Asset Bundling and Information Acquisition of Investors with Different Expertise

Liang Dai*

December 19, 2016

Abstract

This paper investigates how a profit-maximizing asset originator can coordinate the information acquisition of investors with different expertise by means of asset bundling. Bundling is beneficial to the originator when it discourages investors from analyzing idiosyncratic risks and focuses their attention on aggregate risks. But it is optimal to sell aggregate risks separately in order to exploit investors’ heterogeneous expertise in learning about them and thus lower the risk premium. This analysis rationalizes the common securitization practice of bundling loans by asset class, which is at odds with existing theories based on diversification. The analysis also offers an alternative perspective on conglomerate formation (a form of asset bundling), and its relation to empirical evidence in that context is discussed.

*Antai College of Economics and Management, Shanghai Jiaotong University. liangdai@sjtu.edu.cn. I am extremely grateful to Stephen Morris, Valentin Haddad, Hyun Song Shin and Wei Xiong for their continuous guidance and support. I also thank seminar participants at Princeton University, Shanghai Advanced Institute of Finance, The University of Hong Kong, Chinese University of Hong Kong, Peking University (Guanghua SEM), Fudan University, Shanghai University of Finance and Economics, China Meeting of Econometric Society 2016 and Asia Meeting of Econometric Society 2016 for helpful comments and discussions. All remaining errors are mine.
1 Introduction

Securitization plays an important role in the U.S. economy. As of April 2011, outstanding securitized assets totaled $11 trillion, which was substantially more than the amount of all outstanding marketable U.S. Treasury securities (Gorton and Metrick, 2013). One salient feature of securitization is that the creation of asset-backed securities (ABS) always involves pooling loans of the same asset class; i.e., a pool consists exclusively of mortgages, auto receivables or credit card receivables. Different asset classes are not mixed, even if the originator in fact instigates loans of many different asset classes. Existing theories based on diversification\(^1\) do not square well with this feature, as one would expect the benefit of diversification to be greater when different asset classes are mixed.

In this paper I demonstrate that this feature is no longer a puzzle if we recognize the important role played by the heterogeneous expertise of investors in acquiring different information about asset payoffs, which existing theories of securitization abstract from. Pooling all loans of the same asset class prohibits buyers from cherry-picking individual loans, and thus prevents them from using their expertise to exploit other buyers regarding the risks peculiar to the loans picked. This encourages all buyers to acquire information only about risks common to all the loans being sold. Since they face less uncertainty after learning about these risks, buyers demand a lower risk premium from the originator. Different asset classes are sold separately. This enables mortgage specialists to freely trade mortgages and to profit from mortgage-specific information and thus induces them to specialize in acquiring information in their area of expertise. The comparative advantages in information acquisition of different buyers are thus better utilized and result in a lower total risk premium required, benefiting the originator.\(^2\)

---

\(^1\)For example, Subrahmanyam (1991) shows that the introduction of a basket of securities reduces the problem of adverse selection by offsetting demand from informed traders that have private information about individual securities. Demarzo (2005) shows that when the seller has better information, pooling makes the value of the ABS created less sensitive to the his private information about individual assets. When instead the buyer has better information, pooling prevents her from cherry-picking only good assets. Adverse selection is reduced in both cases, as private information about individual assets is diversified.

\(^2\)Parlour and Plantin (2008) point out that "Interestingly, however, the secondary loan market does not
This paper develops a model that formalizes this explanation and further studies a broader theoretical issue: How can a self-interested asset originator coordinate the information acquisition of investors that have different areas of expertise? Because potential investors in any financial asset inherently have different learning expertise, this seems to be a fundamental question in understanding the workings of the financial market, in addition to rationalizing the puzzle as an application, but it has received little attention in the literature to date. As a first step, this paper focuses on asset bundling, a technique commonly used by asset originators. The application of asset bundling in financial market practice is not limited to securitization. Indeed, a conglomerate can also be viewed as a bundle of its several lines of business, in the sense that its stakeholders cannot selectively invest in and receive cash flows from any particular business that it operates. Thus, the model developed can also be used to study conglomerate formation.

My model features two key ingredients: the interaction of heterogeneous investors and their endogenous learning behavior. Asset payoffs are determined by different risks; e.g., sector-specific shocks, region-specific shocks, asset-specific shocks. There is one asset originator and a continuum of investors with different learning expertise. Each risk-averse investor allocates his limited attention to learning about these risks before trading the assets. How he does that is endogenously shaped by the bundling choice of the asset originator and by his interaction with other investors. The asset originator, who wants to maximize the revenue of the sale, bundles his original assets to channel the allocation of investors’ learning capacity in the way that minimizes the total risk premium.

Three key theoretical channels novel in the literature are highlighted in the model, leading to the upside and downside of asset bundling.

The upside of asset bundling is driven by a discipline channel: asset bundling restricts
speculation on risks that are supposedly diversified away, and gives investors less incentive
to acquire information about them. As such, the originator successfully persuades investors
to learn only about risks that cannot be reduced by diversification. Since investors have
better knowledge of such risks after studying them, they demand a lower risk premium in
equilibrium, benefiting the originator.

The downside of asset bundling is driven by two different economic forces. First, asset
bundling mechanically restricts the asset span available to investors, thus preventing them
from holding their respective favorite portfolios. Hence in equilibrium, they demand lower
prices to compensate. This is a trade-restriction channel. Second, asset bundling induces
each investor to specialize less in acquiring information about the risk that he has expertise in.
Because the expertise of investors is less utilized, there are more risks priced in equilibrium.
This is a specialization-destruction channel.

These theoretical channels work not only in the context of securitization, but also in the
context of conglomerate formation. By relabeling the asset originator as an entrepreneur
who owns several lines of business and decides how to set the firm boundaries, my model
can also be viewed as one of conglomerate formation. It offers a new investor-side (instead
of firm-side) perspective of conglomerate formation that can generate both a diversification
premium (by the discipline channel) and a discount (by the trade-restriction channel and the
specialization-destruction channel), and yields empirical predictions consistent with existing
evidence in the literature. As such, my model also builds a conceptual connection between
securitization and conglomerate formation, two seemingly remote contexts that are both
important in their own right.

My model follows Van Nieuwerburgh and Veldkamp (2009, 2010), which study the en-
dogenous information acquisition of investors with heterogeneous expertise, and uses their
modeling approach. My work differs from theirs, as my focus is on the implications of asset
design and asset pricing rather than on the portfolio choices of individual investors.

There are a few papers that also study the endogenous information acquisition of in-
vestors. Peng and Xiong (2006) show how the limited attention of a representative investor leads to categorical learning and return comovement. In a multiple asset, noisy rational expectations model with rational inattentive investors, Mondria (2010) shows how investors’ attention allocation generates asset price comovement. For technical simplification, these papers do not incorporate the interaction of heterogeneous investors. Subrahmanyam (1991) demonstrates how markets of baskets of securities reduce adverse selection cost. Recently, Goldstein and Yang (2015) identify strategic complementarities in the trading and information acquisition of investors informed about different components of the same asset. These two papers endow traders with exogenous information in their baseline models, and traders are ex ante identical in the extensions with endogenous information acquisition.

My work is also related to the literature on security design. In addition to rationalizing the feature of bundling loans by asset classes of securitization, my model complements this literature in two aspects. First, it studies the interaction of heterogeneous security buyers, which existing security design models (e.g. Demarzo and Duffie 1999; Demarzo 2005) typically abstract from. Second, existing security-design models (e.g. Townsend 1979; Dang et al. 2013) usually focus on the extensive margin of information acquisition; i.e., how to reduce the costly information acquisition of security buyers. My model focuses instead on the intensive margin: given the resources available to security buyers for information acquisition, how can the seller induce buyers to use those resources in his preferred way? A more detailed discussion on the relation of my work to this literature is given in Section 5.2.

My work is also related to the literature on financial innovation (e.g. Marin and Rahi 2000; Duffie and Rahi 1995). I obtain a similar result that more complete, but less than perfectly complete financial markets may not be Pareto optimal, as shown in Section 5.3. In this literature, each investor’s private knowledge (i.e., knowledge NOT obtained from prices) of assets being traded is typically exogenous. My model complements their work by exploring how asset design can endogenously affect each investor’s incentive to acquire private knowledge of asset fundamentals.
Lastly, my work complements the literature on corporate diversification by offering an alternative perspective on conglomerate formation. A detailed discussion can be found in Section 6.

The rest of this paper is organized as follows. Section 2 introduces the setup of the baseline model. Section 3 illustrates the discipline channel by studying a polar case in which only one risk is non-diversifiable. Section 4 illustrates the trade-restriction channel and the specialization-destruction channel by studying another polar case in which all sources of risks are non-diversifiable and play a symmetric role. Section 5 introduces a generalization of the baseline model that establishes the optimality of categorization strategy and discusses several issues of the baseline model. Section 6 discusses the application of the model in the context of corporate diversification and relevant empirical evidence in the existing literature. Section 7 concludes.

2 Baseline Model

2.1 Risks and Asset Payoffs

There are two orthogonal sources of risks (hereafter "risks"): $f_1$, $f_2$, and two risky assets, with a supply of one each, and payoffs

$$
\begin{pmatrix}
X_1 \\
X_2
\end{pmatrix}
= 
\begin{pmatrix}
\gamma_{11} & \gamma_{12} \\
\gamma_{21} & \gamma_{22}
\end{pmatrix}
\begin{pmatrix}
f_1 \\
f_2
\end{pmatrix},
$$
or in matrix form, $X = \Gamma f$, such that $\Gamma = (\gamma_{ij})$ is an orthogonal matrix.\(^3\) $w_i \equiv \gamma_{1i} + \gamma_{2i}, i = 1, 2$ is the loading of total asset payoff ($X_1 + X_2$) on $f_i$. The orthogonality of $\Gamma$ implies that $w_1^2 + w_2^2 = 2$. Without loss of generality, hereafter we consider only the range in which $w_1 \geq 1$ and $0 \leq w_2/w_1 \leq 1$.

There is also a risk-free asset with an unlimited supply, and its gross return is normalized to 1.

\(^3\)One can always make $\Gamma$ orthogonal by redefining risks $f$ through the eigenvalue decomposition of $Var(X)$. 

6
2.2 Originator

There is a risk-neutral originator, who owns all the risky assets and wants to sell them. His objective is to maximize the expected total revenue. To do so, he chooses how to bundle the assets (i.e., creating new tradable non-redundant asset(s) that are linear combinations of the original assets, such that the former completely absorb the latter), and then sells all of them. This means he can create a single new asset with payoff \( Y = X_1 + X_2 \) and supply of 1, or instead he can create two new assets, each with supply of 1, and payoffs \( Y_k = t_{k,1}X_1 + t_{k,2}X_2, \ k = 1, 2 \), such that \( t_{1,i} + t_{2,i} = 1, \ i = 1, 2 \); i.e., the original assets are exhausted, and \( t_{1,1}/t_{2,1} \neq t_{1,2}/t_{2,2} \).

Each bundling strategy can be uniquely represented by a matrix \( T \): \( T = (1, 1) \) if a single asset is created, and \( T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \) if two tradable assets are created. By construction, \( T \) has full rank, \( 1'T = 1' \), and payoff(s) of the tradable asset(s) \( Y = TX \). The originator’s problem can be expressed as \( \max T E_0[1'p_T] \), where \( p_T \) denotes the price(s) of asset(s) formed by strategy \( T \).

2.3 Investors

There are two types \( i \in \{1, 2\} \) of risk-averse investors, each with a continuum of mass 1/2. Each investor starts with a flat prior with mean zero about the risks \( f \), and does two things sequentially after observing the bundling choice of the originator: 1) acquires information about the risks \( f \) to maximize his expected utility at the trading stage; 2) chooses a portfolio of tradable assets \( q \) to maximize his mean-variance utility:

\[
\max_q E[\rho q'(Y - p) - \frac{\rho^2}{2} q'Var(Y)q].
\]

 Unlike the security-design literature, here it is assumed that the originator cannot retain any asset. This precludes signaling and focuses on the effect of bundling on the information acquisition choice of the investors.
2.3.1 Expertise and Information Acquisition

Modeling of investors’ information acquisition is based on Van Nieuwerburgh and Veldkamp (2009). Before choosing a portfolio, each investor observes two private signals about risks $f$. One signal has exogenous precision, and the investor is to choose the precision of the other. Conditional on $f$, signals are independent across investors.

The exogenous signal models the different expertise of investors. Specifically, investor $\alpha$ of type $i$’s (hereafter $(\alpha, i)$) exogenous signal $s^{\alpha, i} \sim N(f, (\Lambda_0^i)^{-1})$, where $\Lambda_0^i = diag(\lambda_{0,1}^i, \lambda_{0,2}^i)$. It is assumed that $\lambda_{0,i}^i = \lambda > \lambda_{0,-i} > 0$, where $-i$ denotes risk(s) or type(s) other than $i$. i.e., from their exogenous signals, type $i$ investors know $f_i$ better than others.

The endogenous signal $\eta^{\alpha, i} \sim N(f, (\Lambda_\eta^{\alpha, i})^{-1})$ models investors’ information acquisition. To highlight the role of expertise, and also for tractability of belief aggregation, it is assumed that its precision matrix $\Lambda_\eta^{\alpha, i}$ is diagonal, as in Van Nieuwerburgh & Veldkamp (2009, 2010). This rules out the possibility that an investor chooses to observe a signal correlated with more than one risk. Thus, investors can choose how much to learn about each risk but are not allowed to change the risk structures.

Choosing the precision $\Lambda_\eta^{\alpha, i}$ is equivalent to choosing the precision of the posterior after observing both signals, $\Lambda^{\alpha, i} = diag(\lambda_1^{\alpha, i}, \lambda_2^{\alpha, i}) \equiv \Lambda_\eta^{\alpha, i} + \Lambda_0^i$. Each investor $(\alpha, i)$ faces two constraints in this choice:

1) A capacity constraint that limits the quantity of information carried by the endogenous signals, measured by Shannon capacity, to be no more than $K$, $K \geq 1$:

$$\prod_j \lambda_j^{\alpha, i} \leq K \prod_j \lambda_{0,j}^i. \quad (2)$$

---

5 Hereafter, superscripts index investors and subscripts index objects to learn and trade. Two-dimensional superscripts are needed to distinguish investors, as different investors of the same type may behave differently.

6 This comes from $\det(\Lambda_0^i)^{-1}/\det(\Lambda_\eta^{\alpha, i})^{-1} \leq K$. 

---

8
2) A no-forgetting constraint that prevents the investor from forgetting previous exoge-
nous information about one risk in order to free up capacity to learn about other risks:

\[ \lambda_{j}^{\alpha,i} \geq \lambda_{0,j}^{i} \forall j \]  

Note that when \( K = 1 \), the only possible choice of \( \Lambda^{\alpha,i} \) that satisfies both constraints is 
\[ \lambda_{j}^{\alpha,i} = \lambda_{0,j}^{i} \forall j, \] 
which means investors cannot acquire information.

