1}}.
\]

\[\text{Note that by Proposition 1 neither (14) nor (15) should be binding for the optimal contract; otherwise, the buyer would not acquire information.}\]

\[\text{Since we are interested in the shape of the optimal security } s^* \text{ rather than the optimal price } q^*, \text{ we focus on the perturbation of } s^*.\]
Proof: See Appendix A.

The first term of the right hand side of (18) is the buyer’s local response to $\varepsilon \left( \overline{\theta} \right)$. It is of the same sign as the perturbation $\varepsilon \left( \overline{\theta} \right)$. When the repayment at state $\overline{\theta}$ is higher, the buyer is more likely to accept the offer in this state. The second term measures the buyer’s average response to perturbation $\varepsilon \left( \overline{\theta} \right)$ over all states. It is straightforward to verify that the denominator is positive, due to Jensen’s inequality. As a result, if the perturbation increases the buyer’s repayment on average over all states, there is also a greater likelihood of accepting the contract.

Now we can calculate the variation of the seller’s expected payoff $W \left( s, q^* \right)$, according to (11). Taking the derivative of $W \left( s, q^* \right)$ with respect to $\alpha$ at $\alpha = 0$ leads to

$$
\frac{dW \left( s, q^* \right)}{d\alpha} \bigg|_{\alpha=0} = \mathbb{E} \left[ \left( \frac{dm_{s,q^*} \left( \overline{\theta} \right)}{d\alpha} \right) \left[ q^* - \delta^* \cdot s^* \left( \overline{\theta} \right) \right] \right] - \delta^* \cdot \mathbb{E} \left[ m_{s,q^*} \left( \overline{\theta} \right) \varepsilon \left( \overline{\theta} \right) \right].
$$

Substitute (18) into (19) and we get

$$
\frac{dW \left( s, q^* \right)}{d\alpha} \bigg|_{\alpha=0} = \mathbb{E} \left[ r \left( \overline{\theta} \right) \cdot \varepsilon \left( \overline{\theta} \right) \right],
$$

where

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

and

$$
w^* = \frac{\mathbb{E} \left[ \left( q^* - \delta^* \cdot s^* \left( \overline{\theta} \right) \right) \left[ g'' \left( m_{s,q^*} \left( \overline{\theta} \right) \right) \right]^{-1} \right]}{\left[ g'' \left( \overline{\theta} \right) \right]^{-1} - \mathbb{E} \left[ g'' \left( m_{s,q^*} \left( \overline{\theta} \right) \right) \right]^{-1}}.
$$

Note that $w^*$ is a constant that does not depend on $\theta$ and will be endogenously determined in equilibrium. Further, $r \left( \overline{\theta} \right)$ is the Fréchet derivative of the seller’s expected payoff $W \left( s, q^* \right)$ at $s^*$, which measures the marginal contribution of any perturbation to the seller’s expected payoff when the contract is optimal. Specifically, the first term of (21) is the direct contribution of the perturbation of $s^* \left( \overline{\theta} \right)$ disregarding the variation of $m_{s,q^*} \left( \overline{\theta} \right)$; the second term measures the indirect contribution through the variation of $m_{s,q^*} \left( \overline{\theta} \right)$. This Fréchet derivative $r \left( \overline{\theta} \right)$ plays an important role in shaping the seller’s optimal contract.

To further characterize the optimal contract, we discuss the Fréchet derivative $r \left( \overline{\theta} \right)$ in detail. Let

$$
A_0 = \left\{ \overline{\theta} \in \Theta : \overline{\theta} \neq \overline{\theta}^*, s^* \left( \overline{\theta} \right) = 0 \right\},
$$
A_1 = \left\{ \theta \in \Theta : \theta \neq \theta^*, s^*\left(\theta^*\right) \in \left(0, \sum_{n=1}^{N} \theta_n\right) \right\}

and

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

Clearly, \{A_0, A_1, A_2\} is a partition of \Theta \setminus \left\{ \theta^* \right\}. Since \(s^*\) is the optimal security, we have

\[
\frac{dW(s, q^*)}{da}\bigg|_{a=0} \leq 0
\]

for any feasible\(^{17}\) perturbation \(\varepsilon\left(\theta^*\right)\). Hence, condition (20) implies

(22)

\[
r\left(\theta^*\right) \begin{cases} 
\leq 0 & \text{if } \theta^* \in A_0 \\
= 0 & \text{if } \theta^* \in A_1 \\
\geq 0 & \text{if } \theta^* \in A_2
\end{cases}
\]

Since \(g\) is strictly convex, \(r\left(\theta^*\right) \cdot g''\left(m_{s^*, q^*}\left(\theta^*\right)\right)\) is of the same sign as \(r\left(\theta^*\right)\). Thus (22) can be rewritten as

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

(23)

\[
\begin{cases} 
\leq 0 & \text{if } \theta^* \in A_0 \\
= 0 & \text{if } \theta^* \in A_1 \\
\geq 0 & \text{if } \theta^* \in A_2
\end{cases}
\]

Recall condition (16), given the optimal contract \((s^*, q^*)\), the buyer’s best response \(m_{s^*, q^*}\left(\theta^*\right)\) is characterized by

(24)

\[
\delta_b \cdot s^* \left(\theta^*\right) - q^* = \mu \left[ g' \left(m_{s^*, q^*} \left(\theta^*\right)\right) - g' \left(p_{s^*, q^*}\right)\right],
\]

where

\[
p_{s^*, q^*} = \mathbb{E}_{m_{s^*, q^*} \left(\theta^*\right)}
\]

is the buyer’s unconditional probability of accepting the optimal contract \((s^*, q^*)\). Conditions (23)\(^{18}\) and (24) as a system of functional equations jointly determine the optimal

\(^{17}\)A perturbation \(\varepsilon\) is feasible with respect to \(s^*\) if \(\exists \alpha > 0\), s.t. \(\forall \theta^* \in \Theta, s^*\left(\theta^*\right) + \alpha \cdot \varepsilon \left(\theta^*\right) \in \left[0, \sum_{n=1}^{N} \theta_n\right]\).

\(^{18}\)One may object that Equation (23) is just the first order condition of the seller’s optimization problem. It only
security $s^*$ when the optimal contract induces information acquisition.

We solve the system of equations to get the seller’s optimal contract $s^*$ and the buyer’s associated optimal decision rule of information acquisition $m^*, q^*$. Note that the values of $P^*, q^*$ and $w^*$ will also be endogenously determined in equilibrium. To facilitate the analysis and economize on notations, consider two equivalent equations with respect to variables $m$ and $s$, in which $m$ stands for $m^{*, q^*}(\theta)$ and $s$ stands for $s^*(\theta)$.

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

and

\begin{equation}
\delta_b \cdot s - q^* = \mu \cdot \left[ g'(m) - g'(P^*, q^*) \right].
\end{equation}

Let $m = f_1(s)$ and $m = f_2(s)$ be the two continuous functions implicitly defined by (25) and (26), respectively.

