Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 175 (2018) 447–490

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Asset bundling and information acquisition of investors with different expertise

Liang Dai [1]

Received 5 November 2016; final version received 4 February 2018; accepted 5 February 2018
Available online 9 February 2018

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.

JEL classification: D83; D82; G11; G21; G32

Keywords: Asset bundling; Endogenous information acquisition; Heterogeneous learning expertise; Securitization; Conglomerate formation

E-mail address: liangdai@sjtu.edu.cn.

[1] Antai College of Economics and Management, Shanghai Jiao Tong 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. Comments from the editor, Laura Veldkamp, and two anonymous referees have greatly improved the paper. I also thank Yusheng Zhang for his excellent research assistance. All remaining errors are mine.

https://doi.org/10.1016/j.jet.2018.02.003
0022-0531/

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[2] 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 cost advantages in information acquisition of different buyers are thus better utilized and result in a lower total risk premium required, benefiting the originator.[3]

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.

[2] 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.

[3] Parlour and Plantin (2008) point out that "Interestingly, however, the secondary loan market does not seem to apply this rule of maximal diversification in practice. CLOs are often backed by fairly restricted pools, and are commonly specialized by country and/or industry. This presents a puzzle: the market for individual loans is rapidly growing, which contradicts the principle of maximal pooling. This suggests that diversification comes at a cost: potential investors may have different degree of expertise in different asset pools. In this case, selling underdiversified claims may increase the participation of sophisticated investors for industries or names for which they have expertise."

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 endogenous 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 investors. 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 5.4. 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 discusses the general case and several issues of the baseline model, and introduces a generalization of the baseline model that establishes the optimality of categorization strategy. 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

This section introduces the setup of the baseline model in chronological order. Section 2.1 to 2.5 introduces risks and asset payoffs, the originator, investors, the liquidity trader and the equilibrium concept, respectively. Section 2.6 discusses the modeling approach. Section 2.7 summarizes the setup.

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, $\mathbf{X} = \Gamma \mathbf{f}$, such that $\Gamma = (\gamma_{ij})$ is an orthogonal matrix.[4] $w_i \equiv \gamma_{1i} + \gamma_{2i}$, $i = 1, 2$ is the loading of total

[4] One can always make Γ orthogonal by redefining risks $\mathbf{f}$ through the eigenvalue decomposition of $Var(X)$.

asset payoff $(X_1 + X_2)$ on f_i. The orthogonality of Γ 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.

2.2. Originator

There is a risk-neutral originator,[5] 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.[6] This means that 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, $\mathbf{1}'T = \mathbf{1}'$, and payoff(s) of the tradable asset(s) $\mathbf{Y} = T\mathbf{X}$. The originator's problem can be expressed as $\max_T E_0[\mathbf{1}'\mathbf{p}_T]$, where $\mathbf{p}_T$ 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 $\mathbf{f}$, and does two things sequentially after observing the bundling choice of the originator: 1) acquires information about the risks $\mathbf{f}$ to maximize his expected utility at the trading stage; 2) chooses a portfolio of tradable assets $\mathbf{q}$ to maximize his mean-variance utility:

$$\max_{\mathbf{q}} E[\rho\mathbf{q}'(\mathbf{Y} - \mathbf{p}) - \frac{\rho^2}{2}\mathbf{q}'Var(\mathbf{Y})\mathbf{q}]. \tag{1}$$

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 $\mathbf{f}$. One signal has exogenous precision, and the investor is to choose the precision of the other. Conditional on $\mathbf{f}$, signals are independent across investors.

The exogenous signal models the different expertise of investors. Specifically, investor α of type i's (hereafter (α, i)) exogenous signal $\mathbf{s}^{\alpha,i} \sim N(\mathbf{f}, (\Lambda_0^i)^{-1})$,[7] where $\Lambda_0^i = diag(\lambda_{0,1}^i, \lambda_{0,2}^i)$. It is assumed that $\lambda_{0,i}^i = \bar{\lambda} > \underline{\lambda} = \lambda_{0,-i}^i > 0$, 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 $\boldsymbol{\eta}^{\alpha,i} \sim N(\mathbf{f}, (\Lambda_\eta^{\alpha,i})^{-1})$ models investors' information acquisition. To highlight the role of expertise, and also for tractability of demand aggregation, it is assumed that

5 Section 2.6.3 discusses the purpose of this risk neutrality assumption.

6 Unlike the security-design literature, here it is assumed that the originator cannot retain any asset. This assumption can be motivated by an exogenous cost of carry faced by the originator, e.g., due to liquidity needs or hedging demand against risks outside the model. The author thanks an anonymous referee for pointing this out.