### 2.3.2 Comparative Advantage in Information Acquisition

The capacity constraint is a bound on entropy reduction, an information measure with a long history in information theory (Shannon 1948). It is a common distance measure in econometrics (a log-likelihood ratio) and in statistics (a Kullback-Liebler distance), and is used widely in the recent economics literature on rational inattention (see Sims 2010 for a review).

A key property of this technology is that, \( K\lambda - \lambda \), the gain of signal precision by using a given capacity \( K \), increases with prior knowledge \( \lambda \). This turns initial information advantage (\( \bar{\lambda} > \lambda \)) into a comparative advantage in acquiring additional information:

1) For a given investor, the marginal gain of signal precision of one risk from additional input of capacity increases with capacity already used on it;

2) For a given risk \( f_i \), the gain of signal precision of type \( i \) is greater than that of other types from the same input of capacity.

In financial markets, information acquisition often features first-mover advantage: Basic background knowledge, skills and equipment have to be developed or acquired upfront before getting to know about a particular industry or asset class. This turns initial information advantage into a comparative advantage in acquiring additional information: 1) The increase in familiarity with a particular industry or asset class makes it much easier to acquire new information about it; 2) Such first-mover advantage makes it easier for an expert in a par-
ticular industry or asset class to acquire new knowledge about his area of expertise than an ordinary market participant; 3) This is also a major reason for the difference in the expertise of market participants, which is a primitive of this paper. The learning technology in the model captures such first-mover advantage and the resulting comparative advantage in acquiring new information.

2.3.3 Portfolio Choice

Investors trade the assets available as in the markets of Admati (1985). Before portfolio choice, each investor observes the realization of his private signals and market clearing price(s) $p$ of the tradable assets. In equilibrium, the price(s) $p$ serves as an additional endogenous signal of the payoff(s) of these assets $Y$. The investor updates his belief about $Y$ using Bayes Law and decides how much of each asset to buy, $q^a_i$, to maximize his utility (equation 1). The technical details of the pricing formula and of investors’ portfolio choice are given in the appendix.

2.3.4 The Role of Preference

The mean-variance preference (equation 1) follows from risk aversion at the trading stage and from preference for early resolution of uncertainty at the learning stage. Specifically, an investor’s utility function can be expressed as $U = E_1[u_1(E_2[u_2(W)])]$, where $W = W_0 + q'(Y - p)$ denotes terminal wealth, the sum of initial wealth $W_0$ and profit from portfolio investment.

Time 2 refers to the trading stage. $u_2(W) = -\exp(-\rho W)$. $u''_2 < 0$ governs the investor’s risk aversion at the trading stage.

Time 1 refers to the learning stage. $u_1(x) = -\log(-x)$. Since $u''_1 > 0$, the investor prefers early resolution of uncertainty before the trading stage: At the learning stage, the investor anticipates that the additional information gained later may signal either high or low expected utility $E_2[u_2(W)]$ that he will enjoy at the trading stage. Therefore, at the
learning stage, the investor sees \( E_2[u_2(W)] \) as a random variable, and has expected utility \( E_1[u_1(E_2[u_2(W)])] \). If the investor cannot see the additional information before trading, his expected utility at the learning stage is \( E_1[u_1(u_2(W))] \). Since \( u''_1 > 0 \), Jensen’s inequality implies \( E_1[u_1(E_2[u_2(W)])] > E_1[u_1(u_2(W))] \); i.e., the investor likes to resolve uncertainty by learning before the trading stage.

The preference for early resolution of uncertainty at the learning stage makes the investor choose to learn more about those risks he expects to hold more of at the trading stage. His risk aversion at the trading stage makes him hold more of the risks he knows better. These two preferences form a feedback loop and reinforce each other, pressuring the investor to specialize in learning about a single tradable risk. This is the impetus for specialization in information acquisition in the model.

### 2.4 The Liquidity Trader

As in a standard rational expectations equilibrium model (e.g., Grossman and Stiglitz 1980), traders who trade assets for non-speculative reasons, such as liquidity needs or to hedge risk exposure outside the model, are needed to prevent investors from being able to perfectly infer the private information of others from prices and thus having no need to acquire any private information themselves. A representative liquidity trader ("she") is therefore introduced, whose liquidity demand for risks \( f \) is \( \varepsilon_f \sim N(0, \sigma^2 I) \). This implies that the liquidity trader’s demand for original assets \( X \) is \( \varepsilon = \Gamma^{-1} \varepsilon_f \sim N(0, \sigma^2 \Gamma^{-1} \Gamma' -1) = N(0, \sigma^2 I) \), with the last equality due to the orthogonality of \( \Gamma \).

In the model, a bundling strategy \( T \) may restrict the tradable asset span, making the liquidity trader’s desired portfolio of original assets unfeasible. In this case, it is assumed that she chooses the closest available substitute to fulfill her liquidity demand. That is, her demand \( \varepsilon_T \) for tradable asset(s) \( Y = TX \) is assumed to be the linear projection of her desired portfolio \( \varepsilon'X \) onto the tradable assets span: \( \varepsilon_T = (TT')^{-1}T\varepsilon \sim N(0, \sigma^2(TT')^{-1}) \).

---

7 The liquidity demand here can also be interpreted as hedging demand due to exposure \( -\varepsilon_f \) to risks \( f \) outside the model.
2.5 Equilibrium

We say \( \{ T, \{ \Lambda^{\alpha,i} \}, \{ q^{\alpha,i} \}, p_T \} \) is an equilibrium iff:

1) The bundling strategy \( T \) maximizes the originator’s payoff \( E_0[1'p_T] \);

2) Given the originator’s bundling strategy \( T \) and the distribution of his exogenous signal \( s^{\alpha,i} \), each investor \( (\alpha, i) \)’s choice of information acquisition \( \Lambda^{\alpha,i} \) and portfolio choice \( q^{\alpha,i} \) maximizes his utility (equation 1), subject to the capacity constraint (equation 2) and the no-forgetting constraint (equation 3);

3) Given every investor’s portfolio choice \( \{ q^{\alpha,i} \} \), prices \( p_T \) clear the market: \( \int_{\alpha,i} q^{\alpha,i} + \varepsilon_T = 1 \) ; and

4) Beliefs are updated using Bayes’ law, and expectations are rational; i.e., ex ante beliefs about \( q^{\alpha,i} \) are consistent with the true distribution of the optimal portfolio.

We say \( \{ \{ \Lambda^{\alpha,i} \}, \{ q^{\alpha,i} \}, p_T \} \) is a subgame equilibrium induced by a given bundling strategy \( T \) iff conditions 2) to 4) hold.

For tractability, we consider only linear equilibria, in which price(s) \( p_T \) are linear functions of payoff(s) \( Y \) and liquidity trader’s demand \( \varepsilon_T \).

2.6 Summary of Model Setup

The following timeline summarizes the model setup:

<table>
<thead>
<tr>
<th>Timeline</th>
<th>Originator</th>
<th>Investors</th>
<th>Liquidity trader</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Chooses bundling strategy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Decide which information to acquire</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Sells assets</td>
<td>Observe signals and choose portfolio</td>
<td>Demands assets</td>
</tr>
<tr>
<td>3</td>
<td>Consumes payoff</td>
<td>Consume payoff</td>
<td></td>
</tr>
</tbody>
</table>

In principle, the originator can use a continuum of bundling strategies to create two tradable assets. The following proposition shows that they are all equivalent, and thus it suffices to compare two strategies: i) \( T = I \), selling the original assets as they are, and ii)
\[ T = 1', \] pooling them together into a single asset.

**Proposition 2.1** *The investor’s information acquisition problem and the originator’s payoff are invariant to different bundling strategies that lead to the same tradable asset span.*

The proofs of this and all subsequent propositions are given in the appendix.

Intuitively, if different bundling strategies create the same tradable asset span, each investor’s choice set of feasible portfolios is invariant to these bundling strategies. Therefore, his problems of portfolio choice and information acquisition also remain the same. As a result, his decisions in the learning stage and the trading stage do not change, and neither does the total risk premium demanded.

For later discussion, for each risk \( j = 1, 2 \), define \( \lambda_{j,T}^{a,T} \equiv \int_{\alpha,i} \lambda_{j,T}^{a,i} \), market average signal precision of risk \( f_j \) induced by bundling strategy \( T \). And \( \Lambda_T^T \equiv diag(\lambda_{1,T}^{a,T}, \lambda_{2,T}^{a,T}) \). Subscript \( T \) is suppressed if no confusion is caused.

### 3 The Upside of Bundling: The Discipline Channel

To illustrate the upside of asset bundling — the discipline channel — I use the polar case in which total payoff of the assets for sale depends only on a single risk: \( X_1 + X_2 = \sqrt{2} f_1 \). In the appendix, I show that qualitatively similar results hold as long as the contribution of risk \( f_2 \) is sufficiently low.

Consider a mortgage lender, who has originated and wants to sell mortgages on all apartments in New York. To him, \( f_1 \) corresponds to common shocks to the prices of all these apartments, and \( f_2 \) to shocks specific to the price of a single building whose contribution to the total value of the mortgages for sale is negligible.

As discussed in Proposition 2.1, we need to compare only two bundling strategies: \( T = I \), selling the original assets as they are, and \( T = 1' \), pooling them together.

The following proposition further simplifies the analysis, which shows that, to compare the seller’s payoffs in the subgames induced by the two bundling strategies respectively, it
suffices to compare the corresponding market average signal precision of risk $f_1$ induced:

**Proposition 3.1** If $X_1 + X_2 = \sqrt{2}f_1$, for both $T = I$ and $T = 1'$, the originator’s payoff is

$$E_0[\sum X_i] - 2\rho\left[\frac{1}{\rho^2 \sigma^2} (\lambda_{i,T}^a)^2 + \lambda_{i,T}^a\right].$$

The originator’s payoff depends only on $\lambda_{i,T}^0$ but not on $\lambda_{2,T}^0$, since his net supply of $f_2$ is zero. And his payoff is a strictly increasing function of $\lambda_{i,T}^0$, because investors demand a lower risk premium for holding $f_1$ in equilibrium if, on average, they face less of such risk.

An immediate result is:

**Corollary 3.1** If $X_1 + X_2 = \sqrt{2}f_1$ and $K = 1$, then $T = I$ and $T = 1'$ generate the same payoff to the originator.

That is, when investors cannot acquire information, the originator is indifferent between bundling the assets and selling them as they are, because the investors’ knowledge of $f_1$ is exogenous.

We now characterize how investors acquire information following the two bundling strategies respectively.

**Proposition 3.2** shows that pooling the assets induces the originator’s desired information acquisition behavior in the investors:

**Proposition 3.2** If $X_1 + X_2 = \sqrt{2}f_1$, in the unique subgame equilibrium induced by $T = 1'$, every investor learns only about $f_1$, regardless of investor type.

The reason behind this result is intuitive: when the original assets are pooled together, the unique new asset formed has payoff $Y = X_1 + X_2 = \sqrt{2}f_1$; i.e., the diversifiable risk $f_2$ is washed out. Thus, each investor’s portfolio choice problem is simply how much of $f_1$, the non-diversifiable risk to take. Anticipating that, all investors know in advance that they can benefit only from information about $f_1$, and thus in equilibrium they only acquire such information. This holds regardless of their expertise.
Since this is the dominant strategy for every investor, the subgame equilibrium induced is unique.

Note that in this subgame equilibrium, $\lambda_{1,T}^n$, the market average signal precision of $f_1$, reaches the greatest possible level. As a result, bundling strategy $T = I$, selling the original assets as they are, can do no better than pooling them. Indeed, the following proposition indicates that selling the original assets as they are is strictly inferior to pooling them when investors have a large enough capacity $K$:

**Proposition 3.3** If $X_1 + X_2 = \sqrt{2}f_1$, in the unique subgame equilibrium induced by $T = I$, each investor learns about only one risk, respectively, and

1) all type 1 investors learn only about $f_1$.

2) $\exists K_0 < \infty$ such that a positive proportion of type 2 investors learn about $f_2$ if $K > K_0$.

Although $f_2$ is diversifiable in aggregation, the loading of each asset on it is generally not zero, like the shock specific to the single building in the example at the beginning of this section. Indeed, given the full rank of the risk-loadings matrix $\Gamma$, when assets are sold as they are, investors could hold any amount of any risk in their portfolios. This allows them to profit from their private information about any risk.

As discussed in Section 2.3.4, an investor’s preference for early resolution of uncertainty and his risk aversion make him specialize in learning about a single risk. In addition, as discussed in Section 2.3.2, any given investor’s marginal gain of signal precision of a risk from additional input of capacity increases with the capacity already used on that risk. This further strengthens his incentive to specialize in learning about a single risk. So now the question is: which risk would he choose?

Investors face two concerns when choosing which risk to trade and learn about. First, only $f_1$ is non-diversifiable and carries a premium in equilibrium, which attracts investors to hold and learn about it. Second, investors want to have information about a risk that is better than the market average: The price of a risk reflects only the knowledge of an average
market investor. An investor’s superior information of the risk helps him take advantage of others who know less about it when trading and generates excess return. Therefore, an investor wants to learn about risks studied by fewer people. This is strategic substitutability in information acquisition, which attracts each investor to trade and learn about the risk in which he has expertise.

These two concerns work in the same direction for type 1 investors, so they must dedicate all their capacity to $f_1$ in equilibrium.

However, these concerns work in the opposite direction for type 2 investors, whose expertise is in $f_2$ instead of $f_1$. When assets are not pooled, it might be rational for some of them to learn about $f_2$.

Consider a type 2 investor, and assume that everyone but him learns only about $f_1$. Since he has expertise in $f_2$, he also has a comparative advantage in learning about it, as discussed in Section 2.3.2. Thus, he is informationally advantageous in $f_2$ and disadvantageous in $f_1$. When everyone has low capacity $K$, investors, on average, still face significant uncertainty about $f_1$ after learning, and thus the premium carried by $f_1$ may still be able to attract this type 2 investor to hold it instead of $f_2$, and to learn about it to minimize his informational disadvantage. However, when capacity $K$ becomes large, the premium carried by $f_1$ decreases (to 0 when $K \rightarrow \infty$), and at the same time, the investor’s comparative advantage in learning about $f_2$ becomes larger and larger. Since others have not yet learned about $f_2$, he would prefer to learn about it in order to exploit others with his superior information when trading.

However, from the originator’s perspective, the fact that his net supply of $f_2$ is zero implies that the capacity used to learn about it is a waste of resources. Thus, we have:

**Proposition 3.4** If $X_1 + X_2 = \sqrt{2} f_1$, in equilibrium the originator chooses $T = 1^\prime$, pooling the assets.

Back to the mortgage lender mentioned at the beginning of this section. He is better off pooling all his mortgages than selling them separately, because pooling prohibits those
mortgage buyers who know one particular building better than others from cherry-picking its mortgage and profiting from information about it. Instead, their attention is drawn to shocks common to all the mortgages for sale, which affects the risk premium. We name this beneficial channel of asset bundling the discipline channel.