Before solving the system of equations, it will be helpful to gain some insight into the shape of the optimal security by examining a problem without the limited liability constraint, in which case, $r(\delta) = 0$ and thus by definition,

\[ m^{*, q^*}(\delta) = f_1(s^*(\delta)) \]

Also note that

\[ m^{*, q^*}(\delta) = f_2(s^*(\delta)) \]

holds for all $\delta \in \Theta$. Hence, without the limited liability constraint, the pair of values $(m^{*, q^*}, s^*)$ satisfies both

\[ m^{*, q^*} = f_1(s^*) \]

and

\[ m^{*, q^*} = f_2(s^*) \].

Since $\left[ m \cdot g''(m) \right] > 0$ and $g''(m) > 0$, it is obvious that $f_1'(s) < 0$ and $f_2'(s) > 0$. Therefore, the curves $m = f_1(s)$ and $m = f_2(s)$ intersect at most once, which suggests that in the absence of the limited liability constraint $s^*(\delta)$ must be a constant. This observation reflects the seller’s objective of minimizing the buyer’s incentive to acquire information. By maintaining a constant repayment level, the seller ensures that the buyer gains nothing by distinguishing different states of the underlying cash flow; thus eliminates any incentive to acquire information.

In equilibrium, however, the limited liability constraint does bind. Otherwise, the seller will always propose $s^*(\delta) = s^*$ at price $q^* = \delta_b \cdot s^*$, and achieve a gain $(\delta_b - \delta_s) \cdot s^*$ from trading, which can be made arbitrarily large by increasing $s^*$. Of characterizes the critical points. In principle, we should characterize the largest critical point, but our argument holds for any critical point, so our results are not open to this critique.
course this contradicts the finiteness of the collateral. Proposition 4 examines the lower boundary of the limited liability constraint \( s(\overrightarrow{\vartheta}) \geq 0 \).

**PROPOSITION 4:** *In equilibrium, \( \Pr(A_0) = 0 \), where \( A_0 = \{ \overrightarrow{\vartheta} \in \Theta : \overrightarrow{\vartheta} \neq \overrightarrow{0}, s^*(\overrightarrow{\vartheta}) = 0 \} \).*

**Proof:** See Appendix A.

By this proposition, \( s(\overrightarrow{\vartheta}) \) never reaches its lower boundary. This reflects the seller’s incentive to trade. At the boundary \( s(\overrightarrow{\vartheta}) = 0 \), while raising \( s(\overrightarrow{\vartheta}) \) decreases the seller’s payoff, it increases the probability of trading even more sharply. Thus on average the seller gains by deviating from the lower boundary. As a result, equity residual or call options are not optimal to raise liquidity.

This analysis suggests that the optimal security \( s^*(\overrightarrow{\vartheta}) \) is either a constant or reaches the upper boundary of the limited liability constraint. Figure III.B shows a typical security that satisfies these conditions.

![A Typical Security](image)

A debt contract definitely satisfies these conditions. Moreover, we show that the optimal security must be a debt.

**PROPOSITION 5:** *If the seller’s optimal contract induces the buyer to acquire information, it must be a debt, i.e., \( s^*(\overrightarrow{\vartheta}) = \min(\sum_{n=1}^{N} \theta_n, D^*) \).*

**Proof:** See Appendix A.

Together with Propositions 2 and 3, Proposition 5 enables us to conclude that issuing debt, i.e., pooling the assets and issuing a senior tranche, is always the unique optimal way to raise liquidity. Pooling follows directly from the upper boundary of the limited liability constraint, reflecting the seller’s desire to trade (i.e., \( \delta_i < \delta_b \)). As a complement
to the conventional wisdom, pooling does not necessarily derive from risk diversification, since we posit that both agents are risk-neutral.

Compared with the typical security in Figure III.B, the unique optimality of debt securities reflects not only the seller’s desire to trade but also her incentive to minimize the buyer’s information acquisition. This is achieved by maintaining a constant repayment level. However, if the underlying cash flows become too low to support that level, \( s^* \left( \overline{\theta} \right) \) reaches the upper boundary of the limited liability constraint and is equal to \( \sum_{n=1}^{N} \theta_n \).

In contrast to the non-uniqueness result of (Dang, Gorton and Holmstrom 2010), we show the unique optimality of debt thanks to our flexible information acquisition framework. In (Dang, Gorton and Holmstrom 2010), only two extreme information structures are available while infinite forms of securities can be designed, which inevitably makes some securities indistinguishable. In our framework, with flexibility, the variety of information structures matches that of potential securities. So the uniqueness of the standard debt can be guaranteed. Quasi-debt is no longer optimal in our model. By flattening the uneven tail above the price of a quasi-debt, not only can the buyer’s information cost be saved but the potential loss of trade due to adverse selection can be mitigated. The resulting surplus can be used by the seller to the advantage of both parties; ultimately therefore, it permits better liquidity provision. 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 having to assume a sufficiently large number of underlying assets, as in (DeMarzo 2005)\(^{19}\).

In addition, while most models in this literature depend on specific distributional assumptions about cash flows, our qualitative result does not. Since the stochastic interdependence among the underlying assets could be complex and violate such assumptions, our model offers a better explanation for the prevalence of pooling and tranching in financial markets.

The security design literature usually assumes the Monotone Likelihood Ratio Property (MLRP) or similar hypotheses on the information structure to guarantee a meaningful result. Our framework justifies this assumption by endogenizing the information structure. By Proposition 5, the optimal security \( s^* \left( \overline{\theta} \right) \) is non-decreasing in the sum of cash flows. Proposition 1 implies that the best information structure \( m_{s^*,q^*} \left( \overline{\theta} \right) \) is increasing in the payoff gain \( \delta \cdot s^* \left( \overline{\theta} \right) - q^* \). Hence \( m_{s^*,q^*} \left( \overline{\theta} \right) \) is also non-decreasing in the sum of the cash flows. Therefore, the larger the sum of cash flows, the greater the probability that the buyer will receive a signal to accept. This can be interpreted as a generalized MLRP for multi-dimensional states.

To facilitate the analysis, the security design literature usually restricts 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 impose such a restriction, but

\(^{19}\) (DeMarzo 2005) shows that the benefit of pooling achieves a theoretical maximum as the number of underlying assets approaches infinity.
show nevertheless that the optimal security naturally turns out to be non-decreasing.

Finally, (Dang, Gorton and Holmstrom 2010) get the debt contract as uniquely optimal when their fixed information cost is zero. This can be taken as a special case of our own model in which \( \mu \), the marginal cost of information acquisition, becomes nil. Another special case worth noting is \( \mu \to \infty \), in which the seller can sell any security without inducing information acquisition. So selling the asset pool, i.e., \( s^*(\overrightarrow{\theta}) = \sum_{n=1}^{N} \theta_n \), is uniquely optimal. This corresponds to the debt security with face value above the highest possible aggregate cash flow.