7 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.

its precision matrix $\Lambda_\eta^{\alpha,i}$ is diagonal, as in Van Nieuwerburgh and 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.[8]

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 (α, 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, where $K \geq 1$[9]:

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

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

$$\lambda_j^{\alpha,i} \geq \lambda_{0,j}^i \ \forall j. \tag{3}$$

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. 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) $\mathbf{p}$ of the tradable assets. In equilibrium, the price(s) $\mathbf{p}$ serves as an additional endogenous signal of the payoff(s) of these assets $\mathbf{Y}$. The investor updates his belief about $\mathbf{Y}$ using Bayes Law and decides how much of each asset to buy, $\mathbf{q}^{\alpha,i}$, to maximize his utility (equation (1)). The technical details of the pricing formula and of investors' portfolio choice are given in the appendix.

2.4. 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 $\mathbf{f}$ is $\boldsymbol{\varepsilon}_f \sim N(0, \sigma^2\mathbf{I})$.[10] This implies that the liquidity trader's demand for original assets $\mathbf{X}$ is $\boldsymbol{\varepsilon} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\varepsilon}_f \sim N(0, \sigma^2\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}'^{-1}) = N(0, \sigma^2\mathbf{I})$, with the last equality due to the orthogonality of Γ.

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 $\boldsymbol{\varepsilon}_T$ for tradable

[8] If signals correlated with more than one factor are allowed, aggregation of individual demand, and subsequently each investor's learning choice problem becomes intractable, and the result hard to interpret even if numerically obtainable.

[9] This comes from $\det[(\Lambda_0^i)^{-1}]/\det[(\Lambda^{\alpha,i})^{-1}] \leq K$.

[10] The liquidity demand here can also be interpreted as hedging demand due to exposure $-\boldsymbol{\varepsilon}_f$ to risks $\mathbf{f}$ outside the model.

asset(s) $\mathbf{Y} = T\mathbf{X}$ is assumed to be the linear projection of her desired portfolio $\boldsymbol{\varepsilon}'\mathbf{X}$ onto the tradable assets span: $\boldsymbol{\varepsilon}_T = (TT')^{-1}T\boldsymbol{\varepsilon} \sim N(0, \sigma^2(TT')^{-1})$.

2.5. Equilibrium

We say $\{T, \{\Lambda^{\alpha,i}\}, \{\mathbf{q}^{\alpha,i}\}, \mathbf{p}_T\}$ is an *equilibrium* iff:

1) The bundling strategy T maximizes the originator's payoff $E_0[\mathbf{1}'\mathbf{p}_T]$;

2) Given the originator's bundling strategy T and the distribution of his exogenous signal $\mathbf{s}^{\alpha,i}$, each investor (α, i)'s choice of information acquisition $\Lambda^{\alpha,j}$ and porfolio choice $\mathbf{q}^{\alpha,i}$ 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 $\{\mathbf{q}^{\alpha,i}\}$, prices $\mathbf{p}_T$ clear the market: $\int_{\alpha,i} \mathbf{q}^{\alpha,i} + \boldsymbol{\varepsilon}_T = \mathbf{1}$; and

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

We say $\{\{\Lambda^{\alpha,i}\}, \{\mathbf{q}^{\alpha,i}\}, \mathbf{p}_T\}$ 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) $\mathbf{p}_T$ are linear functions of payoff(s) $\mathbf{Y}$ and liquidity trader's demand $\boldsymbol{\varepsilon}_T$.