### 4 The Downside of Bundling: The Trade-Restriction Channel and the Specialization-Destruction Channel

What if a bank simultaneously issues and wants to sell off loans of different asset classes, say mortgages and credit cards, that make similar contributions to the total value of the loans? This section shows that the bank is better off selling loans of different classes separately.

Formally, we consider the other polar case in which the two risks contribute equally to the total payoff of the original assets: \( X_1 + X_2 = f_1 + f_2 \). Each risk can be thought of as common shocks to a different asset class. It is shown in the appendix that qualitatively similar results hold as long as the contributions of the two risks are sufficiently close.

Again, without loss of generality, we consider only two bundling strategies, \( T = I \), selling the original assets as they are, and \( T = 1' \), pooling them together. In this context, pooling the assets creates a new asset with payoff \( Y = X_1 + X_2 = f_1 + f_2 \), which implies that each investor has to hold an equal amount of \( f_1 \) and \( f_2 \).

Proposition 4.1 states the originator’s payoffs from the two bundling strategies, respectively.

**Proposition 4.1** Let \( g(x) = E_0[\sum X_i] - 2\rho[x + \frac{1}{\rho}x^2]^{-1}, x > 0 \). If \( X_1 + X_2 = f_1 + f_2 \),

1) The originator’s payoff from choosing \( T = I \) is \( g(\frac{K\lambda + \lambda}{2}) \);

2) If \( K \geq \frac{\lambda}{\lambda} \), the originator’s payoff from choosing \( T = 1' \) is \( g(\sqrt{K\lambda\lambda}) \);

3) If \( K < \frac{\lambda}{\lambda} \), the originator’s payoff from choosing \( T = 1' \) is \( g[(\frac{(K\lambda)^{-1} + \frac{1}{\lambda}}{2})^{-1}] \).
It is easy to see that \( g \) is a strictly increasing function. And by the inequality of arithmetic and geometric means, \( 0 < \left( \frac{(K\lambda\bar{\lambda})^{-1} + \bar{\lambda}^{-1}}{2} \right)^{-1} \leq \sqrt{K\lambda\bar{\lambda}} < \frac{K\lambda + \bar{\lambda}}{2} \). Therefore, pooling the assets \((T = 1')\) is strictly inferior for the originator. The following proposition formally states the result of the comparison:

**Proposition 4.2** If \( X_1 + X_2 = f_1 + f_2 \), the originator is strictly better off choosing \( T = I \) instead of \( T = 1' \).

Two different economic forces lead to the deficiency of bundling: the *trade-restriction channel* and the *specialization-destruction channel*.

### 4.1 The Trade-Restriction Channel

The trade-restriction channel is mechanical and is not related to information acquisition. Thus we illustrate it by shutting down learning; i.e., by considering \( K = 1 \). From Proposition 4.1, we can see that the originator’s payoff from selling the original assets as they are is \( g(\frac{\bar{\lambda}_1 + \bar{\lambda}_2}{2}) \), while his payoff from pooling the assets is \( g(\frac{\bar{\lambda}_1 + \bar{\lambda}_2}{2})^{-1} \), which is strictly lower.

In Corollary 3.1, when investors cannot acquire information, pooling the assets or not generates the same payoff to the originator. But here, pooling yields a strictly lower payoff because each investor knows one risk better than the other because of his particular expertise; i.e., from his exogenous signals, and thus wants to trade that risk more aggressively than the other. But this is precluded by the bundling strategy of pooling. The following proposition characterizes investors’ expected holdings of risks, and shows that bundling \((T = 1')\) results in a less efficient allocation of risks across investors than not bundling \((T = I)\):

**Proposition 4.3** If \( X_1 + X_2 = f_1 + f_2 \) and \( K = 1 \), \( \forall i \)

1) If \( T = I \), then each type \( i \)’s expected holding of risk \( f_i \) and \( f_{-i} \) are \( \frac{\lambda_p + \lambda_p}{\frac{\lambda_1 + \lambda_2}{2} + \lambda_p} \) and \( \frac{\lambda_1 + \lambda_2}{2} + \lambda_p \), respectively, where

\[
\lambda_p = \frac{1}{\sigma^2 \bar{\lambda}_i} \left( \frac{\lambda_1 + \lambda_2}{2} \right)^2;
\]

2) If \( T = 1' \), then each type \( i \)’s expected holding of risk \( f_i \) and \( f_{-i} \) are both 1.
The two fractions in 1) have straightforward economic meanings. The numerators $(\bar{\lambda} + \lambda_p)$ and $(\bar{\Lambda} + \lambda_p)$ are a type $i$ investor’s knowledge about $f_i$ and $f_{-i}$, respectively: $\bar{\lambda}$ (or $\bar{\Lambda}$) from his private signal, and $\lambda_p$ from the prices. Similarly, the denominators are an average market investor’s knowledge about each risk.

Intuitively, type $i$ investors know risk $f_i$ better than the other type, and are willing to hold $f_i$ for a lower risk premium. The originator is therefore better off having them hold more of $f_i$. When assets are bundled, an investor is restricted to holding equal amounts of $f_1$ and $f_2$. Since the knowledge of risks is symmetric across different types of investor, and since both risks contribute symmetrically to the payoff of the single tradable asset $Y = X_1 + X_2 = f_1 + f_2$, each investor in expectation takes an equal share of each risk. However, when assets are not bundled, investors can freely trade any risk. Since they are risk averse, type $i$ investors would choose to hold more of $f_i$, the risk they know better, and less of $f_{-i}$, the risk they know less. This can be seen from $\frac{\bar{\lambda} + \lambda_p}{\frac{\bar{\lambda} + \lambda_p}{2} + \lambda_p} > 1 > \frac{\bar{\Lambda} + \lambda_p}{\frac{\bar{\Lambda} + \lambda_p}{2} + \lambda_p}$, as $\bar{\lambda} > \frac{\bar{\lambda} + \bar{\Lambda}}{2} > \bar{\Lambda}$.

The trade-restriction channel is driven by the differences in each investor’s knowledge of different risks. If each investor knows each risk equally well ($\bar{\lambda} = \bar{\Lambda}$), then $\frac{\bar{\lambda} + \lambda_p}{\frac{\bar{\lambda} + \lambda_p}{2} + \lambda_p} = 1 = \frac{\bar{\Lambda} + \lambda_p}{\frac{\bar{\Lambda} + \lambda_p}{2} + \lambda_p}$. Thus bundling or not bundling yields the same allocation of risks across investors. The following proposition further confirms this point by showing that if each investor knows each risk equally well, the originator is indifferent between pooling the assets or not:

**Proposition 4.4** If $X_1 + X_2 = f_1 + f_2$, $K = 1$ and $\bar{\lambda} = \bar{\Lambda}$, then the originator’s payoff is $E_0[\sum X_i] - 2\rho[\frac{1}{\rho^2\sigma^2}\bar{\lambda}^2 + \bar{\lambda}]^{-1}$, whether $T = 1$ or $T = 1'$ is chosen.

### 4.2 The Specialization-Destruction Channel

The specialization-destruction channel affects the originator’s payoff through its impact on the information acquisition behavior of investors. We now characterize how investors acquire information in the subgames engendered by the two bundling strategies.

Proposition 4.5 shows that, if the original assets are sold separately, each investor focuses on his area of expertise:
Proposition 4.5 If $X_1 + X_2 = f_1 + f_2$, in the unique subgame equilibrium induced by $T = I$, each type $i$ investor learns only about $f_i, \forall i$.

Intuitively, when assets are not bundled, investors can freely trade individual risks. As discussed in Proposition 3.3, each investor devotes all his capacity to only one risk. Here, both risks play a symmetric role and carry the same premium in equilibrium, so strategic substitutability in information acquisition determines that each investor specializes in his area of expertise and also determines the uniqueness of subgame equilibrium.

What happens if the assets are pooled, $T = 1'$? Proposition 4.6 shows that investors are then induced to spend most of their capacity on the risk in which they have no expertise:

Proposition 4.6 If $X_1 + X_2 = f_1 + f_2$, in the unique subgame equilibrium induced by $T = 1'$, each investor tries his best to equalize his knowledge of different risks:

1) If $K \geq \bar{\lambda}/\lambda > 1$, then $\lambda_{i}^{\alpha, i} = \lambda_{2}^{\alpha, i} = \sqrt{K \bar{\lambda} \lambda}$, $\forall \alpha, i$;

2) If $1 \leq K < \bar{\lambda}/\lambda$, then $\lambda_{i}^{\alpha, i} = \bar{\lambda}$, $\lambda_{i}^{\alpha, i} = K \lambda$, $\forall \alpha, i$.

Anticipating that he has to hold equal amounts of each risk, each investor tries his best to equalize his knowledge of different risks through information acquisition, in order to adapt to the trading restriction. Such equalization can be achieved perfectly only if the investor’s capacity reaches a threshold $\bar{\lambda}/\lambda$, which depends on the magnitude of his expertise. When his capacity is below the threshold, he devotes all his capacity to the risk in which he has no expertise. Since this is the dominant strategy for every investor, the subgame equilibrium...

---

Footnote 8: Here, the result that every investor’s optimal choice of capacity allocation is his dominant strategy is peculiar to the setup of the trading stage based on Admati (1985). Following Van Nieuwerburgh and Veldkamp (2009, 2010), the same setup based on Admati (1985) is used to model the trading stages following different bundling strategies to guarantee that they are comparable to each other. Here, if $X_1 + X_2 = f_1 + f_2$ and assets are pooled, according to Admati (1985), the price of the bundle takes the form $p = A + Y + C \varepsilon = A + X_1 + X_2 + C \varepsilon = A + f_1 + f_2 + C \varepsilon$. This implies that the price informativeness of $f_1$ and $f_2$ are by construction the same. As a result, the learning complementarities in Goldstein and Yang (2015) do not hold here: From the point view of an investor, when a greater number of other investors choose to learn about $f_1$ rather than $f_2$, the price will not reveal $f_1$ more than $f_2$ and further affect his own capacity allocation. If we allow the price informativeness of $f_1$ and $f_2$ to vary endogenously and differentially as in Goldstein and Yang (2015) or Bond and Goldstein (2015), then an investor’s optimal capacity allocation here may no longer be his dominant strategy, but the key economic force emphasized here, the specialization-destruction channel, still persists.
induced is unique.

When assets are not pooled, each investor focuses on acquiring information in his area of expertise, and his comparative advantage is fully utilized. If assets are pooled, however, each investor expends most of his capacity on the risk in which he has no expertise. As a result, similar to the implication of comparative advantage theory in international trade, investors, on average, face more residual uncertainty after learning about every risk when assets are pooled, which leads to a higher risk premium in equilibrium. We call this adverse channel of asset bundling the specialization-destruction channel.

As capacity $K$ increases, investors are more able to adapt their knowledge to the trading restriction, and thus the trade-restriction channel weakens and the specialization-destruction channel strengthens. When $K < \frac{\lambda}{\Delta}$, each investor lacks the capacity to equalize his knowledge of each risk, and both channels are in play. When $K \geq \frac{\lambda}{\Delta}$, investors have enough capacity to achieve perfect equalization of knowledge. In this case, the trade-restriction channel completely disappears, and only the specialization destruction channel plays a role.

Therefore, the bank at the beginning of this section is better off selling the two asset classes separately. This allows the mortgage specialists among the investors to trade mortgages more aggressively relative to credit card loans, and thus induces them to focus on acquiring information about their specialty. This reduces the residual uncertainty faced by investors on average after learning about the mortgages being sold, and thus lowers the risk premium demanded. A symmetric argument applies to credit card specialists.\footnote{As in Van Nieuwerburgh and Veldkamp (2009), specialization in information acquisition does not imply specialization in risk holding. In equilibrium, type $i$ investors still want to hold some $f_{-i}$ for diversification of liquidity trader risk, which is assumed to be i.i.d. across risks.}
5 Discussion

5.1 A Generalization: The Optimality of Categorization Strategy

This subsection demonstrates that the economic forces illustrated in the previous two sections carry through to more general environments. We generalize the baseline model to an $n$-risk-$n$-asset setup with $n$ corresponding types of investor: the payoffs of the original $n$ assets are $X = \Gamma f$, where $\Gamma$ is an $n$-by-$n$ orthogonal risk loading matrix. Each type of investors has mass $1/n$. Each type $i$ investor’s exogenous signal $s^{\alpha,i} \sim N(f_i, (\Lambda_0^{(i)})^{-1})$, $\Lambda_0^{(i)} = \text{diag}(\lambda_0^{(i,1)}, \ldots, \lambda_0^{(i,n)})$, $\lambda_0^{(i,j)} = \tilde{\lambda} > \lambda = \lambda_0^{(i,j)} \forall j \neq i$. Each bundling strategy that creates $1 \leq m \leq n$ tradable assets is uniquely represented by a full rank $m \times n$ matrix $T_{m \times n}$ such that $1_m' T = 1_n'$ and that the tradable assets have payoffs $Y = TX$. Each tradable asset has a supply of 1. Everything else is analogous to the baseline model.

Proposition 2.1 shows that the essence of the choice of bundling strategies is the resulting tradable asset span. In the appendix it is demonstrated that this proposition also holds in this generalized setup. In the baseline model, the originator effectively has only two feasible choices of tradable asset spans: either to allow investors to freely trade any amount of the two risks, or to restrict investors to trading equal amounts of both. The $n$-risk generalized setup significantly expands the originator’s set of feasible choices of tradable asset spans to a continuum.

We consider the intermediate case: $\exists 1 \leq i^* \leq n$, such that $\forall i \leq i^*, w_i = w > 0$, $^{10}$and $w_i = 0 \ \forall i > i^*$. That is, $f_1, \ldots, f_{i^*}$ are non-diversifiable and play symmetric roles, while $f_{i^*+1}, \ldots, f_n$ are diversifiable. This corresponds to the scenario in which a bank tries to sell loans of $i^*$ different asset classes, each with many loans and each contributing similarly to the total value of the loans for sale. This nests in the two special cases discussed in the previous two sections, in which $n = 2$, and $i^* = 1$ and 2, respectively.

The aim of this subsection is to establish the optimality of categorization strategy, which

---

$^{10}$The orthogonality of loading matrix $\Gamma$ implies $w = \sqrt{n/i^*}$.
corresponds to pooling loans by asset class in the context of securitization and is defined formally as follows:

**Definition 5.1** *Categorization strategy is represented by the \((i^* \times n)\) dimensional matrix \(T\) such that:*

\[
T = \begin{pmatrix}
    \begin{pmatrix}
        \begin{pmatrix}
            w & 0 & \cdots & 0 & 0
        
    \end{pmatrix} & 0 & \cdots & 0 & 0 \\
    0 & w & \ddots & \vdots \\
    \vdots & \ddots & \ddots & 0 \\
    0 & \cdots & 0 & w & 0
    \end{pmatrix}
\end{pmatrix}
\]

This strategy creates \(i^*\) tradable assets, such that \(Y = TX = T\Gamma = (w_f_1, w_f_2, ..., w_f_{i^*})'\). It removes all the diversifiable risks asset-by-asset, and each tradable asset takes all the loading of a different non-diversifiable risk.