**Understanding the Origin of Uniqueness**

For readers familiar with CSV, a natural question regards the uniqueness of the optimal contract. Both (Townsend 1979) and (Dang, Gorton and Holmstrom 2010) employ CSV. Why, then, does the former but not the latter get debt as uniquely optimal? In the previous subsection, we attributed the non-uniqueness in (Dang, Gorton and Holmstrom 2010) to the rigidity of CSV. This argument is correct when the comparison is with our model, but not fully convincing when (Townsend 1979) is also considered. To analyze the different results in (Dang, Gorton and Holmstrom 2010), (Townsend 1979) and our own model, let us first highlight the essence of flexibility. In principle, general flexible choice (not necessarily restricted to information acquisition alone) enables an economic agent to make state-contingent responses. In other words, the agent can make one best response in one state, another best response in another state. In all three of the models under consideration, the contract designer has flexibility, in that she can assign state-contingent repayment by designing any form of security. What matters in shaping the different results on uniqueness is the potential flexibility of the other party. By comparing these three models, we can see that the origin of the uniqueness lies not only in flexibility as such, but in symmetric flexibility. Here, symmetric flexibility requires that both parties to a potential trade have the same level of flexibility.

In our framework, ex-ante symmetric information (in the form of two-sided ignorance) prevents the buyer from making a state-contingent choice if she follows the CSV approach. But in our framework the buyer can choose the state-contingent probability (i.e., \( m(\overrightarrow{\theta}) \)) of accepting the offer; that is, she can engage in flexible information acquisition. In this sense the buyer has as much flexibility as the seller. Given this symmetry of flexibility, the uniqueness of the optimal contract is guaranteed. In (Dang, Gorton and Holmstrom 2010), by contrast, the buyer can only take the conventional CSV approach to information acquisition with only two options, namely, acquiring a signal or not. Moreover, ex-ante two-sided ignorance precludes conditioning the buyer’s action on any private information the seller may have. Hence in (Dang, Gorton and Holmstrom 2010) the CSV keeps the buyer from making state-contingent decisions. The desired symmetric flexibility therefore fails, and so, as a consequence does the uniqueness of the optimal contract. Interestingly, (Townsend 1979) also uses the CSV approach with two options to model information acquisition (to audit or not), but the unique optimality of the standard debt contract still emerges. Why is this the case? Unlike (Dang, Gor-
ton and Holmstrom 2010) and our framework, (Townsend 1979) gives the entrepreneur an information advantage over the lender, in that the former knows the realized profit of the project. Thanks to the revelation principle, the lender who acquires information in the interim stage can decide whether to audit or not in any state, based on the truth told by the entrepreneur. In other words, although the lender in (Townsend 1979) still only has two information acquisition options, like the buyer in (Dang, Gorton and Holmstrom 2010), the two options are state-contingent in (Townsend 1979) but not in (Dang, Gorton and Holmstrom 2010). Therefore, the symmetry of flexibility is still established in (Townsend 1979), and the uniqueness of the optimal contract is ensured as well. Figure III.B shows the relation among these three models.

<table>
<thead>
<tr>
<th>OUR MODEL</th>
<th>flexible info. acquisition $\rightarrow$ state-contingent response $m(\theta)$</th>
<th>ex-ante symmetric info./double-sided ignorance</th>
</tr>
</thead>
<tbody>
<tr>
<td>DANG, GORTON &amp; HOLMSTROM (2010)</td>
<td>CSV: rigid info. acquisition</td>
<td>ex-ante symmetric info./double-sided ignorance</td>
</tr>
<tr>
<td>TOWNSEND (1979)</td>
<td>CSV: rigid info. acquisition</td>
<td>asymmetric info. + revelation principle $\rightarrow$ state-contingent audit</td>
</tr>
</tbody>
</table>

**HOW OUR MODEL RELATES TO DANG, GORTON & HOLMSTROM (2011) AND TOWNSEND (1979)**

**Two Main Factors Driving the Optimality of Debt**

This subsection discusses two main factors that drive our results. We show that in their absence issuing debt may not be optimal.

The first feature of our model is fixed aggregate risk. Before designing the contract, the seller already owns assets $\overline{\theta}$. Hence the aggregate risk owned by seller and buyer is invariant with respect to the success or failure of the transaction. This fixed aggregate risk produces conflicting interests between the two parties: the buyer gains by any information acquisition, but at the expense of the seller through adverse selection. That is, the buyer tries to get information that can enable her to reject offers when the repayment is lower than the price and accept when it is higher. But any particular quantity or quality of information does not affect the aggregate risk.
The importance of this factor can be seen clearly in our derivation of the optimal security. 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 trade. Conditional on $\vec{\vartheta}$, the buyer’s and seller’s payoff gains are

$$\delta_b \cdot s(\vec{\vartheta}) - q$$

and

$$q - \delta_s \cdot s(\vec{\vartheta})$$

respectively. Neither of these depends explicitly on $\vec{\vartheta}$. The future cash flows $\vec{\vartheta}$ can affect incentives only through the security $s(\vec{\vartheta})$. This is why we can define the functions $m = f_1(s)$ and $m = f_2(s)$ rather than $m = f_1(s, \vec{\vartheta})$ or $m = f_2(s, \vec{\vartheta})$ in (25) and (26). The simple shape of debt follows from this independence of $f_1$ and $f_2$ on $\vec{\vartheta}$.

To classify this point, we consider a similar problem with variable aggregate risk. The seller is an entrepreneur who wants to raise capital $k$ for a project that generates future cash flow $\theta$. As before, the entrepreneur designs a security $s(\theta)$ and proposes a take-it-or-leave-it offer $(s, k)$ to an investor, equivalent to the buyer who acquires information. The entrepreneur’s project gets funded and generates future cash flow $\theta$ only if the investor accepts the offer. Hence, the aggregate risk depends on whether the transaction succeeds or not. In this case, the buyer’s payoff gain remains unchanged but the seller’s becomes

$$\delta_s \cdot (\theta - s(\theta))$$

which depends explicitly on $\theta$. As a result, we have $m = f_1(s, \theta)$ rather than $m = f_1(s)$, and the debt contract, which has a flat tail, is no longer optimal. Unlike the case of fixed aggregate risk, even if $s(\theta)$ is off the boundaries of the limited liability constraint, the entrepreneur chooses to vary $s(\theta)$ in order to induce the investor to acquire some information. Since aggregate risk is variable, a certain type of information acquisition benefits the entrepreneur and the investor taken together. It helps screen out bad projects, those with low cash flows.\(^{20}\)

The second factor that drives our results is homogeneous information acquisition. That is, no state is more difficult than others in terms of information acquisition. This property, which follows the approach of rational inattention, gives rise to the independence of our qualitative results from the stochastic interdependence among the underlying assets. Recalling the binary decision problem in Section II, the decision maker’s optimal strategy

---

\(^{20}\)In fact, this is a story of venture finance. Besides raising capital, the security design also serves to elicit information from the investors (venture capitalists), who are often more knowledgeable about the venture’s market prospects. A working paper of the author ((Ming Yang and Yao Zeng 2012)) addresses this problem in a variable aggregate risk framework and shows that the two parties’ incentives can be aligned rather than conflicting if securities such as convertible preferred stock and convertible debt are adopted.
$m$ is characterized by equation (7)

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

where

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

The right hand side of equation (7) is the Fréchet derivative\(^{21}\) of information cost. It does not depend explicitly on $\theta$. This is the homogeneity we referred to. As an 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 (26). This dependence reflects the buyer’s varying difficulty in discerning different states. Hence the optimal contract, unlike debt, may not have a flat tail.