2.6. Discussion on modeling approach

This paper aims at studying a benchmark in which the originator faces endogenous adverse selection problem and wants to resolve it as much as possible through his bundling choice. Adverse selection stems from asymmetric initial information of different investors and their subsequent endogenous information acquisition, and results in different price discount. The benchmark is meaningful only if investors with different initial information would engage in different learning choices if assets are not bundled. The setup featuring specialized learning (instead of generalized learning, in the language of Van Nieuwerburgh and Veldkamp (2010)) is a way to get those choices to differ, and the resulting asymmetric information acquisition is realistic. To set up the main tension in the benchmark, the model adopts one that introduces a combination of investors' preference and information technology that makes it optimal for each investor to engage in specialized learning.[11] In addition, the assumed risk neutrality of the originator largely simplifies the analysis and helps to focus on the main tension. This subsection discusses the details of the modeling approach.

2.6.1. The role of investors' 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. This is the impetus for specialization in information acquisition in the model.

Specifically, an investor's utility function can be expressed as $U = E_1[u_1(E_2[u_2(W)])]$, where $W = W_0 + \mathbf{q}'(\mathbf{Y} - \mathbf{p})$ 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.

[11] Readers interested in how different combinations of preferences and information technologies lead to specialized or generalized learning are referred to Van Nieuwerburgh and Veldkamp (2010).

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.

2.6.2. *Increasing return to scale in information acquisition*

While investors' preference induces specialized learning, it does not provide guidance on *which* risk an investor should focus his attention on. Learning technology fills this gap. Together with initial information advantage (i.e., $\bar{\lambda} > \underline{\lambda}$), the learning technology incentivizes investors to specialize in their respective expertise.

Specifically, 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 increasing return to scale. That is, $K\lambda - \lambda$, the gain of signal precision by using a given capacity K, increases with prior knowledge λ. This turns initial information advantage ($\bar{\lambda} > \underline{\lambda}$) into a cost 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 a 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 cost 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 particular 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 cost advantage in acquiring new information.

2.6.3. *Risk neutrality of the originator*

This paper studies how the originator can mitigate endogenous adverse selection through his bundling choice. Together with risk aversion of investors, risk neutrality of the originator is assumed as a convenient way to set up the desired adverse selection. In addition, since the fundamentals of these assets $E_0[\sum Y_i] = E_0[\sum X_i]$ are exogenous, the risk neutrality assumption also perfectly aligns the preference of the originator (who wants to maximize the total price of the assets) with that of a market maker who wants to maximize the market liquidity of the assets in the sense of minimizing the expected total price discount, $E_0[\sum(Y_i - P_i)]$. If the originator is instead modeled as risk averse, besides technical complication, there would be additional effects distracting readers from the main points analyzed in the paper.

2.7. Summary of model setup

The following timeline summarizes the model setup:

Timeline	Originator	Investors	Liquidity trader
0	Chooses bundling strategy		
1		Decide which information to acquire	
2	Sells assets	Observe signals and choose portfolio	Demands assets
3	Consumes payoff	Consume payoff	

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) $T = \mathbf{I}$, selling the original assets as they are, and ii) $T = \mathbf{1}'$, 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 $\lambda^a_{j,T} \equiv \int_{\alpha,i} \lambda^{\alpha,i}_{j,T}$, market average signal precision of risk f_j induced by bundling strategy T. And $\Lambda^a_T \equiv diag(\lambda^a_{1,T}, \lambda^a_{2,T})$. Subscript T is suppressed if no confusion is caused.

3. The upside of bundling: the discipline channel

This section discusses the upside of asset bundling — the discipline channel. That is, asset bundling mechanically washes out idiosyncratic risks and disincentivizes investors to learn about them. As such, investors' learning capacity is channeled to risks that cannot be diversified, and subsequently lowers the risk premium investors demand for holding them.

To illustrate the discipline channel, this section discusses the polar case in which total payoff of the assets for sale depends only on a single risk: $X_1 + X_2 = \sqrt{2} f_1$. This can be motivated by 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. It is shown that in this case, the originator is better off pooling the assets. Later in Section 5.1.1, it is shown that qualitatively similar results hold as long as the contribution of risk f_2 is sufficiently low.