To establish global payoff optimality, we should ideally compare this strategy with all its opponents. However, the information acquisition problem for each individual investor following an arbitrary bundling strategy \(T\) is intractable. So instead, I prove that categorization strategy achieves a weaker sense of optimality: it implements the capacity allocation and achieves the originator’s payoff of an optimality benchmark. In this hypothetical benchmark, before investors trade assets, the originator could directly force them to acquire information in the way that maximizes his payoff, instead of indirectly inducing them to do so by means of asset bundling as previously discussed. This benchmark is meant to capture the best outcome the originator can achieve by affecting how investors acquire information about his assets.

Formally, this optimality benchmark is defined as:

\[
\max_{\{\lambda^{\alpha,i}\}} E_0[\sum_i p(X_i)] = E_0[1'p_{\mathbf{I}_n}] \tag{4}
\]

subject to capacity constraint (2) and no-forgetting constraint (3) \(\forall \alpha, i\)
In other words, this benchmark seeks to solve the following problem: Suppose the originator must sell the original assets as they are, but can \textit{directly assign} a feasible capacity allocation to each investor before he makes his portfolio choice, what is the optimal assignment that maximizes the originator’s payoff?

This benchmark is considered for the following reasons.

First, the main focus of this paper is to study how the asset originator induces investors to acquire information in a way that maximizes his profit. This benchmark explicitly highlights such a consideration.

Second, one can also rationalize this approach with a bounded-rationality story: As is the case for economists, it is too complicated for the asset originator in this model to compare the whole continuum of bundling strategies one by one. He therefore takes a shortcut: He first determines his desired feasible capacity allocation for each investor and then checks whether a simple and commonly used bundling strategy could induce that allocation.

Third, in the intermediate case, this benchmark can be analytically solved, and its unique solution has a clear economic interpretation.

Fourth, the result that categorization strategy implements the benchmark in the intermediate case also has a clear economic interpretation, which combines the intuitions introduced in Sections 3 and 4.

The following proposition characterizes the optimality benchmark:

\textbf{Proposition 5.1} If $1 \leq i^* \leq n$, such that $\forall i \leq i^*, w_i = w > 0$, and $w_i = 0 \forall i > i^*$, then the solution to the optimality benchmark is such that:

1) each investor of type $i \leq i^*$ specializes in learning about $f_i$; and

2) each investor of type $i > i^*$ specializes in one non-diversifiable risk $j \leq i^*$, such that there is an equal mass of investors specializing in learning about each such risk.

Intuitively, a solution to the optimality benchmark completely utilizes the expertise of investors on non-diversifiable risks. Since the average precision of private information about
diversifiable risks does not enter the objective function, no capacity should be spent on them. Having each type \( i > i^* \) investor specializing in exactly one non-diversifiable risk takes advantage of comparative advantage in information acquisition. Last, since all non-diversifiable risks carry equal weight in the objective function, and the objective function is concave in \( \lambda_i \), each non-diversifiable risk should receive the same capacity.

The next proposition gives the conclusion of this subsection: Categorization strategy implements the capacity allocation and the originator’s payoff of the optimality benchmark.

**Proposition 5.2** If \( \exists 1 \leq i^* \leq n \), such that \( \forall i \leq i^* , w_i = w > 0 \), and \( w_i = 0 \ \forall i > i^* \), then in any equilibrium induced by categorization strategy, the aggregate capacity allocation and the resulting originator’s payoff are the same as the solution to the optimality benchmark.

The intuition of this result combines that developed in Sections 3 and 4. The removal of diversifiable risks asset-by-asset prohibits investors from taking them, and deters information acquisition about them. This employs the beneficial discipline channel. The full span on non-diversifiable risks allows investors to take any amount of any of them, and induces perfect specialization in information acquisition about them. This avoids the harmful trade-restriction and specialization-destruction channels.

### 5.2 Relation to the Security-Design Literature

Now that the three main economic forces in my model have been illustrated, we are in a good position to discuss the model’s relation to the literature on security design. Bundling loans into different pools and issuing securities backed by them is the defining characteristic of securitization, which plays a significant role in the U.S. economy. Originators and investors typically have asymmetric information (as in my model), raising the concern of adverse selection and moral hazard. In the literature relevant to securitization, the theoretical literature on security design probes how to mitigate such information friction. Section 6 of Gorton and Metrick (2013) provides an excellent survey. In addition to rationalizing the
feature of pooling loans by asset class introduced at the beginning of Section 1, my model also contributes to this literature in the following aspects:

First, for simplification, by assuming a 1-seller-1-buyer setup, existing security design models (e.g. Demarzo and Duffie 1999; Demarzo 2005; Farhi and Tirole 2014) typically abstract from the interaction of heterogeneous security buyers, which is the core of my model. Indeed, the beneficial discipline channel of asset bundling introduced in Section 3 is achieved by effectively prohibiting investors with expertise in idiosyncratic risks from using their superior information to exploit others.

Second, a typical security-design model in which the seller has information advantage over the buyer (e.g. Townsend 1979; Dang et al. 2013; Yang 2013) looks at how to deter or reduce the buyer’s costly information acquisition, which is wasteful from a welfare perspective. My model augments this by addressing a complementary question: Given that the buyers ("investors") have decided to expend a fixed amount of resources ("capacity") to acquire information, how can the seller ("originator") induce them to do so in his preferred way? Indeed, the harmful specialization-destruction channel of asset bundling introduced in Section 4 results precisely from the waste of capacity on risks that investors have no expertise in studying.

5.3 Optimal Bundling Strategies May Not Favor Investors

So far, we have been focusing on maximization of originator’s payoff. By definition the respective optimal bundling strategies maximize the originator’s utility. Since the fundamentals of these assets $E_0[\sum Y_i] = E_0[\sum X_i]$ are exogenous, the respective optimal bundling strategies also maximize the market liquidity of the risky assets, in the sense of minimizing the expected total price discount, $E_0[\sum (Y_i - P_i)]$. A natural follow-up question would be: in general, do the respective optimal bundling strategies that we have identified generate a Pareto improvement relative to the subgame equilibria induced by other bundling strategies? This subsection shows that the answer is no.
Besides the originator, there are two types of agents in the model: the investors and the representative liquidity trader. Since the liquidity trader has no well-defined utility function, the following welfare analysis will focus on the investors.

The respective optimal bundling strategies minimize the risks faced by investors and in turn the total risk premium. Since investors are risk averse, one might think that, in general, optimal bundling strategies also maximize investors’ utility. However, it turns out that this conjecture is wrong.

The following proposition uses the case studied in Section 3 (in which total payoff of the assets depends only on a single risk $f_1$) to show that there could be a conflict of interest between the originator and investors:

**Proposition 5.3** If $X_1 + X_2 = \sqrt{2}f_1$, the average expected utility of investors in the subgame induced by $T = 1'$ is lower than that induced by $T = I$.

In Section 3, it was shown that, if there is only one non-diversifiable risk, the originator prefers pooling the assets ($T = 1'$) to selling the assets as they are ($T = I$). One might think that pooling the assets prohibits investors from exploiting each other with regard to the diversifiable risk $f_2$, and thus, should make them better off. However, it turns out that, on average, investors are actually worse off, for three reasons: First, while pooling the assets reduces the average uncertainty investors face about $f_1$ after learning, it also reduces the premium they can earn by holding it, and the latter outweighs the former for mean-variance investors\(^{11}\). Second, investors profit from the liquidity trader’s demand, because the liquidity trader always moves price(s) against herself. If investors face less uncertainty about $f_1$, their demand for it becomes more elastic, so that the liquidity trader’s demand causes less movement in the price of $f_1$, and thus investors profit less from her. Third, pooling also restricts the asset span available to the liquidity trader, making her less aggressive in

\(^{11}\)The proof of Proposition 9.1 in the appendix shows that an investor’s expected utility is $\frac{1}{2}E\{(Y - p)’[Var(Y)]^{-1}(Y - p)\}$; i.e., expected return enters quadratically in the numerator, while variance enters linearly in the denominator.
fulfilling her "true" demand for the original assets, and causing her to lose less to investors.\textsuperscript{12}

6 Asset Bundling and Corporate Diversification

Although the theoretical model in this paper is motivated by securitization, the key economic forces highlighted also work in other contexts. By relabeling the asset originator in the model as an entrepreneur who owns several lines of business and decides how to set the firm boundaries, the model provides an alternative perspective of corporate diversification. This section discusses this perspective and its relation to existing empirical evidence.

This section is not meant to test the model against existing theories of conglomerate formation, but rather serves two different purposes. First, it shows that although the model is motivated by a feature of securitization, the main economic channels highlighted also apply in other contexts. Second, it shows that by taking the model’s perspective a conceptual connection can be built between securitization and conglomerate formation, two issues that are both important in their own right but seemingly remote from each other conceptually.\textsuperscript{13}

6.1 A Market-Side Theory of Corporate Diversification

"Conglomerate firm production represents more than 50 percent of production in the United States. Given the size of production by conglomerate firms, understanding the costs and benefits of this form of organization has important implications...For corporate finance, the primary questions about diversification are: 'When does corporate diversification affect firm value?' and 'When diversification adds value, how does it do so?'" (Maskimovic and Phillips

\textsuperscript{12}If $X_1 + X_2 = f_1 + f_2$, the third effect works in the opposite direction to the first two, making the impact of the optimal bundling strategy on investors ambiguous.

\textsuperscript{13}Economists have been interested in building such a connection. For example, Leland (2007) focuses on the pure financial synergies of corporate mergers and decomposes them into an "LL" Effect and a Leverage Effect. The former is due to the loss of separate limited liability and is always negative. The latter results from the change in optimal leverage and could be either positive or negative, depending on whether the larger tax benefit outweighs the greater cost of financial distress. The tradeoff between them yields an optimal financial scope of the firm. The same analysis is also applied to provide an explanation for structured finance. My model provides an alternative perspective.
The literature on corporate diversification took off with the discovery by Lang and Stulz (1994) and Berger and Ofek (1995) of the diversification discount: a typical conglomerate is valued by the stock market at a discount compared with a collection of comparable single-segment firms. This discount represents an economically important puzzle. Consequently, a large number of studies tried to explain the diversification discount and determine whether the discount is a real empirical phenomenon or an artifact of the measurement process. Maksimovic and Phillips (2013) provide a comprehensive survey of these two strands of literature.

The firm boundary of a conglomerate can be viewed from two complementary perspectives. One is from inside the boundary (the firm side): the different businesses of the firm are managed by the same manager (or team), and the firm’s organizational structure affects its market value through its impact on the cash flows generated by its various businesses. The other is from outside the boundary (the market side): the stakeholders of the firm cannot selectively invest in and receive cash flow from any particular business disproportionately relative to the others run by the firm. This view takes the cash flows as given and looks into how the firm boundary affects how financial market participants acquire relevant information and value the firm.

These two complementary perspectives create a clear bifurcation of all existing theories of corporate diversification. Most theories take the first view; for example, Maksimovic and Phillips (2002), Stein (1997), Matsusaka and Nanda (2002), Rajan et al. (2000). Only a

Maksimovic and Phillips (2002) introduce a neoclassical model that trades off the diseconomies of total firm size due to scarcity of management talent against diminishing returns to scale in each business, and predicts that an entrepreneur with similar productivities in all his businesses would choose the form of conglomerate, and single-segment firms otherwise. In Stein (1997), the headquarters of a financially constrained conglomerate, who has better knowledge of all its businesses than the external financial market, could create value by shifting more funds to its best businesses. And such winner-picking works better if errors in knowledge of different businesses are more correlated. In Matsusaka and Nanda (2002), the cost of internal funding is lower than that of external funding, hence a firm with an internal capital market enjoys the option of avoiding costly external financing when any individual business lacks funds, but is also subject to a more severe overinvesting agency problem. Rajan et al. (2000) highlight the deficiency of ex post bargaining for profits among segments of a conglomerate when an ex ante division rule cannot be committed to. Ex ante transfer of production risks from a less productive segment to a more productive one on one hand increases
very few take the second view. Krishnaswami and Subramaniam (1999) provide evidence that a spinoff enhances value because it mitigates the information asymmetry in the market about the profitability and operating efficiency of the different divisions of the firm. Hund et al. (2010) suggest that, if multiple segment firms have lower uncertainty about mean profitability than single segment firms, rational learning about mean profitability provides an alternative explanation for the diversification discount that does not rely on suboptimal managerial decisions or a poor firm outlook. Vijh (2002) presents an explanation closest to this paper: the increased difficulty facing shareholders investing in shares of conglomerates in their effort to create efficient asset portfolios compared with investing in single-industry firms. However, these theories generate only a diversification discount, not a diversification premium. This seems to be a drawback. From a normative perspective, to help us understand the pros and cons of corporate diversification theoretically, a theory that can generate both a diversification discount and a premium is a must; And from a positive perspective, the existing empirical literature shows that there are circumstances in which a diversification premium is observed; e.g., as documented by Villalonga (2004).

My model suggests an alternative market-side perspective of corporate diversification. It complements the first category by viewing the firm boundary of a conglomerate from a different angle and embellishes the second category by remedying the aforementioned drawback: the discipline channel introduced in Section 3 could generate a diversification premium, while the trade restriction-channel and the specialization-destruction channel introduced in Section 4 could generate a discount.

6.2 Empirical Predictions

My model also yields two empirical predictions that are consistent with existing evidence in the literature. It seems reasonable to assume that when a company’s different lines of allocative efficiency, but on the other hand intensifies the concern of the more productive segment that more profit has to be shared with its deficient counterpart and reduces production incentive. For more theories that take the first view, see Maksimovic and Phillips (2013).
business are less similar, it is more likely that investors in the financial market have different expertise in acquiring information about each of them. Thus, my model yields two empirical predictions.

First, for a cross-section of diversified firms formed for reasons exogenous to my model, the more similar the businesses that a firm operates, the smaller (greater) the magnitude of diversification discount (premium) should be observed. This is consistent with:

a) Berger and Ofek’s original (1995) paper: "The value loss is smaller when the segments of the diversified firm are in the same two-digit SIC code."

b) Villalonga (2004): Diversified firms in the period 1989-96 (defined by the Business Information Tracking Series (BITS), a database that covers the whole U.S. economy at the establishment level) actually trade at a premium on average. But a subsample of firms that are covered and defined by Compustat as conglomerates, which captures purely unrelated diversification, shows a discount.

c) John and Ofek (1995): For firms that increase their focus by selling assets, the average cumulative excess return to the seller on the two days preceding and on the day of the divestiture announcement is positive and is positively related to different measures of increase in focus.