For further insight, we present an example with non-homogeneous information cost. 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 $\theta \in (a, 1]$, while for $\theta \in [0, a]$, the information can be acquired at marginal cost

$$\frac{\mu}{\Pr (\theta \in [0, a])} .$$

Let $\delta_\theta = 1$ and the seller’s optimal contract be $(s, q)$. Given this contract, the buyer’s optimal strategy $m (\theta)$ is characterized by

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

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

$$p_1 = \frac{\int_{[0,a]} m (\theta) dP (\theta)}{\Pr (\theta \in [0, a])} .$$

\(^{21}\)For the readers not familiar with this concept, the Frechet derivative can be thought of as the gradient of the cost function.
For $\theta \in (a, 1)$, the buyer accepts the offer if and only if $s(\theta) - q \geq 0$, so we must have $s(\theta) = q$.

However, information remains costly in region $[0, a]$, so by our previous argument a debt contract is optimal within this region. Finally (see Figure 1), the optimal contract on interval $[0, 1]$ is no longer a debt security.

C. The Distribution of Bargaining Power

This subsection shows that our results are not affected by the distribution of bargaining power between seller and buyer. We exchange their roles, positing that the buyer proposes the contract and the seller acquires information and makes the decision. We specify the environment and summarize the results in three propositions. Most of the proofs are analogous to those in our main model, and can accordingly be omitted here.

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

$$W(s, q) = \mathbb{E} \left( m_{s,q} \left( \bar{\theta} \right) \cdot \left[ \delta_b \cdot s \left( \bar{\theta} \right) - q \right] \right).$$

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

$$\bar{p}_{s^*,q^*} = \mathbb{E} m_{s^*,q^*} \left( \bar{\theta} \right)$$
be the corresponding probability of the trade taking place.

**PROPOSITION 6:** $\overline{p}_{s^*,q^*} > 0$, i.e., trade occurs with positive probability. 

**Proof:** See Appendix A.

**PROPOSITION 7:** If the buyer’s optimal contract induces the seller to always accept it without acquiring information, it must be a debt security

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

with price $q^*$, where

$$D^* = \mu \delta_{x}^{-1} \cdot \left[\ln \delta_{b} - \ln \delta_{s}\right] + \delta_{x}^{-1} q^*,$$

and $q^*$ is the unique fixed point of

$$h\left(q\right) = \mu \ln \mathbb{E} \exp\left(\mu^{-1} \delta_{x} \cdot \min\left(\sum_{n=1}^{N} \theta_n, \mu \delta_{x}^{-1} \cdot \left[\ln \delta_{b} - \ln \delta_{s}\right] + \delta_{x}^{-1} q\right)\right).$$

**PROOF:** The proof is practically identical to that of Proposition 3.

**PROPOSITION 8:** If the buyer’s optimal contract induces the seller to acquire information, it must be a debt security 

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

**Proof:** The proof is practically identical to that of Proposition 5.

Propositions 3, 5, 7 and 8 show that it is always optimal to issue a debt backed by the entire pool of assets, no matter which party designs the contract and which acquires information. In the light of our previous analysis, this result is intuitive. Switching the roles of buyer and seller does not alter either the aggregate risk or the homogeneity of information acquisition cost.

**IV. Conclusions and Discussion**

This paper studies liquidity provision in a framework of flexible information acquisition. There is no information asymmetry before bargaining. The buyer has expertise in acquiring information about the fundamental as in the rational inattention approach. She collects the most payoff-relevant information according to the contract proposed, which may endogenously generate adverse selection. Hence, the seller deliberately designs the security to induce the buyer to acquire information least harmful to the seller’s interest. Issuing debt, i.e., a senior tranche backed by the whole pool of assets, is shown to

---

$^{22}$However, switching the roles of the buyer and the seller does affect the face value and price of the debt, and hence both parties’ expected payoffs.
be the uniquely optimal way to raise liquidity, regardless of stochastic interdependence among the underlying assets and the distribution of bargaining power between parties. Compared with the security design literature, our results are clearer. We neither restrict coverage to non-decreasing securities nor impose various assumptions on information structures, such as MLRP. Instead, these properties of the optimal security are justified in equilibrium. Our results are driven by two main factors: fixed aggregate risk and homogeneous information cost, without which the debt may not be optimal.

The role of fixed aggregate risk sheds light on a general classification of information, namely, determining what information is socially valuable and what is not. In particular, flexibility enables economic agents to acquire these two types of information separately, which generates different welfare implications. At the level of the society, acquisition of information that is not socially valuable not only wastes social resources but also produces endogenous adverse selection, reducing social welfare; hence, the proper organizational form of the society should deter the acquisition of such information. The acquisition of socially valuable information, however, increases social welfare and thus should be encouraged. Our main model with fixed aggregate risk resembles an exchange economy in which none of the information is socially valuable, so that debt is optimal because it deters information acquisition better than other contract types. On the other hand, our example with variable aggregate risk resembles a production economy in which some information is socially valuable, as it helps prevent investing in bad states. Consequently, the acquisition of this information should be encouraged, so that debt may not always be the optimal contract. This classification of information also provides a new perspective on coexistence of debt and equity, the main forms of financing contracts. For start-ups and high risk projects, equity-like securities could be more desirable, in that it encourages the acquisition of socially valuable information, which helps to screen projects and control the social aggregate risk. For mature corporations with robust growth, however, whose priority is liquidity, debt could be better, as it deters unnecessary acquisition of information that is not socially valuable. This consideration jibes in part with the pecking-order theory, and future work may further unify the life-cycle evolution of capital structure of corporations with a theory of flexible information acquisition.