As discussed in Proposition 2.1, we need to compare only two bundling strategies: $T = \mathbf{I}$, selling the original assets as they are, and $T = \mathbf{1}'$, 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 $T = \mathbf{I}$ and $T = \mathbf{1}'$, the originator's payoff is $E_0[\sum X_i] - 2\rho[\frac{1}{\rho^2\sigma^2}(\lambda^a_{1,T})^2 + \lambda^a_{1,T}]^{-1}$.*

The originator's payoff depends only on $\lambda^a_{1,T}$ but not on $\lambda^a_{2,T}$, since his net supply of f_2 is zero. And his payoff is a strictly increasing function of $\lambda^a_{1,T}$, 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 $T = \mathbf{I}$ and $T = \mathbf{1}'$ 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.

Given Proposition 3.1, 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 *(Discipline Channel). If $X_1 + X_2 = \sqrt{2} f_1$, in the unique subgame equilibrium induced by $T = \mathbf{1}'$, 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, $\lambda^a_{1,T}$, the market average signal precision of f_1, reaches the greatest possible level. As a result, bundling strategy $T = \mathbf{I}$, 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 $T = \mathbf{I}$, 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 Γ, 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 2.6.1, 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 2.6.2, 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 cost advantage in learning about it, as discussed in Section 2.6.2. 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 $K \to \infty$), and at the same time, the investor's cost 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. Yet, 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 $T = \mathbf{1}'$, 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

This section discusses the downside of asset bundling. First, pooling the assets 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 the 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 the specialization-destruction channel.

To illustrate these two channels, 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, say mortgages and credit cards, that make similar contributions to the total value of the assets. Again by Proposition 2.1, without loss of generality we consider only two bundling strategies, $T = \mathbf{I}$, selling the original assets as they are, and $T = \mathbf{1}'$, 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. This section shows that bundling the assets is strictly inferior in this case. Later, it is shown in Section 5.1.1 that qualitatively similar results hold as long as the contributions of the two risks are sufficiently close.

We first establish that the originator is strictly worse off pooling the assets. Proposition 4.1 states the originator's payoffs from the two bundling strategies, respectively.

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

1) The originator's payoff from choosing $T = \mathbf{I}$ *is* $g(\frac{K\bar{\lambda}+\underline{\lambda}}{2})$;

2) If $K \geq \bar{\lambda}/\underline{\lambda}$, *the originator's payoff from choosing* $T = \mathbf{1}'$ *is* $g(\sqrt{K\bar{\lambda}\underline{\lambda}})$;

3) If $K < \bar{\lambda}/\underline{\lambda}$, *the originator's payoff from choosing* $T = \mathbf{1}'$ *is* $g[(\frac{(K\underline{\lambda})^{-1}+\bar{\lambda}^{-1}}{2})^{-1}]$.

It is easy to see that g is a strictly increasing function. And by the inequality of arithmetic and geometric means, $0 < (\frac{(K\underline{\lambda})^{-1}+\bar{\lambda}^{-1}}{2})^{-1} \leq \sqrt{K\bar{\lambda}\underline{\lambda}} < \frac{K\bar{\lambda}+\underline{\lambda}}{2}$. Therefore, pooling the assets ($T = \mathbf{1}'$) 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* $T = \mathbf{I}$ *instead of* $T = \mathbf{1}'$.

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 $g(\frac{\bar{\lambda}+\underline{\lambda}}{2})$, while his payoff from pooling the assets is $g[(\frac{\underline{\lambda}^{-1}+\bar{\lambda}^{-1}}{2})^{-1}]$, 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 ($T = \mathbf{1}'$) results in a less efficient allocation of risks across investors than not bundling ($T = \mathbf{I}$):

Proposition 4.3 *(Trade-Restriction Channel). If $X_1 + X_2 = f_1 + f_2$ and $K = 1$, $\forall i$*

1) If $T = \mathbf{I}$, then each type i's expected holding of risk f_i and f_{-i} are $\frac{\bar{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p}$ and $\frac{\underline{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p}$, respectively, where $\lambda_p = \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2$;

2) If $T = \mathbf{1}'$, 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 ($\underline{\lambda} + \lambda_p$) are a type i investor's knowledge about f_i and f_{-i}, respectively: $\bar{\lambda}$ (or $\underline{\lambda}$) from his private signal, and λ_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 $\frac{\bar{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p} > 1 > \frac{\underline{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p}$, as $\bar{\lambda} > \frac{\bar{\lambda}+\underline{\lambda}}{2} > \underline{\lambda}$.