Second, regarding conglomerate formation, a diversified firm is more likely to be formed across similar businesses. This is consistent with:

a) Comment and Jarrell (1995) who show a steady trend toward greater focus during the 1980s, as measured by a revenue-based Herfindahl index, and this is associated with greater shareholder wealth.

b) Two recent empirical papers by Hoberg and Phillips (2010, 2012) which document that multiple-industry firms are more likely to operate in industries that are similar, measured by overlaps in industry product language, and that mergers and acquisitions are more likely between firms with similar product market language.
7 Conclusion

This paper investigates how a self-interested asset originator can use asset bundling to coordinate the information acquisition of investors with different expertise. Three key economic forces novel in the literature are highlighted in the model. The upside of asset bundling is driven by the discipline channel: asset bundling gives investors less incentive to acquire information about risks that are eventually diversified away. As such, the originator successfully induces investors to learn only about risks that cannot be diversified. Since investors have better knowledge of such risks after learning, they demand a lower risk premium in equilibrium, benefiting the originator. The downside of asset bundling is driven by two different economic forces: Asset bundling mechanically restricts the asset span available to investors, and thus prevents them from holding their respective favorite portfolios. Hence in equilibrium, they demand lower prices to compensate. This is the trade-restriction channel. Asset bundling also induces each investor to specialize less in information acquisition about the risk he has expertise in. Since investors’ expertise is less utilized, more risks are priced in equilibrium. This is the specialization-destruction channel. These forces rationalize the common practice of bundling loans by asset class in securitization, which is at odds with existing theories based on diversification. The analysis also offers an alternative perspective on conglomerate formation (a form of asset bundling), and the relation to empirical evidence in that context is discussed.

There are many other ways that an asset originator can affect how investors with different expertise interact and acquire information; e.g., by designing appropriate auction rules, or choosing what information to reveal to the public, and how. These alternatives would also be interesting to study in future research.

8 References


Peng L. and W. Xiong. 2006. "Investor Attention, Overconfidence and Category Learn-


9 Appendix

9.1 Problems and propositions in the n-risk setup

The derivations in this subsection are in the n-risk setup introduced in Subsection 5.1. The 2-risk setup of the baseline model introduced in Section 2 is a special case, and the results here also apply.

Let \( w_i \equiv \Gamma_i \), loading of total asset payoff \( \sum X_i \) on risk \( f_i \), \( i = 1, 2, \ldots, n \). Thus, \( \sum X_i = \sum_i w_i f_i \).

9.1.1 Price(s) of tradable assets and the originator’s payoff

The posterior of investor \( \alpha \) of type \( i \) about the payoff(s) of tradable assets \( Y = TX \) given his two private signals \( s^{\alpha,i} \) and \( \eta^{\alpha,i} \) is \( N(\mu^{\alpha,i}, (\Omega^{\alpha,i})^{-1}) \), where

\[
\mu^{\alpha,i} = T(\Lambda^{\alpha,i})^{-1}(\Lambda_0^{(i)} s^{\alpha,i} + \Lambda_\eta^{\alpha,i} \eta^{\alpha,i}),
\]

and

\[
\Omega^{\alpha,i} = (T \Gamma(\Lambda^{\alpha,i})^{-1} \Gamma')^{-1}.
\]

Define \( \Omega^a = \int_{\alpha,i} \Omega^{\alpha,i} \), the average precision of all investors’ private signals of \( Y \).

Note that, given the bundling strategy \( T_{m \times n} \) and everyone’s information acquisition choice \( \{\Lambda^{\alpha,i}\} \), the rest of the problem fits in the setup of Admati (1985),\(^1\) which gives the equilibrium prices as a function of asset payoffs \( Y \) and supply from liquidity traders \( \varepsilon \):

\[
p_T = A_T + Y + C_T \varepsilon_T, \quad \text{where}
\]

\[
A_T = -\rho \left[ \frac{1}{\rho^2 \sigma^2} \Omega^a (TT') \Omega^a + \Omega^a \right]^{-1} 1
\]

\[
C_T = \rho (\Omega^a)^{-1}
\]  \( \tag{5} \)

Note that \( (p_T - A_T) \sim N(Y, \Omega_{\varepsilon_T}^{-1}) \), where

\[
\Omega_{\varepsilon_T} = [C_T Var(\varepsilon_T) C_T']^{-1} = \frac{1}{\rho^2 \sigma^2} \Omega^a (TT') \Omega^a.
\]

Thus, we have:

\(^1\) Investors in Admati (1985) have common priors, while we treat priors as though they were private signals.
Proposition 9.1 The originator’s payoff is 
\[ E_0[1' p_T] = E_0[1'(A_T + Y + C_T \varepsilon_T)] = E_0[\sum X_i] + 1'A_T, \]
where \( A_T = -\rho[\Omega_{T,p} + \Omega^a]^{-1} = -\rho[\frac{1}{\tilde{\rho}_i^2} \Omega^a (T'T') \Omega^a + \Omega^a]^{-1}1. \)

9.1.2 Portfolio choice and information acquisition of investors

The following proposition articulates the objective function of an investor:

Proposition 9.2 The information acquisition problem of investor \((\alpha, i)\) is

\[ \max_{\Lambda^{\alpha,i}} Tr[\Omega^{\alpha,i} \Omega_{T,p}^{-1}] + A_T^{\alpha,i} A_T \]
\[ \text{s.t.} \quad (2) \text{ and } (3) \]

Proof: Observing the price(s) \( p_T \), the investor further updates his belief of \( Y \). His posterior becomes \( N(\hat{\mu}^{\alpha,i}, (\hat{\Omega}^{\alpha,i})^{-1}) \), where \( \hat{\mu}^{\alpha,i} = (\hat{\Omega}^{\alpha,i})^{-1}(\Omega^{\alpha,i} \mu^{\alpha,i} + \Omega_{T,p} p_T) \), and \( \hat{\Omega}^{\alpha,i} = \Omega^{\alpha,i} + \Omega_{T,p} \).

From his utility function (1), given his capacity choice, \((\alpha, j)\)'s optimal portfolio choice is

\[ q^{\alpha,j} = \frac{1}{\rho} \hat{\Omega}^{\alpha,j}(\hat{\mu}^{\alpha,j} - p_T) \]

Thus, ex ante, his expected utility is

\[ U = E_0[\frac{1}{2} (\hat{\mu}^{\alpha,j} - p_T)' \hat{\Omega}^{\alpha,j}(\hat{\mu}^{\alpha,j} - p_T)] \]

He knows the distribution of his exogenous signal \( s^{\alpha,i} \) ex ante. Conditional on it, \((\hat{\mu}^{\alpha,i} - p_T)\) is a normal vector, with mean \(-A_T\), and variance \( \Omega_{T,p}^{-1} - (\hat{\Omega}^{\alpha,i})^{-1} \). To derive this variance, note that \( V ar(\hat{\mu}^{\alpha,j}|\Lambda_0^i) = TT(\Lambda_0^i)^{-1}T' - (\hat{\Omega}^{\alpha,i})^{-1}, \)
\( V ar(p_T|\Lambda_0^i) = TT(\Lambda_0^i)^{-1}T' + \Omega_{T,p}^{-1}, \)
and \( cov(p_T, \hat{\mu}^{\alpha,i}|\Lambda_0^i) = TT(\Lambda_0^i)^{-1}T' \).
If a generic random vector \( z \sim N(\mu, \Sigma) \), then \( E[z'z] = \mu'\mu + Tr(\Sigma) \). Hence,

\[
2U = Tr[\hat{\Omega}^{\alpha,i}(\Omega_{T,p}^{-1} - (\hat{\Omega}^{\alpha,i})^{-1})] + A_T'\hat{\Omega}^{\alpha,i}A_T = Tr[\hat{\Omega}^{\alpha,i}\Omega_{T,p}^{-1}] + A_T'\hat{\Omega}^{\alpha,i}A_T - m
\]

The last equality is due to \( \hat{\Omega}^{\alpha,i} = \Omega^{\alpha,i} + \Omega_{T,p} \). Since the third term is exogenous to investor \((\alpha, i)\), this proves the proposition. \( Q.E.D. \)

Now, we can prove Proposition 2.1.

**Proof of Proposition 2.1:** For \( 2 \times m \) bundling strategy \( T_1 \) and \( T_2 \), if they lead to the same asset span, then \( \exists \) a full-rank \( m \times m \) matrix \( M \) such that \( 1'M = 1' \), and that \( T_2 = MT_1 \), then

\[
1'A_{T_2} = 1'[\frac{1}{\rho^2}\int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}(T_2 T_2') \int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1} + \int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}]^{-1} 
\]

\[
= 1'[\frac{1}{\rho^2}\int_{\alpha,i} (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}(MT_1 T_1'M') \int_{\alpha,i} (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1} 
\]

\[
+ \int_{\alpha,i} (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}]^{-1} 
\]

\[
= 1'[\frac{1}{\rho^2}\int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1 T_1') \int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1} + \int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}]^{-1} M'1 
\]

So the originator has the same payoff.

Concerning the investor’s information acquisition problem,

The first term in (6):

\[
Tr[\Omega_{T_2}^{\alpha,i}\Omega_{T_2,p}^{-1}] = Tr\{ (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}[\frac{1}{\rho^2}\int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}(T_2 T_2') \int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}]^{-1} \}
\]

\[
= Tr\{ (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1} 
\]

\[
[\frac{1}{\rho^2}\int_{\alpha,i} (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}(MT_1 T_1'M') \int_{\alpha,i} (MT_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}]^{-1} \}
\]

\[
= Tr\{M'^{-1}(T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}[\frac{1}{\rho^2}\int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1 T_1') \int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}]^{-1} M' \}
\]

\[
= Tr\{M'^{-1}\Omega_{T_1}^{\alpha,i}\Omega_{T_1,p}^{-1} \} = Tr[\Omega_{T_1}^{\alpha,i}\Omega_{T_1,p}^{-1}] 
\]

And because

\[
1'[\frac{1}{\rho^2}\int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}(T_2 T_2') \int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1} + \int_{\alpha,i} (T_2 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}]^{-1} 
\]

\[
= 1'M[\frac{1}{\rho^2}\int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1 T_1') \int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1} + \int_{\alpha,i} (T_1 \Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}] M' 
\]

38
\[ = 1' \left[ \frac{1}{\rho \sigma^2} \int_{\alpha,i} (T_1 \Gamma (\Lambda^{\alpha,i})^{-1} \Gamma' T_1')^{-1} (T_1 T_1') \int_{\alpha,i} (T_1 \Gamma (\Lambda^{\alpha,i})^{-1} \Gamma' T_1')^{-1} + \int_{\alpha,i} (T_1 \Gamma (\Lambda^{\alpha,i})^{-1} \Gamma' T_1')^{-1} \right] M', \]

and \( \Omega^{\alpha,i}_{T_2} = (T_2 \Gamma (\Lambda^{\alpha,i})^{-1} \Gamma' T_2')^{-1} = (MT_1 \Gamma (\Lambda^{\alpha,i})^{-1} \Gamma' M')^{-1} = M^{-1} \Omega^{\alpha,i}_{T_1} M^{-1}, \)

the second term in (6) \( A' T_2 \Omega^{\alpha,i}_{T_2} A T_2 = A' T_1 \Omega^{\alpha,i}_{T_1} A T_1 \)

Therefore, the information acquisition problem is also invariant. This concludes the proof. \( Q.E.D. \)

### 9.1.3 Information acquisition of investors and the originator’s payoff in the subgame induced by \( T = 1' \)

**Proposition 9.3** In the unique subgame equilibrium induced by \( T = 1' \), investor \((\alpha, j)\) minimizes \( \sum_i \{w_i^2(\lambda_i^{\alpha,j})^{-1}\} \), subject to capacity constraint (2) and no-forgetting constraint (3).

Proof: If \( T = 1' \), then all matrices in (6) are scalars, and the investor can only affect \( \Omega^{\alpha,j} \). The objective function is a decreasing function of \( (\Omega^{\alpha,i})^{-1} = 1' \Gamma (\Lambda^{\alpha,j})^{-1} \Gamma 1 = \sum_i \{w_i^2(\lambda_i^{\alpha,j})^{-1}\} \), and thus the investor chooses to minimize it. This is his dominant strategy, so the subgame equilibrium is unique. \( Q.E.D. \)

This is intuitive: when all assets are bundled together, each investor faces a one-dimensional problem: how big a proportion of the whole pool, \( \sum_i w_i f_i \), to hold. Therefore, he minimizes his posterior variance of it, \( \sum_i \{w_i^2(\lambda_i^{\alpha,j})^{-1}\} \), no matter what others are doing.

**Proposition 9.4** In the unique subgame equilibrium induced by \( T = 1' \), the originator’s payoff is

\[
E_0 \left[ \sum X_i \right] - \rho \left[ \frac{n}{\rho \sigma^2} \int_{\alpha,j} \left( \sum_i w_i^2(\lambda_i^{\alpha,j})^{-1} \right)^2 \right] + \int_{\alpha,j} \left( \sum_i w_i^2(\lambda_i^{\alpha,j})^{-1} \right)^{-1} - 1
\]

Proof: By Proposition 9.1, the originator’s payoff is

\[
E_0 \left[ \sum X_i \right] + 1' A_T = E_0 \left[ \sum X_i \right] - \rho \left[ \frac{1}{\rho \sigma^2} \int_{\alpha,j} \left( \sum_i w_i^2(\lambda_i^{\alpha,j})^{-1} \right)^{-1} \right] - 1
\]

\[
= E_0 \left[ \sum X_i \right] - \rho \left[ \frac{1}{\rho \sigma^2} \int_{\alpha,j} \left( \sum_i w_i^2(\lambda_i^{\alpha,j})^{-1} \right)^{-1} \right] - 1
\]

\[
= E_0 \left[ \sum X_i \right] - \rho \left[ \frac{n}{\rho \sigma^2} \int_{\alpha,j} \left( \sum_i w_i^2(\lambda_i^{\alpha,j})^{-1} \right)^{-1} \right] - 1. Q.E.D.
\]
9.1.4 Information acquisition of investors and the originator’s payoff in the subgame induced by $T = I$

**Proposition 9.5** In any equilibrium induced by $T = I$, investor $(\alpha, j)$ specializes in risk $i_0$, where $i_0 \in \arg \max_i \{L_i \lambda_{0,i}^j\}$, where

$$L_i = \left\{ \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2 \right\}^{-1} + \left\{ \rho w_i \lambda_i^\alpha + \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2 \right\}^{-1} \right)^2$$

Proof: If $T = I$, using the fact that $\Gamma$ is an orthogonal matrix, the first term in (6) is

$$Tr\{(\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} \left[ \frac{1}{\rho^2 \sigma^2} \int_{\alpha,j} (\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} \int_{\alpha,j} (\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} \right] \}$$

$$= Tr\{\Gamma \Lambda^{\alpha,j} \Gamma' \left[ \frac{1}{\rho^2 \sigma^2} \int_{\alpha,j} \Lambda^{\alpha,j} \Gamma' \int_{\alpha,j} (\Lambda^{\alpha,j})^{-1} \Gamma' \right] \}$$

$$= Tr\{\Lambda^{\alpha,j} \left[ \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2 \right]^{-1} \} = \sum_{i=1}^{n} \left\{ \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2 \right\}^{-1} \lambda_i^{\alpha,j}.$$  

Since $A_T' = -\rho 1^T \left[ \frac{1}{\rho^2 \sigma^2} \int_{\alpha,j} (\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} \int_{\alpha,j} (\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} \right] 1^T$  

$$= -\rho 1^T \left[ \frac{1}{\rho^2 \sigma^2} (\int_{\alpha,j} \Lambda^{\alpha,j})^2 + \int_{\alpha,j} \Lambda^{\alpha,j}]^{-1} \Gamma' \right],$$  

$$\Omega^{\alpha,j} = (\Gamma (\Lambda^{\alpha,j})^{-1} \Gamma')^{-1} = \Gamma'^{-1} \Lambda^{\alpha,j} \Gamma^{-1} = \Gamma \Lambda^{\alpha,j} \Gamma',$$

the second term in (6) is

$$A_T' \Omega^{\alpha,j} A_T$$

$$= \rho^2 1^T \Gamma \left[ \frac{1}{\rho^2 \sigma^2} (\int_{\alpha,j} \Lambda^{\alpha,j})^2 + \int_{\alpha,j} \Lambda^{\alpha,j}]^{-1} \Lambda^{\alpha,j} \left[ \frac{1}{\rho^2 \sigma^2} (\int_{\alpha,j} \Lambda^{\alpha,j})^2 + \int_{\alpha,j} \Lambda^{\alpha,j}]^{-1} \Gamma' \right]$$

$$= \sum_{i=1}^{n} \left\{ \rho w_i [\lambda_i^\alpha + \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2]^{-1} \right\}^2 \lambda_i^{\alpha,j}$$, since $\Lambda^{\alpha,j}$ is diagonal.