In an analogous way, positing flexible information acquisition can be helpful in revisiting the literature on the endogenous determination of capital structure, by specializing information acquisition. Given flexibility, agents who monitor may have different incentives to acquire qualitatively different information when facing different forms of financial contracts. Hence in certain capital structures, different layers of financial contracts permit a specialization in information acquisition. In other words, layers of capital structure correspond to specialized layers of information to be acquired. This specialization may in turn affect the production of information and the efficiency of monitoring, further reshaping the optimal capital structure. In this way, we can see that flexibility of information acquisition plays a role in determining the capital structure. Further results regarding its effects on corporate finance as well as on social welfare will come from future work.

\[23\) This idea is explored in a working paper, (Yang and Zeng 2012).
REFERENCES


**Mathematical Appendix**

**Derivation of mutual information.**

Before observing her signal, the agent’s uncertainty about $\theta$ is given by Shannon’s entropy of her prior\(^{24}\)

$$H\left(\text{prior}\right) = -\mathbb{E} \ln p\left(\theta\right),$$

where the expectation operator $\mathbb{E}(\cdot)$ is with respect to $\theta$ under prior $P$, and $p$ is the density function of $P$\(^{25}\). After observing signal 1, the agent forms a posterior of $\theta$

$$\frac{m\left(\theta\right) p\left(\theta\right)}{\mathbb{E}m\left(\theta\right)}$$

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

$$H\left(\text{posterior}|1\right) = -\mathbb{E} \left[ \frac{m\left(\theta\right)}{\mathbb{E}m\left(\theta\right)} \ln \left( \frac{m\left(\theta\right) p\left(\theta\right)}{\mathbb{E}m\left(\theta\right)} \right) \right].$$

Similarly, observing signal 0 leads to a posterior

$$\frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \mathbb{E}m\left(\theta\right)}$$

and posterior entropy

$$H\left(\text{posterior}|0\right) = -\mathbb{E} \left[ \frac{1 - m\left(\theta\right)}{1 - \mathbb{E}m\left(\theta\right)} \ln \left( \frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \mathbb{E}m\left(\theta\right)} \right) \right].$$

Then the agent’s expected posterior entropy through choosing information structure $m\left(\cdot\right)$ is

$$H\left(\text{posterior}\right) = -\mathbb{E} \left[ m\left(\theta\right) \ln \left( \frac{m\left(\theta\right) p\left(\theta\right)}{\mathbb{E}m\left(\theta\right)} \right) \right] - \mathbb{E} \left[ \left[1 - m\left(\theta\right)\right] \ln \left( \frac{\left[1 - m\left(\theta\right)\right] p\left(\theta\right)}{1 - \mathbb{E}m\left(\theta\right)} \right) \right].$$

Hence, mutual information $I\left(m\right)$ is

$$I\left(m\right) = H(\text{prior}) - H(\text{posterior}) = \mathbb{E} g\left(m\left(\theta\right)\right) - g\left(\mathbb{E}m\left(\theta\right)\right),$$

---

\(^{24}\)This is essentially the unique measure of uncertainty given three axioms. See (Cover and Thomas 1991) for detailed discussion.

\(^{25}\)Following the convention of information theory, we let $0 \cdot \ln 0 = 0$. This is reasonable since $\lim_{x \to 0} x \cdot \ln x = 0$. 

---

where
\[ g(x) = x \ln x + (1 - x) \ln (1 - x) . \]

**Proof of Proposition 1.**

PROOF: Suppose \( m \) is an optimal strategy. Let \( \varepsilon \) be any feasible perturbation function. The payoff from the perturbed strategy \( m + \alpha \cdot \varepsilon \) is
\[
V^* (m + \alpha \cdot \varepsilon) = \mathbb{E} [ (m(\theta) + \alpha \cdot \varepsilon(\theta)) \cdot \Delta u(\theta)]
- \mu \cdot \left[ \mathbb{E} g(m(\theta) + \alpha \cdot \varepsilon(\theta)) - g(\mathbb{E}[m(\theta) + \alpha \cdot \varepsilon(\theta)]) \right],
\]
where the expectation operator \( \mathbb{E}(\cdot) \) is with respect to \( \theta \) under common prior \( P \). \( \alpha \in \mathbb{R} \), and \( \varepsilon \) is feasible with respect to \( m \) if \( \exists \alpha > 0 \), s.t. \( \forall \theta \in \Theta, m(\theta) + \alpha \cdot \varepsilon(\theta) \in [0, 1] \).

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

Note that
\[
\Delta u(\theta) - \mu \cdot (g'(m(\theta)) - g'(p_1))
\]
is the Fréchet derivative of \( V^* (\cdot) \) at \( m \). Hence the tangent hyperplane at \( m \) can be expressed as
\[
\left\{ (\tilde{m}, v) \in M \times \mathbb{R} : v - V^* (m) = \mathbb{E} \left[ \left[ \Delta u(\theta) - \mu g'(m(\theta)) + \mu g' \left( \int_{\Theta} m(\theta) \cdot dP(\theta) \right) \right] (\tilde{m}(\theta) - m(\theta)) \right] \right\} .
\]

An important observation: since \( V^* (\cdot) \) is a concave functional on \( M \), \( V^* \) is upper bounded by any hyperplane tangent at any \( m \in M \), i.e., \( \forall m, \tilde{m} \in M \),
\[
V^* (\tilde{m}) - V^* (m) \leq \mathbb{E} \left[ \left[ \Delta u(\theta) - \mu g'(m(\theta)) + \mu g' \left( \int_{\Theta} m(\theta) \cdot dP(\theta) \right) \right] (\tilde{m}(\theta) - m(\theta)) \right].
\]

This inequality is strict when
\[
m \in M^o \triangleq M \setminus \{ m \in M : m(\theta) \text{ is a constant a.s.} \}
\]
and \( \Pr (\tilde{m}(\theta) \neq m(\theta)) > 0 \), since \( V^* (\cdot) \) is strictly concave on \( M^o \). This observation is helpful later in our proof.

The optimality of \( m \) requires \( \left. \frac{dV^* (m + \alpha \cdot \varepsilon)}{d\alpha} \right|_{\alpha = 0} \leq 0 \) for all feasible perturbation \( \varepsilon \). Hence we must have
\[
\Delta u(\theta) - \mu \cdot (g'(m(\theta)) - g'(p_1)) \begin{cases} 
\geq 0 & \text{if } m(\theta) = 1 \\
= 0 & \text{if } m(\theta) \in (0, 1) \\
\leq 0 & \text{if } m(\theta) = 0
\end{cases}
\]

(A1)
Note that $\Pr (m (\theta) = 1) > 0$ implies $\Pr (m (\theta) = 1) = 1$. Otherwise,

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

implies

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

for all $\theta$ in set

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

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

$$\left. \frac{dV^* (m + a \cdot \varepsilon)}{da} \right|_{a=0} = \int_B (-\infty) \cdot (-1) \, dP (\theta) = +\infty ,$$

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

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

(A2)  \hspace{1cm} \Delta u (\theta) - \mu \cdot (g' (m (\theta)) - g' (p_1)) = 0 \text{ a.s.} .