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 ($\bar{\lambda} = \underline{\lambda}$), then $\frac{\bar{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p} = 1 = \frac{\underline{\lambda}+\lambda_p}{\frac{\bar{\lambda}+\underline{\lambda}}{2}+\lambda_p}$. 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 $\bar{\lambda} = \underline{\lambda}$, then the originator's payoff is $E_0[\sum X_i] - 2\rho[\frac{1}{\rho^2\sigma^2}\underline{\lambda}^2 + \underline{\lambda}]^{-1}$, whether $T = \mathbf{I}$ or $T = \mathbf{1}'$ 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 $T = \mathbf{I}$, 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, i.e., $T = \mathbf{1}'$? 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 *(Specialization-Destruction Channel). If $X_1 + X_2 = f_1 + f_2$, in the unique subgame equilibrium induced by $T = \mathbf{1}'$, each investor tries his best to equalize his knowledge of different risks:*

1) If $K \geq \bar{\lambda}/\underline{\lambda} > 1$, then $\lambda_1^{\alpha,i} = \lambda_2^{\alpha,i} = \sqrt{K\bar{\lambda}\underline{\lambda}}$, $\forall \alpha, i$;

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

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 $\bar{\lambda}/\underline{\lambda}$, 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,[12] the subgame equilibrium induced is unique.

When assets are not pooled, each investor focuses on acquiring information in his area of expertise, and his cost 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, 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 $K < \bar{\lambda}/\underline{\lambda}$, each investor lacks the capacity to equalize his knowledge of each risk, and both channels are in play. When $K \geqslant \bar{\lambda}/\underline{\lambda}$, 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

[12] 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 $p = A + Y + C\varepsilon = A + X_1 + X_2 + C\varepsilon = A + f_1 + f_2 + C\varepsilon$. Since the coefficients for f_1 and f_2 are both constantly 1 and do not vary with investors' private knowledge of f_1 and f_2, 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.

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.[13]

5. Discussion

5.1. General case

In Section 3 and 4, two polar cases ($w_2 = 0$ and 1, respectively) are used to highlight the three key economic channels of asset bundling. This subsection complements these two sections with a discussion of the general case: $w_2 \in [0, 1]$. Two main results — payoff continuity and threshold for payoff comparison with respect to w_2 — are presented sequentially.

Recall that by construction $w_1^2 + w_2^2 = 2$, and we consider only the range in which $w_1 \geq 1$ and $0 \leq w_2/w_1 \leq 1$. In this range, w_2/w_1 is continuous and strictly increasing in w_2. Therefore, payoff continuity and monotononicity with respect to w_2 and to w_2/w_1 are equivalent and thus stated interchangeably.

5.1.1. Payoff continuity

Proposition 5.1. *The originator's payoffs in the subgames induced by $T = \mathbf{1}'$ and $T = \mathbf{I}$ both change continuously with w_2/w_1.*

When assets are sold as they are, i.e., $T = \mathbf{I}$, type 1 investors all specialize in learning about f_1, and the proportion of type 2 investors specializing in f_2 changes continuously with w_2; when assets are bundled, i.e., $T = \mathbf{1}'$, similar to the situation in Section 4.2, regardless of his type, each investor minimizes the payoff variance of the bundle, $Var[w_1 f_1 + w_2 f_2]$, which is continuous in w_2/w_1, and this gives rise to the desired payoff continuity.

Note that the proposition holds for *any* combination of other parameter values. This also justifies the claim in the foreword of Section 3 (Section 4) that $T = \mathbf{I}$ ($T = \mathbf{1}'$) is optimal when w_2/w_1 is sufficiently small (large).