Thus, let $L_i = \left\{ \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2 \right\}^{-1} + \left\{ \rho w_i [\lambda_i^\alpha + \frac{1}{\rho^2 \sigma^2} (\lambda_i^\alpha)^2]^{-1} \right\}^2$. The objective function (6) is

$$\sum_{i=1}^{n} L_i \lambda_i^{\alpha,j} = \sum_{i=1}^{n} L_i \lambda_{0,i}^j y_i,$$

where $y_i \equiv \lambda_i^{\alpha,j} / \lambda_{0,i}^j$, capacity constraint (2) becomes $\prod_{i=1}^{n} y_i \leq K$, and no-forgetting constraint (3) becomes $y_i \geq 1 \forall i$.

This problem maximizes a sum subject to a product constraint. The second order condition for this problem is positive, meaning the optimum is a corner solution. A simple variational argument can show that investor $(\alpha, j)$ would specialize in risk $i_0$, where $i_0 \in \arg \max_i \{L_i \lambda_{0,j}^j\}$. $Q.E.D.$
Proposition 9.6 In a subgame equilibrium induced by $T = I$, the originator’s payoff is

$$E_0[\sum X_i] - \rho \sum_i w_i^2[\lambda_i^a + \frac{1}{\rho^2 \sigma^2}(\lambda_i^a)^2]^{-1}$$

Proof: By Proposition 9.1, the originator’s payoff is

$$E_0[\sum X_i] + 1^T A_r = E_0[\sum X_i] - \rho 1'\left[\frac{1}{\rho^2 \sigma^2}\Omega^a(TT')\Omega^a + \Omega^a\right]^{-1}1$$

$$= E_0[\sum X_i] - \rho 1'\left[\frac{1}{\rho^2 \sigma^2}1 + \int_{\lambda_i} (\Gamma(\Lambda_{i,j})^{-1})^{-1}1\right]^{-1}1$$

$$= E_0[\sum X_i] - \rho 1'\left[\frac{1}{\rho^2 \sigma^2}(\int_{\lambda_i} \Lambda_{i,j}^{-1}) + \int_{\lambda_i} \Lambda_{i,j}^{-1}\right]^{-1}1$$

$$= E_0[\sum X_i] - \rho \sum_i w_i^2[\lambda_i^a + \frac{1}{\rho^2 \sigma^2}(\lambda_i^a)^2]^{-1}. Q.E.D.$$
If there is a positive mass \( d \) of type \( j \) investors who learn about at least two risks, including risk \( i_1 \) and \( i_2 \). Denote the set of these investors \( D \). And \( \int_D \lambda^{a,j}_{i_1} = K_1 \lambda^j_{0,i_1}, \int_D \lambda^{a,j}_{i_2} = K_2 \lambda^j_{0,i_2} \), where \( K_1 > 1, K_2 > 1, \) and \( K_1 K_2 \leq K \).

If instead we let \( d \in (\frac{K_1 - 1}{K_1 K_2 - 1}, \frac{K_1 - 1}{K_1 K_2 - 1}(1 - \frac{1}{K_1})) \) proportion of \((\alpha, j) \) \( \in D \) putting all their capacity previously used on risk \( i_1 \) and \( i_2 \) only on \( i_1 \), and the rest of \((\alpha, j) \) \( \in D \) only on \( i_2 \), the resulting new average precision of private signals of risks are:

\[
\int_D \lambda^{a,j}_{i_1} = [d K_1 K_2 + 1 - d] \lambda^j_{0,i_1} > K_1 \lambda^j_{0,i_1} = \int_D \lambda^{a,j}_{i_1}
\]

\[
\int_D \lambda^{a,j}_{i_2} = [d + (1 - d) K_1 K_2] \lambda^j_{0,i_2} > K_2 \lambda^j_{0,i_2} = \int_D \lambda^{a,j}_{i_2}
\]

\[
\int_D \lambda^{a,j}_i = \int_D \lambda^{a,j}_i \ \forall i \neq i_1, i_2, \text{ and}
\]

\[
\int_{D} \lambda^{a,j}_i = \int_{D} \lambda^{a,j}_i \ \forall i
\]

This strictly improves the originator’s payoff, which is a contradiction.

The first two steps reduce the dimension of the problem from infinity to \( n(i^* - 1) \). In other words, it suffices to focus on \( \{b^j_i : b^j_i \text{ is the proportion of type } j \text{ investors who specialize in risk } i \} \).

**Step 3:** "If \( \exists i_0 \neq j_0, i_0, j_0 \leq i^* \), such that \( b^j_{i_0} > 0 \), then \( b^j_{j_0} = 0 \ \forall k \neq j_0.\"

If \( b^j_{i_0} = d_1 > 0 \) and \( \exists k_0 \) s.t. \( b^k_{j_0} = d_2 > 0 \). Let \( d = \min\{d_1, d_2\} \) Consider a different allocation \( \{\tilde{b}^j_i : \tilde{b}^j_{j_0} = b^j_{j_0} + d, \tilde{b}^j_{i_0} = d_1 - d, \tilde{b}^k_{j_0} = b^k_{j_0} + d, \tilde{b}^k_{i_0} = d_2 - d, \text{ and } \tilde{b}^j_i = b^j_i \text{ otherwise}\} \). We have \( \tilde{\lambda}^a_{j_0} = \lambda^a_{j_0} + \frac{d K_1 (\lambda^j_{0,i_0} - 1)}{\lambda^j_{0,i_0}} \geq \lambda^a_{i_0} \), and \( \tilde{\lambda}^a_i = \lambda^a_i \forall i \neq i_0, j_0 \). This yields a strict improvement of the originator’s payoff, which is a contradiction.

Let \( f(x) = (x + \frac{1}{\rho^2 (2 \rho - 2)^2})^{-1}, x > 0. \) Then \( f'(x) = -\frac{(\frac{2}{\rho^2 (2 \rho - 2)^2} x^3)}{(x + \frac{1}{\rho^2 (2 \rho - 2)^2})^2} < 0, \) and \( f''(x) = \frac{2(\frac{3}{\rho^2 (2 \rho - 2)^2} x^2 + \frac{3}{\rho^2 (2 \rho - 2)^2} x + 1)}{(x + \frac{1}{\rho^2 (2 \rho - 2)^2})^3} > 0. \)

**Step 4:** "All type \( j \leq i^* \) investors specialize in risk \( j \)."

Assume otherwise. Then \( \exists i_0 \leq i^*, \exists j_0 \leq i^* \) s.t. \( i_0 \neq j_0 \) and \( b^j_{i_0} > 0 \). This means \( b^j_{j_0} \leq 1 - b^j_{i_0} < 1. \) By step 3, \( b^k_{j_0} = 0 \ \forall k \neq j_0, \) and \( b^k_{i_0} = 1 \), and thus \( \tilde{\lambda}^a_{j_0} = \frac{\tilde{\lambda}^a_{i_0}}{\lambda^a_{i_0} + \frac{b^j_{0}(K-1)\lambda^j_{0,i_0}}{n}} < \)
Consider a different allocation \( \{ \tilde{b}_i^j : \tilde{b}_{j_0}^i = b_{j_0}^i + b_{i_0}^j, \tilde{b}_{i_0}^i = 0, \text{ and } \tilde{b}_i^j = b_i^j \text{ otherwise} \} \).

\[
\sum_{i=1}^{i^*}[\lambda_i^a + \frac{1}{\rho \sigma^2}(\tilde{\lambda}_i^a)^2]^{-1} - \sum_{i=1}^{i^*}[\lambda_i^a + \frac{1}{\rho \sigma^2}(\lambda_i^a)^2]^{-1} = f[\lambda_{i_0}^a + \frac{b_{i_0}^j}{b_{j_0}^i}(K-1)\tilde{\lambda}_i^a] + f[\lambda_{i_0}^a + \frac{b_{j_0}^i}{b_{i_0}^j}(K-1)\tilde{\lambda}_i^a] - [f(\lambda_{i_0}^a) + f(\lambda_{i_0}^a)] < 0.
\]

The second to last inequality is due to \( f' < 0 \), and the last inequality is due to \( f'' > 0 \).

This means that the proposed alternative allocation strictly improves the originator’s payoff, a contradiction.

**Step 5:** "\( \forall (\alpha, j) \text{ s.t. } j > i^* \), he specializes in one systematic risk \( i \leq i^* \), such that there is an equal mass of investors specializing in each systematic risk."

Suppose \( \exists i_1, i_2 \leq i^* \), s.t. \( \sum_{i=1}^{i_1} \tilde{b}_i^j - \sum_{i=1}^{i_2} \tilde{b}_i^j \equiv 2b > 0 \). By step 4, \( \sum_{i=1}^{i^*} \tilde{b}_i^j = \tilde{b}_{i_1}^j = 1 = \sum_{i=1}^{i^*} \tilde{b}_i^j = \tilde{b}_{i_2}^j \), and thus \( \sum_{i=1}^{i_1} \tilde{b}_i^j - \sum_{i=1}^{i_2} \tilde{b}_i^j = \sum_{i=i_1}^{i_2} \tilde{b}_i^j = \sum_{i=i_1}^{i_2} b_i^j = 2b > 0 \). This implies \( \lambda_{i_1}^a - \lambda_{i_2}^a = \frac{2b(K-1)\tilde{\lambda}_i^a}{\rho \sigma^2} > 0 \).

Consider an alternative allocation \( \{ \tilde{b}_i^j \} \), such that \( \sum_{i=i_1}^{i_2} \tilde{b}_i^j = \sum_{i=i_1}^{i_2} b_i^j - b, \sum_{i=i_1}^{i_2} \tilde{b}_i^j = \sum_{i=i_1}^{i_2} b_i^j + b, \text{ and } \tilde{b}_i^j = b_i^j \text{ otherwise} \). Then

\[
\sum_{i=1}^{i^*}[\tilde{\lambda}_i^a + \frac{1}{\rho \sigma^2}(\tilde{\lambda}_i^a)^2]^{-1} - \sum_{i=1}^{i^*}[\lambda_i^a + \frac{1}{\rho \sigma^2}(\lambda_i^a)^2]^{-1} = f[\lambda_{i_1}^a + \frac{b_{i_1}^j}{b_{i_2}^j}(K-1)\tilde{\lambda}_i^a] + f[\lambda_{i_2}^a + \frac{b_{i_2}^j}{b_{i_1}^j}(K-1)\tilde{\lambda}_i^a] - [f(\lambda_{i_1}^a) + f(\lambda_{i_2}^a)] < 0.
\]

The inequality is again due to \( f'' > 0 \). This means that the proposed alternative allocation strictly improves the originator’s payoff, a contradiction. This concludes the proof of the whole proposition. \( Q.E.D. \)

**Proof of Proposition 5.2:** We first prove that the categorization strategy induces the capacity allocation of the solution to the optimality benchmark.

Note that \( (\Omega^{\alpha,j})^{-1} = TT'(\Lambda^{\alpha,j})^{-1}\Gamma'T' \)
\[
\begin{align*}
&= \left( wI_i^* \ 0_{(n-i^*)i^*} \right) \left( \begin{array}{cccc}
(\lambda_1^{\alpha,j})^{-1} & 0 & \cdots & 0 \\
0 & (\lambda_2^{\alpha,j})^{-1} & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & (\lambda_n^{\alpha,j})^{-1}
\end{array} \right) \left( wI_i^* \ 0_{(n-i^*)i^*} \right) \\
&= w^2 \left( \begin{array}{cccc}
(\lambda_1^{\alpha,j})^{-1} & 0 & \cdots & 0 \\
0 & (\lambda_2^{\alpha,j})^{-1} & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & (\lambda_n^{\alpha,j})^{-1}
\end{array} \right)
\end{align*}
\]

\[TT' = T\Gamma' T' = \left( wI_i^* \ 0_{(n-i^*)i^*} \right) \left( wI_i^* \ 0_{(n-i^*)i^*} \right) = w^2I_i^*,\]

\[
\Omega_{T,p} = \frac{1}{\rho^2\sigma^2} \sum_{a} TT' \Omega^a = \frac{1}{\rho^2\sigma^2} \int_{\alpha,j} \left( T\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma'^{-1} T\Gamma'(\Lambda^{\alpha,j})^{-1}\Gamma'^{-1} \right) TT' \int_{\alpha,j} \left( T\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma'^{-1} T\Gamma'(\Lambda^{\alpha,j})^{-1}\Gamma'^{-1} \right) \]

\[
= \frac{1}{\rho^2\sigma^2} w^{-2} \left( \begin{array}{cccc}
\lambda_1^a & 0 & \cdots & 0 \\
0 & \lambda_2^a & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_i^a
\end{array} \right) w^2 w \left( \begin{array}{cccc}
\lambda_1^a & 0 & \cdots & 0 \\
0 & \lambda_2^a & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_i^a
\end{array} \right) \\
= \frac{1}{\rho^2\sigma^2} w^{-2} \left( \begin{array}{cccc}
(\lambda_1^a)^2 & 0 & \cdots & 0 \\
0 & (\lambda_2^a)^2 & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & (\lambda_i^a)^2
\end{array} \right)
\]