By definition,

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

thus (A2) implies

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

Let

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

and

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

then there exists $p_1 \in [0, 1]$ such that $m (\cdot, p_1) \in M_1$ is an optimal strategy. Note that $J (p_1) = p_1$ is a necessary condition for the optimality of $m (\cdot, p_1)$. 
Since \( m ( \cdot, p_1 ) \in M_1 \subset M \), the original problem is reduced to
\[
\max_{p \in [0, 1]} V^* (m (\cdot, p)) = \mathbb{E} \left[ \Delta u (\theta) \cdot m (\theta, p) \right] - c (m (\cdot, p)) .
\]
The first order derivative with respect to \( p \) is
\[
\frac{dV^* (m (\cdot, p))}{dp} = \mathbb{E} \left( \left[ \Delta u (\theta) - \mu \cdot g' (m (\theta, p)) + \mu \cdot g' (J (p)) \right] \cdot \frac{\hat{m} (\theta, p)}{\hat{p}} \right).
\]
By definition,
\[
\Delta u (\theta) - \mu \cdot g' (m (\theta, p)) = -\mu \cdot g' (p) ,
\]
thus
\[
\frac{dV^* (m (\cdot, p))}{dp} = \mu \cdot \left[ g' (J (p)) - g' (p) \right] \cdot \mathbb{E} \frac{\hat{m} (\theta, p)}{\hat{p}} .
\]
Since
\[
\frac{\hat{m} (\theta, p)}{\hat{p}} > 0
\]
for all \( \theta \in \Theta \),
\[
\frac{dV^* (m (\cdot, p))}{dp} \geq 0
\]
if and only if
\[
g' (J (p)) - g' (p) \geq 0 .
\]
Since \( g' \) is strictly increasing, we have
\[
\frac{dV^* (m (\cdot, p))}{dp} \geq 0
\]
if and only if
\[
J (p) \geq p .
\]
In order to be a global maximum, \( m (\cdot, p_1) \) must first be a local maximum within \( M_1 \).
This requires
\[
(A4) \quad J (p_1) = p_1 .
\]
But \( A4 \) is not sufficient. The sufficient condition for \( m (\cdot, p_1) \) to be a local maximum within \( M_1 \) is
\[
\exists \text{ neighborhood } (p_1 - \beta, p_1 + \beta) ,
\]
s.t. \( J (p) \geq p \) for all \( p \in (p_1 - \beta, p_1] \)
and \( J (p) \leq p \) for all \( p \in [p_1, p_1 + \beta) .
\]
Note that
\[ J(0) = 0, \quad J(1) = 1, \]
\[ \frac{dJ}{dp}\bigg|_{p=0} = \mathbb{E} \exp \left( \mu^{-1} \Delta u(\theta) \right) \]
and
\[ \frac{dJ}{dp}\bigg|_{p=1} = \mathbb{E} \exp \left( -\mu^{-1} \Delta u(\theta) \right). \]

We proceed by discussing four possible cases.

**Case i):**
\[ \mathbb{E} \exp \left( \mu^{-1} \Delta u(\theta) \right) > 1 \]
and
\[ \mathbb{E} \exp \left( -\mu^{-1} \Delta u(\theta) \right) > 1. \]

In this case, \( J(p) > p \) for \( p \) close enough to 0 and \( J(p) < p \) for \( p \) close enough to 1. Since \( J(p) \) is continuous, the set \( \{ p \in (0, 1) : J(p) = p \} \) is non-empty. For any \( p_1 \in \{ p \in (0, 1) : J(p) = p \} \), the Fréchet derivative at \( m(\cdot, p_1) \) is
\[ \Delta u(\theta) - \mu \cdot g'(m(\theta, p_1)) + \mu \cdot g'(J(p_1)) = 0 \]
and thus \( m(\cdot, p_1) \) is a critical point of functional \( V^*(\cdot) \). Since \( m(\cdot, p_1) \in M^0 \), the observation mentioned above implies
\[ V^*(m) - V^*(m(\cdot, p_1)) < 0 \]
for all \( m \in M \) such that \( \Pr (m(\theta) \neq m(\theta, p_1)) > 0 \). Hence, \( V^*(m(\cdot, p_1)) \) is strictly higher than the values achieved at any other \( m \in M \), i.e.,
\[ \{ p \in (0, 1) : J(p) = p \} = \{ p_1 \} \]
and \( m(\cdot, p_1) \) is the *unique* global maximum. This actually proves (6).

**Case ii):**
(A5) \[ \mathbb{E} \exp \left( \mu^{-1} \Delta u(\theta) \right) > 1 \]
and
(A6) \[ \mathbb{E} \exp \left( -\mu^{-1} \Delta u(\theta) \right) \leq 1. \]

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

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

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

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

which implies

\[ \mathbb{E} \exp \left( -2\mu^{-1} \Delta u (\theta) \right) > \left[ \mathbb{E} \exp \left( -\mu^{-1} \Delta u (\theta) \right) \right]^2 = 1 \]
due to Jensen’s inequality and the assumption that \( \Delta u (\theta) \) is not constant. Thus we have

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

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

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

\[ p_1 = \sup \{ p \in (0, 1) : J (p) \leq p \}. \]

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

\[ V^\ast \left( m (\cdot, 1) \right) > V^\ast \left( m (\cdot, 0) \right). \]

Hence, \( p_1 = 1 \), i.e., \( m (\theta) = 1 \) a.s. is the unique optimal strategy. This actually proves (4).

**case iii):**

(A9)

\[ \mathbb{E} \exp \left( \mu^{-1} \Delta u (\theta) \right) \leq 1 \]

and

(A10)

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

Note that this case can be summarized by (A9) alone, since (A10) is a direct implication of (A9) according to Jensen’s inequality. In this case, by the same argument as in case ii), \( m (\theta) = 0 \) a.s. is the unique optimal strategy. This actually proves (5).
case iv):

(A11) \[ \mathbb{E} \exp (\mu^{-1} \Delta u (\theta)) \leq 1 \]

and

(A12) \[ \mathbb{E} \exp (-\mu^{-1} \Delta u (\theta)) \leq 1 . \]

Case iv) cannot exists since Jensen’s inequality implies

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

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

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

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

Proof of Proposition 2.

PROOF: We prove by constructing a debt that generates positive expected payoff to the seller. Let \( \beta \in (\delta, \delta_b^{-1}, 1) \) and

\[ f (q) = \mathbb{E} \min \left( \sum_{n=1}^{N} \theta_n, \beta \delta_s q \right) . \]

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

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

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

\[ f'' (q) = \text{Pr} \left( \sum_{n=1}^{N} \theta_n \geq \beta \delta_s^{-1} q \right) \cdot \beta \delta_s^{-1} \]

\[ > \beta^{-1} \delta_s \delta_b^{-1} \cdot \beta \delta_s^{-1} = \delta_b^{-1} . \]

Note that

\[ f (0) = 0 , \]

which implies that

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

for all \( q \in (0, q_0) \).
Consider a debt
\[ s\left(\overrightarrow{\theta}\right) = \min\left(\sum_{n=1}^{N} \theta_n, D\right) \]
with face value \( D = \beta \delta^{-1} q \) and price \( q \in (0, q_0) \). The buyer’s payoff gain from accepting this offer over rejecting it is
\[ \Delta u \left(\overrightarrow{\theta}\right) = \delta_b \cdot s \left(\overrightarrow{\theta}\right) - q. \]