5.1.2. Threshold result for payoff comparison

The second result concerning the general case is that, when investors' learning capacity K is sufficiently large and/or risk aversion ρ is sufficiently low,[14] there is a threshold $w_2^* \in [0, 1)$ such that bundling ($T = \mathbf{1}'$) is optimal for the originator if and only if $w_2 \leq w_2^*$.[15]

[13] 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 independent and identically distributed across risks.

[14] The assumption of large learning capacity K and/or low risk aversion ρ is realistic for institutional investors, who are major holders of securized assets and stocks (see Section 6 for the discussion of the model in the context of conglomerate formation). If this assumption fails, when assets are not bundled, the proportion of type 2 investors specializing in their expertise is not necessarily 1, and cannot be solved in closed form. Also, liquidity trader risk and trade-restriction channel can be significant, complicating the analysis. This makes the originator's incentive of bundling the assets not necessarily monotonic in w_2/w_1.

[15] The author thanks the editor, Laura Veldkamp, and an anonymous referee for suggesting this direction.

Proposition 5.2. *$\exists\ \bar{\lambda}/\underline{\lambda} \leq K^* < \infty$ and $0 < \rho^* < \infty$ such that if $K > K^*$, or if $K > \bar{\lambda}/\underline{\lambda}$ and $\rho < \rho^*$, $\exists\ w_2^* \in (0, \sqrt{\frac{2\lambda}{K\bar{\lambda}+\underline{\lambda}}})$, such that the originator is strictly better off choosing $T = \mathbf{1}'$ ($T = \mathbf{I}$) if $w_2 < w_2^*$ ($w_2 > w_2^*$), and is indifferent if $w_2 = w_2^*$. And w_2^* is given by*

$$[(w_1^*)^2(K\bar{\lambda})^{-1} + (w_2^*)^2\underline{\lambda}^{-1}]^{-1} + [(w_1^*)^2(K\underline{\lambda})^{-1} + (w_2^*)^2\bar{\lambda}^{-1}]^{-1} = \frac{K\bar{\lambda} + \underline{\lambda}}{2}, \tag{4}$$

$$\text{where } (w_1^*)^2 = 2 - (w_2^*)^2.$$

To make sense of Proposition 5.2, we first analyse investors' learning behavior when assets are bundled and not bundled, respectively.

When assets are not bundled (i.e., $T = \mathbf{I}$), as discussed in Proposition 3.3, each investor devotes all his capacity to only one risk. The choice of the risk to specialize in learning about depends on equilibrium risk premium and on the informational advantage that he can create with his learning capacity. The latter dominates the former when investors are not too risk averse (i.e., with low ρ), and/or when they have large learning capacity K to resolve uncertainty they face. This makes both types specialize in their respective expertise regardless of relative portfolio weight of risks w_2/w_1, and the resulting payoff of the originator invariant with w_2/w_1.

When assets are bundled (i.e., $T = \mathbf{1}'$), the optimal capacity allocation of each investor minimizes the variance of the bundle he faces, $Var[w_1 f_1 + w_2 f_2]$. Similar to the situation in Section 4.2, this is achieved by trying his best to equalize $Var[w_1 f_1]$ and $Var[w_2 f_2]$. When w_2/w_1 is low, an investor cannot achieve such equalization perfectly even if he devotes all his learning capacity to the more important risk f_1, and therefore specializes in learning about f_1; When w_2/w_1 is high, he allocates his capacity to both risks to perfectly equalize $Var[w_1 f_1]$ and $Var[w_2 f_2]$. A type 1 investor is able to perfectly equalize $Var[w_1 f_1]$ and $Var[w_2 f_2]$ by allocating his capacity to both risks when $w_2/w_1 > \sqrt{\underline{\lambda}/K\bar{\lambda}}$,[16] and devotes all his capacity to f_1 otherwise. A similar situation holds for type 2 investors, but with a higher threshold $\sqrt{\bar{\lambda}/K\underline{\lambda}}$, since they know less about f_1 than type 1 to begin with. The resulting payoff of the originator turns out to be strictly decreasing in w_2/w_1.