And \( A_T = -\rho[\Omega^a + \Omega_{T,p}]^{-1} \mathbf{1} \)

\[
= -\rho \left( w^{-2} \left( \begin{array}{cccc}
\lambda_1^a & 0 & \cdots & 0 \\
0 & \lambda_2^a & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_i^a
\end{array} \right) + \frac{1}{\rho^2\sigma^2} w^{-2} \left( \begin{array}{cccc}
(\lambda_1^a)^2 & 0 & \cdots & 0 \\
0 & (\lambda_2^a)^2 & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & (\lambda_i^a)^2
\end{array} \right) \right]^{-1} \mathbf{1}
\]

\[
= -\rho w^2 \left( \begin{array}{cccc}
[\lambda_1^a + \frac{1}{\rho^2\sigma^2} (\lambda_1^a)^2]^{-1} & 0 & \cdots & 0 \\
0 & \vdots & \ddots & 0 \\
0 & 0 & \vdots & [\lambda_i^a + \frac{1}{\rho^2\sigma^2} (\lambda_i^a)^2]^{-1}
\end{array} \right) \mathbf{1}
\]

So the first term in the objective function (6) is \( Tr[\Omega^{\alpha,j}\Omega_{T,p}^{-1}] \)
\[
\begin{align*}
&= Tr\{w^{-2} \begin{pmatrix}
\lambda_1^{\alpha,j} & 0 \\
\vdots & \ddots \\
0 & \lambda_i^{\alpha,j}
\end{pmatrix} \times \left[ \frac{1}{\rho^2 \sigma^2} w^{-2} \begin{pmatrix}
\lambda_i^{\alpha,j} & 0 \\
\vdots & \ddots \\
0 & (\lambda_i^{\alpha,j})^2
\end{pmatrix} \right]^{-1} \} \\
&= \sum_{i=1}^{i^*} \left\{ \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2 \right\}^{-1} \lambda_i^{\alpha,j}
\end{align*}
\]

And the second term in (6) is \( A_T^\alpha \Omega^{\alpha,j} A_T \)
\[
= 1'(-\rho w^2) \begin{pmatrix}
[\lambda_1^{\alpha,j} + \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2]^{-1} & 0 \\
0 & \ddots \\
0 & \ddots & \ddots
\end{pmatrix} \\
\times w^{-2} \begin{pmatrix}
\lambda_1^{\alpha,j} & 0 \\
\vdots & \ddots \\
0 & \lambda_i^{\alpha,j}
\end{pmatrix}
\]
\[
\times (-\rho w^2) \begin{pmatrix}
[\lambda_1^{\alpha,j} + \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2]^{-1} & 0 \\
0 & \ddots \\
0 & \ddots & \ddots
\end{pmatrix} 1
\]
\[
= \sum_{i=1}^{i^*} \rho^2 w^2 [\lambda_1^{\alpha,j} + \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2]^{-2} \lambda_i^{\alpha,j}.
\]

Let \( V_i \equiv \{ \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2 \}^{-1} + \rho^2 w^2 [\lambda_1^{\alpha,j} + \frac{1}{\rho^2 \sigma^2} (\lambda_i^{\alpha,j})^2]^{-2}, \ i = 1, 2, \ldots, i^* \), and \( V_j \equiv 0 \) if \( i > i^* \). With an argument analogous to the proof of Proposition 9.4, it can be proved that investor \((\alpha, j)\) devotes all his capacity to a single risk \( i_0 \), where \( i_0 \in \arg \max_{i \leq i^*} \{ V_i \lambda_{i_0,j} \} \).

\( V_k \equiv 0 \ \forall k > i^* \) implies that no investor learns about diversifiable risks. \( V_i = V_j > 0 \ \forall i, j \leq i^* \) implies that \( V_i \lambda_{0,i} = V_i \lambda_{0,j} > V_i \lambda_{1,j} = V_j \lambda_{0,j} > 0 \) \( \forall i, j \leq i^*, i \neq j, \forall k > i^* \), which verifies that every type \( i \leq i^* \) investor would specialize in learning about his own risk \( f_i \). And since \( V_j \lambda_{0,j} = V_i \lambda_{0,j} \ \forall i, j \leq i^* \), every type \( k > i^* \) investor is indifferent between any two non-diversifiable risks. Thus we have verified that a solution to the optimality benchmark is a subgame equilibrium.

Now we show that every type \( j \leq i^* \) investor would specialize in his own risk \( j \) in any subgame equilibrium. Suppose among type \( j_1, j_2, \ldots, j_m \leq i^* \) investors, there are respectively a strictly positive proportion \( b_{j_1}, b_{j_2}, \ldots, b_{j_m} \) who are not learning about their own risks, and
WLOG $b_{j_1} \geq b_{j_2} \geq ... \geq b_{j_m}$. By Proposition 9.4, for each of such risk $j_k, k = 1, 2, ..., m$, there exists a different risk $\tilde{j}_k \neq j_k$ such that $V_{j_k} \bar{\lambda} \leq V_{\tilde{j}_k} \bar{\lambda}$. This implies $V_{j_k} \bar{\lambda} < V_{j_k} \bar{\lambda} < V_{\tilde{j}_k} \bar{\lambda}$, i.e. none of non-type $j_k$ investors would specialize in risk $j_k$. This implies $\Lambda^a_{j_k} \geq \Lambda^a_j \forall j \leq i^*$, and thus $V_{j_1} \geq V_{j_2} \forall j \leq i^*$. Now we have $V_{j_i} \bar{\lambda} > V_{j_k} \bar{\lambda} > V_{\tilde{j}_k} \bar{\lambda}$, a contradiction.

The mass of type $j > i^*$ investors learning about each non-diversifiable risk has to be equal. This is because, the risk $i_0$ that has the strictly least investors specializing in it has $V_{i_0} > V_i \forall i \neq i_0$, attracting all type $j > i^*$ investors to specialize in it in equilibrium, a contradiction.

Lastly, with the expression of $A_T$ derived above in the proof of this proposition, it is straightforward to verify that the originator’s payoff induced by categorization strategy, $E_0[\sum X_i] + 1' A_T$, is identical to that of the optimality benchmark. This concludes the proof. Q.E.D.

### 9.2 Propositions in the 2-risk setup

**Proposition 3.1** is a special case of Proposition 9.4 and 9.6, where $n = 2, w_1^2 = 2$ and $w_2^2 = 0$.

**Proof of Proposition 3.2** is straightforward from Proposition 9.3, as $\sum_i \{w_i^2 (\lambda_i^a)^{-1}\} = 2(\lambda_1^a)^{-1}$

**Proof of Proposition 3.3**: We use Proposition 9.5 to prove this proposition. The result that every investor learns about only one risk directly follows.

If some type 1 investors prefer learning about $f_2$ to $f_1$, then we must have $L_1 \bar{\lambda} \leq L_2 \bar{\lambda}$. This implies $L_1 \bar{\lambda} < L_2 \bar{\lambda}$, which means all type 2 investors strictly prefer to learn about $f_2$. We have $\lambda_1^a < \lambda_2^a$, and thus $L_1 > L_2$ as $w_1 > w_2 = 0$. This implies $L_1 \bar{\lambda} > L_2 \bar{\lambda}$, a contradiction. So all type 1 investors learn only about $f_1$.

Suppose all type 2 investors also learn only about $f_1$. Then $L_2 \bar{\lambda}_0^2 = \{\frac{1}{\rho \sigma^2} (\frac{\bar{\lambda}^2}{2})\}^{-1} \bar{\lambda}$, and $L_1 \bar{\lambda}_0^2 \rightarrow 0$ as $K \rightarrow \infty$. Thus, $\exists K_0 < \infty$ such that a positive proportion of type 2 investors learn about $f_2$ if $K > K_0$. Q.E.D.
Proof of Proposition 4.5: Again by Proposition 9.5, every investor learns about only one risk.

That every investor specializes in his expertise is a subgame equilibrium, since this implies $L_1 = L_2$, and thus $L_1 \bar{\lambda} > L_2 \bar{\lambda}$ and $L_1 \bar{\lambda} < L_2 \bar{\lambda}$, justifying each investor’s choice.

The equilibrium is unique. Otherwise, say WLOG if some type 2 investors learn about $f_1$ in equilibrium, then $L_1 \bar{\lambda} \geq L_2 \bar{\lambda}$, which implies $L_1 \bar{\lambda} > L_2 \bar{\lambda}$; i.e., all type 1 investors learn only about $f_1$, and thus $\lambda_1^a > \lambda_2^a$, and $L_1 < L_2$ since $w_1 = w_2$. This further implies $L_1 \bar{\lambda} < L_2 \bar{\lambda}$, a contradiction. Q.E.D.

Proof of Proposition 4.6: We first derive the equilibrium capacity allocation.

By Proposition 9.3, the optimization problem for investor $(\alpha, j)$ is

$$\min_{\{\lambda_{i,j}\}} \sum_i (\lambda_{i,j})^{-1}$$

s.t. $(\lambda_{i,j})^{-1} \leq (\lambda_{i,j})^{-1}$ and $\prod_i (\lambda_{i,j})^{-1} \geq \frac{1}{K} (\bar{\lambda})^{-1}, \forall i$. The first order condition for this problem is $1 - \frac{v}{(\lambda_{i,j})^{-1}} \prod_i (\lambda_{i,j})^{-1} + z_i = 0$, where $v$ is the Lagrange multiplier on the capacity constraint and $z_i$ is the Lagrange multiplier on the no-forgetting constraint for risk $i$. We guess and verify that if $K$ exceeds a cutoff $K^*$, the no-forgetting constraint does not bind ($z_i = 0\forall i$). This implies $(\lambda_{i,j})^{-1} = \frac{v}{(\bar{\lambda})^{-1}}$. Taking a product on both sides and imposing the capacity constraint again yields $v = (\frac{1}{K} (\bar{\lambda})^{-1})^{-\frac{1}{2}}$, and thus $(\lambda_{i,j})^{-1} = (\frac{1}{K} (\bar{\lambda})^{-1})^{\frac{1}{2}}$ which strictly decreases in $K$, verifying the guess.

The cutoff $K^*$ solves $\bar{\lambda} = \sqrt{K^* \bar{\lambda} \bar{\lambda}}$. So $K^* = \bar{\lambda} \bar{\lambda}$. And the result "If $1 \leq K < \bar{\lambda} \bar{\lambda}$, then $\lambda_{i,j} = \begin{cases} \bar{\lambda}, & \text{if } i = j \\ K \bar{\lambda}, & \text{if } i \neq j \end{cases}, \forall i, \alpha, j"$ follows from the no-forgetting constraint, which states that if $z_i > 0$, then $\lambda_{i,j} = (\lambda_{i,j}^{a,j})^{-1}$. Q.E.D.

Proof of Proposition 4.1:

By Proposition 4.5, in the subgame equilibrium induced by $T = I$, $\lambda_1^a = \lambda_2^a = \frac{K \bar{\lambda} + \lambda}{2}$. So by Proposition 9.6, the originator’s payoff is

$$E_0 \left[ \sum X_i \right] - 2 \rho \left[ \frac{K \bar{\lambda} + \lambda}{2} + 1 + \frac{1}{\rho^2 \sigma^2} \left( \frac{K \bar{\lambda} + \lambda}{2} \right)^2 \right]^{-1} = g \left( \frac{K \bar{\lambda} + \lambda}{2} \right)$$

By Proposition 4.6, in the subgame equilibrium induced by $T = 1'$, if $K \geq \bar{\lambda} \bar{\lambda}$, $(\lambda_{i,j}^{a,j})^{-1} + \frac{\bar{\lambda} \bar{\lambda} - K \bar{\lambda} - \lambda}{2}$
\((\lambda_2^\alpha)^{-1} = 2(\sqrt{K\lambda})^{-1} \forall \alpha, j\). So by Proposition 9.4, the originator’s payoff is

\[ E_0[\sum X_i] - 2\rho[\sqrt{K\lambda} + \frac{1}{\rho^2\sigma^2}(\sqrt{K\lambda})^2]^{-1} = g(\sqrt{K\lambda}) \]

If \(K < \lambda/\bar{\lambda}\), \((\lambda_1^\alpha)^{-1} + (\lambda_2^\alpha)^{-1} = (K\lambda)^{-1} + \bar{\lambda}^{-1} \forall \alpha, j\), so the originator’s payoff is

\[ E_0[\sum X_i] - 2\rho\left[\frac{(K\lambda)^{-1} + \bar{\lambda}^{-1}}{2}\right]^{-1} + \frac{1}{\rho^2\sigma^2}\left[\frac{(K\lambda)^{-1} + \bar{\lambda}^{-1}}{2}\right]^{-2} = g(\frac{(K\lambda)^{-1} + \bar{\lambda}^{-1}}{2}) \]

Q.E.D.

**Proof of Proposition 4.3:** By (7), the expected portfolio holdings of investor \((\alpha, i)\) are

\[ E[\mathbf{q}_T^{\alpha,i}] = \frac{1}{\rho} \hat{\Omega} \mathbf{p}_T - \mathbf{p}_T \]. As discussed in the proof of Proposition 9.2, \(E[\hat{\mu}^{\alpha,i} - \mathbf{p}_T] = -\hat{\lambda} = \rho\left[\frac{1}{\rho\sigma^2} \Omega^a(T'\Omega^a + \Omega^a)^{-1}\right] \mathbf{1}^\prime\), and \(\hat{\Omega}^{\alpha,i} = \Omega^{\alpha,i} + \frac{1}{\rho\sigma^2} \Omega^a(T'\Omega^a)\Omega^a\).

If \(T = \mathbf{1}\), by orthogonality of \(\Gamma\), \(\Omega^{\alpha,i} = [\Gamma(\Lambda_0^{-1})\Gamma]^{-1} = \Gamma\Lambda_0^{-1}\Gamma\). \(\Omega^a = \int_{\alpha,i} \Omega^{\alpha,i} = \Gamma\Lambda_0^{-1}\Gamma\),

where \(\Lambda_0 = \int_{\alpha,i} \Lambda_0 = \hat{\lambda} + \hat{\lambda}/2\mathbf{1}. \frac{1}{\rho\sigma^2} \Omega^a(T'\Omega^a)\Omega^a = \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1}.

Hence, \(E[\mathbf{q}_T^{\alpha,i}] = \Gamma(\Lambda_0^{-1} + \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1}^{-1} \Gamma^{-1}\Gamma\mathbf{1}\)

\[ = \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1}^{-1} \Gamma(\Lambda_0^{-1} + \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1} \mathbf{1} \]

The last equality is due to \(\Gamma\mathbf{1} = \mathbf{1}\), because \(w_1 = w_2 = 1\).

A type \(i\) investor’s expected risk holding is \(\Gamma' E[\mathbf{q}_T^{\alpha,i}] = \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1}^{-1} \Gamma(\Lambda_0^{-1} + \frac{1}{\rho\sigma^2} \hat{\lambda} + \hat{\lambda}/2)^2\mathbf{1} \).

Since \(\Lambda_0^1 = \text{diag}(\hat{\lambda}, \hat{\lambda})\) and \(\Lambda_0^2 = \text{diag}(\lambda, \lambda)\), this proves the first statement.