By Jensen’s inequality, we have
\[
\begin{align*}
\mathbb{E} \exp \left( \mu^{-1} \Delta u \left(\overrightarrow{\theta}\right) \right) &\geq \exp \left( \mu^{-1} \mathbb{E} \Delta u \left(\overrightarrow{\theta}\right) \right) \\
&= \exp \left( \mu^{-1} \left[ \delta_b \cdot f \left( q \right) - q \right] \right) \\
&> 1,
\end{align*}
\]
which implies \( \overline{p}_{s,q} > 0 \) according to Proposition 1. Then, the seller’s expected payoff from this contract is
\[
W \left( s, q \right) = \mathbb{E} \left( m_{s,q} \left( \overrightarrow{\theta}\right) \cdot \left[ q - \delta_s \cdot s \left(\overrightarrow{\theta}\right) \right] \right)
\]
\[
\geq \mathbb{E} \left( m_{s,q} \left( \overrightarrow{\theta}\right) \cdot \left[ q - \delta_s \cdot \beta \delta^{-1} q \right] \right)
\]
\[
= \left( 1 - \beta \right) q \cdot \overline{p}_{s,q} > 0.
\]

By proposing the optimal contract \( (s^*, q^*) \), the seller’s expected payoff should be no less than \( W \left( s, q \right) > 0 \). This directly implies \( \overline{p}_{s^*,q^*} > 0 \), since \( \overline{p}_{s^*,q^*} = 0 \) always generates zero expected payoff to the seller.

**Proof of Proposition 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 \( s^* \). Let
\[
J \left( \alpha \right) = \mu \ln \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s \left(\overrightarrow{\theta}\right) \right) + \delta_s \cdot \mathbb{E} s \left(\overrightarrow{\theta}\right).
\]

Taking first order variation leads to
\[
\begin{align*}
\left. \frac{dJ}{d\alpha} \right|_{\alpha=0} &= \mathbb{E} \left[ \left( \delta_s - \delta_b \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s^* \left(\overrightarrow{\theta}\right) \right) \right) ^{-1} \exp \left( -\mu^{-1} \delta_b \cdot s^* \left(\overrightarrow{\theta}\right) \right) \cdot \varepsilon \left(\overrightarrow{\theta}\right) \right] \\
&= (A13) \mathbb{E} \left[ r \left(\overrightarrow{\theta}\right) \cdot \varepsilon \left(\overrightarrow{\theta}\right) \right].
\end{align*}
\]
Let

\[ A_0 = \left\{ \vec{\theta} \in \Theta : \vec{\theta} \neq \vec{0}, s^* (\vec{\theta}) = 0 \right\}, \]

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

and

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

Then \( \{A_0, A_1, A_2\} \) is a partition of \( \Theta \setminus \{\vec{0}\} \). Since \( s^* \) is the optimal security,

\[ \frac{dJ}{da} \bigg|_{a=0} \geq 0 \]

holds for any feasible perturbation \( \varepsilon (\vec{\theta}) \). Hence, we have

\[
\begin{align*}
\hat{r} (\vec{\theta}) & \begin{cases} 
\geq 0 & \text{if } \vec{\theta} \in A_0 \\
= 0 & \text{if } \vec{\theta} \in A_1 \\
\leq 0 & \text{if } \vec{\theta} \in A_2
\end{cases}.
\end{align*}
\]

For any \( \vec{\theta}' \in A_0 \), (A14) implies \( \hat{r} (\vec{\theta}') \geq 0 \), i.e.,

\[ \delta_s \geq \delta_b \left[ \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s^* (\vec{\theta}) \right) \right]^{-1}. \]

Together with (12), this inequality implies

\[ \ln \delta_s \geq \ln \delta_b + \mu^{-1} q^*. \]

Hence,

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

which is a contradiction. Therefore,

\[
(A15) \quad \Pr (A_0) = 0.
\]

For any \( \vec{\theta}' \in A_1 \), (A14) implies \( \hat{r} (\vec{\theta}') = 0 \), i.e.,

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

i.e.,

\[ \ln \delta_s = \ln \delta_b + \mu^{-1} q^* - \mu^{-1} \delta_b \cdot s^* (\vec{\theta}'). \]
Therefore,

\[(A16)\]

\[s^* \left( \tilde{\theta}^* \right) = \mu \delta_b^{-1} \cdot [\ln \delta_b - \ln \delta_s] + \delta_b^{-1} q^*\]

is a constant for all \(\tilde{\theta}^* \in A_1\).

For any \(\tilde{\theta}^* \in A_2\), \((A14)\) implies \(r \left( \tilde{\theta}^* \right) \leq 0\), i.e.,

\[\delta_s \leq \delta_b \left[ \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot s^* \left( \tilde{\theta}^* \right) \right) \right]^{-1} \exp \left( -\mu^{-1} \delta_b \cdot \sum_{n=1}^{N} \theta'_n \right),\]

i.e.,

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

Therefore,

\[(A17)\]

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

Let

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

Then, \((A15)\), \((A16)\) and \((A17)\) imply that

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

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

Finally, let

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

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

By (12), we have

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

Hence \(q^*\) is a fixed point of \(h \left( q \right)\).
It is obvious that \( h(0) > 0 \). Also note that
\[
\begin{align*}
\frac{\partial h}{\partial q}(q) & = \left[ \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot \min \left( \sum_{n=1}^{N} \theta_n, \mu \delta_b^{-1} \cdot (\ln \delta_b - \ln \delta_s) + \delta_b^{-1} q \right) \right) \right]^{-1} \\
& \cdot \mathbb{E} \left[ \exp \left( -\mu^{-1} \delta_b \cdot \min \left( \sum_{n=1}^{N} \theta_n, \mu \delta_b^{-1} \cdot (\ln \delta_b - \ln \delta_s) + \delta_b^{-1} q \right) \right) \right] \cdot 1^{\left\{ \sum_{n=1}^{N} \theta_n \geq \mu \delta_b^{-1} \cdot (\ln \delta_b - \ln \delta_s) + \delta_b^{-1} q \right\}} \\
& \leq 1
\end{align*}
\]
and
\[
\lim_{q \to \infty} \frac{\partial h}{\partial q}(q) = \left[ \mathbb{E} \exp \left( -\mu^{-1} \delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \right]^{-1} \cdot \mathbb{E} \left[ \exp \left( -\mu^{-1} \delta_b \cdot \sum_{n=1}^{N} \theta_n \right) \right] \cdot \lim_{q \to \infty} 1^{\left\{ \sum_{n=1}^{N} \theta_n \geq \mu \delta_b^{-1} \cdot (\ln \delta_b - \ln \delta_s) + \delta_b^{-1} q \right\}} = 0.
\]

Hence, \( h(q) \) has a unique fixed point \( q^* > 0 \).