The originator's payoff from bundling the assets strictly decreases with w_2/w_1, while his payoff of not bundling them is invariant with w_2/w_1. This, together with the fact that bundling is superior when $w_2 = 0$ (established in Section 3) and inferior when $w_2 = 1$ (established in Section 4), implies the existence of the desired threshold w_2^* for payoff comparison.

To make sense of equation (4), first recall from Section 4 that the trade-restriction channel disappears when investors have large learning capacity K (as assumed in Proposition 5.2). Therefore, the originator tradeoffs the discipline channel and the specialization-destruction channel when determining whether to bundle his assets. The threshold w_2^* balances the two channels. To make the discipline channel prevail, that is, to make it incentive compatible for all investors to specialize in learning about f_1 (the more important risk) when assets are bundled, w_2^* cannot

[16] When w_2/w_1 is large, $w_1^2/\bar{\lambda} = Var[w_1 f_1] < Var[w_2 f_2] = w_2^2/\underline{\lambda}$ for type 1 investors before learning, so they will first allocate capacity to f_2. But given the assumption that $K \geq \bar{\lambda}/\underline{\lambda}$, their capacity is more than enough to equalize $Var[w_1 f_1]$ and $Var[w_2 f_2]$, and thus it is never optimal for them to devote all their capacity to f_2.

be too large. This explains why $w_2^* < \sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}$ in Proposition 5.2.[17] Furthermore, w_2^* balances the two channels by equalizing market average knowledge of the market portfolio induced by bundling and not bundling, respectively. When $w_2 = w_2^*$, if assets are bundled, the discipline channel makes both types of investors specialize in f_1, and information aggregation happens with respect to the only tradable asset, the market portfolio $w_1 f_1 + w_2 f_2$. Hence, the resulting market average knowledge of the market portfolio is

$$\frac{1}{2}\left\{[(w_1^*)^2(K\bar{\lambda})^{-1} + (w_2^*)^2\underline{\lambda}^{-1}]^{-1} + [(w_1^*)^2(K\underline{\lambda})^{-1} + (w_2^*)^2\bar{\lambda}^{-1}]^{-1}\right\}.$$

The specialization-destruction channel of bundling implies that if assets are not bundled, each type of investors would instead specialize in their respective expertise when $w_2 = w_2^*$. In this case, contrary to the case of bundling, information aggregation happens with respect to each risk. This results in market average knowledge of both risks being $\frac{K\bar{\lambda}+\underline{\lambda}}{2}$, and that of the market portfolio being

$$\left\{(w_1^*)^2\left(\frac{K\bar{\lambda}+\underline{\lambda}}{2}\right)^{-1} + (w_2^*)^2\left(\frac{K\bar{\lambda}+\underline{\lambda}}{2}\right)^{-1}\right\}^{-1} = \frac{1}{2}\frac{K\bar{\lambda}+\underline{\lambda}}{2}.$$

At the threshold $w_2 = w_2^*$, whether assets are bundled or not, the resulting market average knowledge of the market portfolio is the same in order to make the originator indifferent. This interprets equation (4). Further manipulation of equation (4) yields a closed-form solution for w_2^* in Proposition 5.3.

Proposition 5.3. $w_2^* = \sqrt{\frac{-B+\sqrt{B^2-4AC}}{2A}}$, *where* $A = \frac{(KT+1)(KT-1)(K-T)}{2}$, $B = K^2T^3 - K^2T^2 + KT + K - 2T$, $C = 2T(1-K)$, *and* $T = \bar{\lambda}/\underline{\lambda}$.

Note that w_2^* only depends on the magnitude of expertise $\bar{\lambda}/\underline{\lambda}$ and learning capacity K. It does not depend on investors' risk aversion ρ or liquidity trader risk σ, because by assumption, investors either do not care much about risk (i.e., with low ρ), or have large learning capacity K to resolve uncertainty. Other than through $\bar{\lambda}/\underline{\lambda}$, neither does w_2^* depend on the *absolute* level of prior knowledge $\underline{\lambda}$ or $\bar{\lambda}$, because it equally affects the payoff of bundling the assets and that of not bundling them.