If \(T = \mathbf{1}', \forall (\alpha, i), \Omega^{\alpha,i} = (\hat{\lambda} + \hat{\lambda})^{-1} = \Omega^a\). Thus \(E[\mathbf{q}_T^{\alpha,i}] = 1 \forall (\alpha, i)\). Since 1 unit of the tradable asset contains 1 unit of each risk, this proves the second statement. Q.E.D.

**Proof of Proposition 4.4** is straightforward from Proposition 4.1.

**Proof of Proposition 5.3:** It is shown in the proof of Proposition 9.2 that \(2U^{\alpha,i} =\)

\( \text{Tr} [\Omega^{\alpha,i} \Omega_{T,p}^{-1}] + A_T \Omega^{\alpha,i} A_T + A_T \Omega_{T,p} A_T \)

Thus,

\(2U^a \equiv \int_{\alpha,i} 2U^{\alpha,i} = \int_{\alpha,i} \text{Tr} [\Omega^{\alpha,i} \Omega_{T,p}^{-1}] + \int_{\alpha,i} (A_T \Omega^{\alpha,i} A_T + A_T \Omega_{T,p} A_T)\)

\(= \text{Tr} [\Omega^a \Omega_{T,p}^{-1}] + A_T (\Omega^a + \Omega_{T,p}) A_T\)

\(= \text{Tr} [\rho^2 \sigma^2 (T'T)^{-1} (\Omega^a)^{-1}] + \rho^2 \mathbf{1}' (\Omega^a + \Omega_{T,p}) \mathbf{1}\)

Note that the second term= \(-\rho \mathbf{1}' A_T\). We know from Proposition 9.1 that \(\mathbf{1}' A_T\) is the
greatest for the originator’s favored bundling strategy. Hence, the second term is smaller for $T = 1'$ than for $T = I$. This comparison corresponds to the first reason in the explanation of Proposition 5.3 in the text.

The liquidity trader’s expected loss to the investors is $E_0[\varepsilon_T'(p_T - Y)] = E_0[\varepsilon_T'(A_T + C_T \varepsilon_T)] = E_0[\varepsilon_T' C_T \varepsilon_T] = Tr[\rho \sigma^2 (TT')^{-1} (\Omega^a)^{-1}]$, which is proportional to the first term of $2U^a$.

If $w_1^2 = 2$ and $w_2^2 = 0$, $Tr[\rho \sigma^2 (TT')^{-1} (\Omega^a)^{-1}] = \begin{cases} \rho^2 \sigma^2 [(\lambda_{1,T=I}^a)^{-1} + (\lambda_{1,T=I}^a)^{-1}], & \text{if } T = I \\ \rho^2 \sigma^2 (\lambda_{1,T=1'}^a)^{-1}, & \text{if } T = 1' \end{cases}$.

By Proposition 3.3, $\lambda_{1,T=1'}^a \geq \lambda_{1,T=I}^a$. Thus, $\rho^2 \sigma^2 (\lambda_{1,T=1'}^a)^{-1} \leq \rho^2 \sigma^2 (\lambda_{1,T=I}^a)^{-1} < \rho^2 \sigma^2 (\lambda_{1,T=I}^a)^{-1} + (\lambda_{1,T=1'}^a)^{-1}$.

The first inequality corresponds to the second reason in the explanation of Proposition 5.3 in the text, and the second inequality to the third reason. Q.E.D.

### 9.3 Extrapolation between the two polar cases

In Sections 3 and 4, two polar cases were used to show the upside and downside of asset bundling. Recall from Section 2 that $\sum X_i = w_1 f_1 + w_2 f_2$, where $w_1 \geq w_2 \geq 0$ are loadings of total payoff of assets on each risk. The polar case in Section 3 corresponds to the case of extreme asymmetry in the contributions of the two risks to the total payoff of the assets: $w_2/w_1 = 0$, while the polar case in Section 4 corresponds to the case of extreme symmetry: $w_2/w_1 = 1$. Can we say something about the cases in between: $0 < w_2/w_1 < 1$?

The following proposition shows that, with an increase in the contribution of $f_2$ relative to that of $f_1$, $w_2/w_1$, the originator’s payoff in the subgame induced by $T = 1'$; i.e., pooling the assets, changes continuously. So does his payoff in the subgame induced by $T = I$.

**Proposition 9.7** In the baseline model, the originator’s payoffs in the subgames induced by $T = 1'$ and $T = I$ both change continuously with $w_2/w_1$.

This proposition establishes the continuity of changes in payoffs with respect to changes in risk loadings. This implies that the results in Section 3 and 4 are robust to a small
perturbation of risk loadings. We have established in Section 3 that, if risk \( f_2 \) is diversifiable, the originator is strictly better off pooling the assets than selling them as they are when investors have enough capacity to acquire information. The continuity established here implies that the same conclusion holds as long as the contribution of \( f_2 \) is sufficiently small. Similarly, we know from Proposition 4.1 that the originator is strictly worse off pooling the assets when the two risks are completely symmetric; i.e., \( w_1 = w_2 \). And the continuity established here implies that the same conclusion holds as long as the contributions of the two risks are sufficiently close.

**Proof of Proposition 9.7**: First, recall that \( w_1^2 = 2 - w_2^2 \), which monotonically decreases with \( w_2 \). So \( w_2 \) increases monotonically with \( w_2/w_1 \) in the range we consider: \( 0 \leq w_2/w_1 \leq 1 \) and \( w_1 > 1 \). Thus, it suffices to focus on change in \( w_2 \).

If \( T = 1' \), to show that the originator’s payoff changes continuously with \( w_2 \), by Proposition 9.4, it suffices to show that \([w_1^2(\lambda_1^{\alpha,i})^{-1} + w_2^2(\lambda_2^{\alpha,i})^{-1}]\) changes continuously with \( w_2 \) for each type.

With an argument analogous to the proof of Proposition 4.3, one can show that each investor tries his best to equalize \( w_1^2(\lambda_1^{\alpha,i})^{-1} \) and \( w_2^2(\lambda_2^{\alpha,i})^{-1} \) when acquiring information.

For a type 1 investor, if \( w_1^2(\lambda_1^{\alpha,i})^{-1} \geq w_2^2(\lambda_2^{\alpha,i})^{-1} \), he goes for \( f_1 \) first before learning about \( f_2 \) : if \( w_2^2(K\lambda) - 1 > w_2^2(\lambda_1^{\alpha,i})^{-1} \), then \( \lambda_1^{\alpha,i} = K\lambda, \lambda_2^{\alpha,i} = \lambda \);

if \( w_2^2(K\lambda) - 1 \leq w_2^2(\lambda_1^{\alpha,i})^{-1} \), then \( w_2^2(\lambda_1^{\alpha,i})^{-1} = w_2^2(\lambda_2^{\alpha,i})^{-1} = w_1 w_2 \sqrt{(K\lambda\lambda)^{-1}} \).

If \( w_1^2(\lambda_1^{\alpha,i})^{-1} \leq w_2^2(\lambda_2^{\alpha,i})^{-1} \), he goes for \( f_2 \) first before learning about \( f_1 \) : if \( w_2^2(\lambda_1^{\alpha,i})^{-1} < w_2^2(K\lambda)^{-1} \), then \( \lambda_1^{\alpha,i} = \lambda, \lambda_2^{\alpha,i} = K\lambda \);

if \( w_2^2(\lambda_1^{\alpha,i})^{-1} \geq w_2^2(K\lambda)^{-1} \), then \( w_1^2(\lambda_1^{\alpha,i})^{-1} = w_2^2(\lambda_2^{\alpha,i})^{-1} = w_1 w_2 \sqrt{(K\lambda\lambda)^{-1}} \).

Therefore, \( w_1^2(\lambda_1^{\alpha,i})^{-1} + w_2^2(\lambda_2^{\alpha,i})^{-1} = \begin{cases} w_2^2(K\lambda)^{-1} + w_2^2(\lambda_1^{\alpha,i})^{-1}, & \text{if } \frac{w_2^2}{w_1^2} < \frac{1}{K\lambda} \\ 2w_1 w_2 \sqrt{(K\lambda\lambda)^{-1}}, & \text{if } \frac{1}{K\lambda} \leq \frac{w_2^2}{w_1^2} \leq \frac{K\lambda}{\lambda} \\ w_1^2(\lambda_1^{\alpha,i})^{-1} + w_2^2(\lambda_2^{\alpha,i})^{-1}, & \text{if } \frac{w_2^2}{w_1^2} > \frac{K\lambda}{\lambda} \end{cases} \).

Since \( \frac{w_2^2}{w_1^2} \) strictly increases with \( w_2 \) in the range of interest, \( w_1^2(\lambda_1^{\alpha,i})^{-1} + w_2^2(\lambda_2^{\alpha,i})^{-1} \) is continuous in \( w_2 \) in each segment, and its value coincides at the two thresholds.
For a type 2 investor, since \( w_1^2 \lambda^{-1} > w_2^2 \lambda^{-1} \), he always goes for \( f_1 \) first before learning about \( f_2 \):

if \( w_1^2(K\lambda)^{-1} > w_2^2 \lambda^{-1} \), then \( \lambda_1^{a,2} = K\lambda, \lambda_2^{a,2} = \bar{\lambda} \);

if \( w_1^2(K\lambda)^{-1} \leq w_2^2 \lambda^{-1} \), then \( w_1^2(\lambda_1^{a,2})^{-1} = w_2^2(\lambda_2^{a,2})^{-1} = w_1 w_2 \sqrt{(K\lambda\lambda)^{-1}} \).

Therefore, \( w_1^2(\lambda_1^{a,1})^{-1} + w_2^2(\lambda_2^{a,1})^{-1} = \left\{ \begin{array}{ll}
  w_1^2(K\lambda)^{-1} + w_2^2 \lambda^{-1}, & \text{if } \frac{w_2}{w_1} < \frac{1}{K\lambda} \\
  2w_1 w_2 \sqrt{(K\lambda\lambda)^{-1}}, & \text{if } \frac{w_2}{w_1} \geq \frac{1}{K\lambda}
\end{array} \right. \), which is again continuous in \( w_2 \) in each segment, and its value coincides at the threshold.

Thus, we have proved that if \( T = 1' \), then the originator’s payoff changes continuously with \( w_2 \).

Now we prove that the same conclusion holds if if \( T = I \).

As in Proposition 9.5, let \( L_i = \{ \frac{1}{\rho \sigma^2} (\lambda_i^a)^2 \}^{-1} + \{ \rho w_i (\lambda_i^0 + \frac{1}{\rho \sigma^2} (\lambda_i^0)^2) \}^{-1}, i = 1, 2 \).

Since \( w_1 \geq w_2 \geq 0 \), by an argument analogous to the proof of Proposition 4.3, all type 1 investors must learn only about \( f_1 \) in equilibrium.

Each type 2 investor learns about only \( f_1 \) or only \( f_2 \). Let \( b \in [0, 1] \) denote the proportion of those who learn about \( f_2 \). It suffices to show that \( b \) changes continuously with \( w_2 \).

Then \( \lambda_1^a = \frac{K(\lambda+\bar{\lambda})}{2} - b \frac{(K-1)\lambda}{2}, \) which strictly decreases in \( b \), and \( \lambda_2^a = \frac{\lambda+\bar{\lambda}}{2} + b \frac{(K-1)\lambda}{2}, \) which strictly increases in \( b \).

\( L_1, L_2 \) can be viewed as functions of \( b, w_2 \) and \( K \), \( L_1(b, w_2, K) \) strictly increases in \( b \), decreases in \( w_2 \) and \( K \). \( L_2(b, w_2, K) \) strictly decreases in \( b \), increases in \( w_2 \), and decreases in \( K \).

\( \forall K \geq 1 \), it is easy to verify that \( \Delta L_1(1,0,K) > \bar{\lambda} L_2(1,0,K), \) and \( \Delta L_1(1,1,K) < \bar{\lambda} L_2(1,1,K). \) Hence \( \exists 0 < \bar{w} < 1 \), such that \( \Delta L_1(1,\bar{w},K) = \bar{\lambda} L_1(1,\bar{w},K). \)

If \( w_2 > \bar{w}, \Delta L_1(1,w_2,K) < \bar{\lambda} L_2(1,w_2,K) \), by Proposition 9.5, each type 2 investor strictly prefers to learn about \( f_2 \), thus \( b = 1 \).

It is easy to see \( \forall K \geq 1, \Delta L_1(0,1,K) < \bar{\lambda} L_2(0,1,K). \)

Since \( \Delta L_1(0,0,\infty) < \bar{\lambda} L_2(0,0,\infty) \) and \( \Delta L_1(0,0,1) > \bar{\lambda} L_2(0,0,1), \) \( \exists \bar{K} \in (1,\infty) \) s.t. \( \Delta L_1(0,0,\bar{K}) = \bar{\lambda} L_2(0,0,\bar{K}). \)
If $K < \bar{K}$, we have $\Delta L_1(0, 0, K) > \bar{\lambda}L_2(0, 0, K)$ and $\Delta L_1(0, 1, K) < \bar{\lambda}L_2(0, 1, K)$. Hence $\exists 0 < w < 1$, such that $\Delta L_1(0, w, K) = \bar{\lambda}L_1(0, w, K)$.

Thus, if $K < \bar{K}$ and $w_2 < w$, $\Delta L_1(0, w_2, K) > \bar{\lambda}L_2(0, w_2, K)$, by Proposition 9.5, each type 2 investor strictly prefers to learn about $f_1$, thus $b = 0$.

If $K < \bar{K}$ and $\bar{w} \geq w_2 \geq w$ or if $K \geq \bar{K}$ and $w_2 \leq \bar{w}$, we have $\Delta L_1(0, w_2, K) \leq \bar{\lambda}L_2(0, w_2, K)$ and $\Delta L_1(1, w_2, K) \geq \bar{\lambda}L_2(1, w_2, K)$. Hence $\exists \bar{b}(w_2, K) \in [0, 1]$ s.t. $\Delta L_1(\bar{b}, w_2, K) = \bar{\lambda}L_2(\bar{b}, w_2, K)$. By Proposition 9.5, each type 2 investor is indifferent between learning about $f_1$ and $f_2$, thus a proportion $\bar{b}(w_2, K)$ of them learning about $f_2$ constitutes an equilibrium. $\bar{b}(w_2, K)$ increases continuously in $w_2$, $\bar{b}(\bar{w}, K) = 1$, and $\bar{b}(w, K) = 0$ if $K < \bar{K}$.

To summarize, if $K \geq \bar{K}$, the equilibrium $b = \begin{cases} 
\bar{b}(w_2, K), & \text{if } w_2 \leq \bar{w} \\
1, & \text{if } w_2 > \bar{w}
\end{cases}$. If $K < \bar{K}$, the equilibrium $b = \begin{cases} 
0, & \text{if } w_2 < w \\
\bar{b}(w_2, K), & \text{if } w \leq w_2 \leq \bar{w} \\
1, & \text{if } w_2 > \bar{w}
\end{cases}$.

In both cases, $b$ varies continuously with $w_2$. This concludes the proof. $Q.E.D.$