**Proof of Lemma 1.**

**Proof:** Taking derivative with respect to \( \alpha \) at \( \alpha = 0 \) for both sides of (16) leads to
\[
\begin{align*}
\mu^{-1} \delta_b \cdot e(\overline{\theta}) & = g'' \left( m_{s^*,q^*} \left( \overline{\theta} \right) \right) \cdot \frac{d m_{s,q^*} \left( \overline{\theta} \right)}{d \alpha} \bigg|_{\alpha = 0} \\
& + g'' \left( \overline{p}_{s^*,q^*} \left( \overline{\theta} \right) \right) \cdot \mathbb{E} \left( \frac{d m_{s,q^*} \left( \overline{\theta} \right)}{d \alpha} \right) \bigg|_{\alpha = 0}.
\end{align*}
\]

Take expectation of both sides and manipulate we get
\[
\begin{align*}
\mathbb{E} \left( \frac{d m_{s,q^*} \left( \overline{\theta} \right)}{d \alpha} \right) \bigg|_{\alpha = 0} & = \mu^{-1} \delta_b \left[ 1 - \mathbb{E} \left[ g'' \left( m_{s^*,q^*} \left( \overline{\theta} \right) \right) \right] \cdot g'' \left( \overline{p}_{s^*,q^*} \left( \overline{\theta} \right) \right) \right]^{-1} \cdot \mathbb{E} \left( g'' \left( m_{s^*,q^*} \left( \overline{\theta} \right) \right) \right)^{-1} e \left( \overline{\theta} \right).
\end{align*}
\]

Combining the above two equations leads to (18).

**Proof of Proposition 4.**

**Proof:** We first prove \( f_1(0) > f_2(0) \). If not, \( f_1(s) < f_2(s) \) for all \( s > 0 \). Hence
\[ \forall \vec{\theta} \neq \vec{\theta}^*, \]
\[ r \left( \vec{\theta} \right) \cdot g'' \left( m_{s^*, q^*} \left( \vec{\theta} \right) \right) \]
\[ = -\delta_s \cdot f_2 \left( s^* \left( \vec{\theta} \right) \right) g'' \left( f_2 \left( s^* \left( \vec{\theta} \right) \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot s^* \left( \vec{\theta} \right) + w \right) \]
\[ < -\delta_s \cdot f_1 \left( s^* \left( \vec{\theta} \right) \right) g'' \left( f_1 \left( s^* \left( \vec{\theta} \right) \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot s^* \left( \vec{\theta} \right) + w \right) \]
\[ = 0, \]

where the inequality holds since \[ m \cdot g'' \left( m \right) \] > 0. Then (23) implies \( s^* \left( \vec{\theta} \right) \) = 0 almost surely. Therefore, there is no trade, which contradicts Proposition 2.

For any \( \vec{\theta} \in A_0 \),
\[ r \left( \vec{\theta} \right) \cdot g'' \left( m_{s^*, q^*} \left( \vec{\theta} \right) \right) \]
\[ = -\delta_s \cdot f_2 \left( 0 \right) g'' \left( f_2 \left( 0 \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot 0 + w \right) \]
\[ > -\delta_s \cdot f_1 \left( 0 \right) g'' \left( f_1 \left( 0 \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot 0 + w \right) \]
\[ = 0, \]

where the inequality holds since \( f_1 \left( 0 \right) > f_2 \left( 0 \right) \) and \( m \cdot g'' \left( m \right) \] > 0. This result contradicts (23), which states \( r \left( \vec{\theta} \right) \cdot g'' \left( m_{s^*, q^*} \left( \vec{\theta} \right) \right) \leq 0 \) for \( \vec{\theta} \in A_0 \).

**Proof of Proposition 5.**

PROOF: Let \( (s, m) \) be the unique intersection of \( f_1 \left( s \right) \) and \( f_2 \left( s \right) \). For any \( \vec{\theta} \) such that \( \sum_{n=1}^{N} \theta_n < \bar{s} \), we have
\[ m_{s^*, q^*} \left( \vec{\theta} \right) = f_2 \left( s^* \left( \vec{\theta} \right) \right) < f_2 \left( \bar{s} \right) = f_1 \left( \bar{s} \right) < f_1 \left( s^* \left( \vec{\theta} \right) \right). \]

Then
\[ r \left( \vec{\theta} \right) \cdot g'' \left( m_{s^*, q^*} \left( \vec{\theta} \right) \right) \]
\[ > -\delta_s \cdot f_1 \left( s^* \left( \vec{\theta} \right) \right) \cdot g'' \left( f_1 \left( s^* \left( \vec{\theta} \right) \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot s^* \left( \vec{\theta} \right) + w \right) \]
\[ = 0, \]

where the inequality holds since \( m \cdot g'' \left( m \right) \] > 0. Condition (23) then implies
\[ s^* \left( \vec{\theta} \right) = \sum_{n=1}^{N} \theta_n \]

for all \( \vec{\theta} \) such that \( \sum_{n=1}^{N} \theta_n < \bar{s} \).
For any $\vec{\theta}$ such that $\sum_{n=1}^{N} \theta_n > \bar{s}$, if $s^* (\vec{\theta}) = \sum_{n=1}^{N} \theta_n$, then (23) implies

$$0 \leq r (\vec{\theta}) \cdot g'' \left( m_{s^*, q^*} (\vec{\theta}) \right) = -\delta_s \cdot f_2 \left( s^* (\vec{\theta}) \right) \cdot g'' \left( f_2 \left( s^* (\vec{\theta}) \right) \right) + \mu^{-1} \delta_b \left( q^* - \delta_s \cdot s^* (\vec{\theta}) + w \right)$$

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

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

$$= 0,$$

which is a contradiction. Hence Proposition 4 implies $s^* (\vec{\theta}) = \bar{s}$ for all $\vec{\theta}$ such that $\sum_{n=1}^{N} \theta_n > \bar{s}$.

For any $\vec{\theta}$ such that $\sum_{n=1}^{N} \theta_n = \bar{s}$, $s^* (\vec{\theta}) = \bar{s}$ is a direct implication of Proposition 4.

Therefore, the optimal security is a debt with face value $\bar{s}$, i.e.,

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

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

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

is a special debt (with a very large face value).

**Proof of Proposition 6.**

PROOF: Let $\beta \in (\delta_s \delta_b^{-1}, 1)$ and

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

where the expectation operator $\mathbb{E} (\cdot)$ is with respect to $\theta$ under 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 $q \in (0, q_0)$,

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

$$> \beta \delta_b \delta_s^{-1} \cdot \beta^{-1} \delta_s \delta_b^{-1} = 1.$$

Note that

$$f(0) = 0,$$

which implies

$$f(q) > q$$

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

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

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

$$q - \delta_s \cdot s(\theta)$$

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

$$> 0$$

for all $\theta \in \Theta$, the seller will accept this offer without acquiring information. Hence, the buyer’s expected payoff from proposing $(s, q)$ is

$$W(s, q) = f(q) - q > 0.$$

By proposing the optimal contract $(s^*, q^*)$, the buyer’s expected payoff should be no less than $W(s, q) > 0$. This directly implies $\overline{p}_{s^*, q^*} > 0$, since $\overline{p}_{s^*, q^*} = 0$ always generates zero expected payoff to the buyer.