5.2. A generalization: the optimality of categorization strategy

This subsection demonstrates that the economic forces illustrated in Section 3 and 4 carry through to more general environments and can be combined. We generalize the baseline model to an n-risk-n-asset setup with n corresponding types of investor, and show the optimality of categorization strategy, which corresponds to pooling loans by asset class in the context of securitization.

[17] From the analysis just now on optimal learning choices when assets are bundled, both types specialize in f_1 if and only if $w_2/w_1 \le \sqrt{\underline{\lambda}/K\bar{\lambda}}$, or equivalently, $w_2^* \le \sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}$ since $w_1^2 + w_2^2 = 2$ by construction. Moreover, it is shown in the appendix that the specialization-destruction channel strictly dominates the discipline channel when $w_2^* = \sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}$.

The generalized setup is as follows: Payoffs of the original n assets are $\mathbf{X} = \Gamma\mathbf{f}$, where Γ is an $n \times n$ orthogonal risk loading matrix. Each type of investors has mass $1/n$. Each type i investor's exogenous signal $\mathbf{s}^{\alpha,i} \sim N(\mathbf{f}, (\Lambda_0^{(i)})^{-1})$, $\Lambda_0^{(i)} = diag(\lambda_{0,1}^{(i)}, ..., \lambda_{0,n}^{(i)})$, $\lambda_{0,i}^{(i)} = \bar{\lambda} > \underline{\lambda} = \lambda_{0,j}^{(i)}$ $\forall j \neq i$. 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 $\mathbf{1}_m' T = \mathbf{1}_n'$ and that the tradable assets have payoffs $\mathbf{Y} = T\mathbf{X}$. 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 $\forall i \leq i^*$, $w_i = w > 0$,[18] and $w_i = 0$ $\forall i > i^*$. That is, $f_1, ..., f_{i^*}$ are non-diversifiable and play symmetric roles, while $f_{i^*+1}, ..., f_n$ 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 defined formally as follows:

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

$$T\Gamma = \begin{pmatrix} w\mathbf{I}_{i^*} & \mathbf{0}_{(n-i^*)*i^*} \end{pmatrix} = \begin{pmatrix} w & 0 & \cdots & 0 & 0 & & 0 \\ 0 & w & & \vdots & & \ddots & \\ \vdots & & \ddots & 0 & & & \\ 0 & \cdots & 0 & w & 0 & & 0 \end{pmatrix}.$$

This strategy creates i^* tradable assets, such that $\mathbf{Y} = T\mathbf{X} = T\Gamma\mathbf{f} = (wf_1, wf_2, ..., wf_{i^*})'$. 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.

[18] The orthogonality of loading matrix Γ implies that $w = \sqrt{n/i^*}$.

Formally, this optimality benchmark is defined as:

$$\max_{\{\Lambda^{\alpha,i}\}} E_0[\sum_i p(X_i)] = E_0[\mathbf{1}'\mathbf{p}_{\mathbf{I}_n}] \tag{5}$$

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 *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:

Proposition 5.4. *If* $\exists\ 1 \le i^* \le n$*, such that* $\forall i \le i^*$, $w_i = w > 0$*, and* $w_i = 0\ \forall i > i^*$*, then the solution to the optimality benchmark is such that:*

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

2) each investor of type $i > i^*$ *specializes in one non-diversifiable risk* $j \le i^*$*, 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 cost advantage in information acquisition. Last, since all non-diversifiable risks carry equal weight in the objective function, and the objective function is concave in λ_i^a, 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 5.5. *If* $\exists\ 1 \le i^* \le n$*, such that* $\forall i \le i^*$, $w_i = w > 0$*, 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.

5.3. *Asymmetric prior precision*

In order to focus on the heterogeneity in the *dimension* of learning expertise, heterogeneity in the *magnitude* of learning expertise is precluded in the baseline model. That is, we assume symmetric prior precision: $\lambda_{0,i}^{i} = \bar{\lambda} > \underline{\lambda} = \lambda_{0,-i}^{i} > 0\ \forall i$. This subsection relaxes this assumption in the following way and probes into its impact on the three key economic channels introduced in Section 3 and 4:

Assumption 5.1. *$\lambda_{0,1}^{1} = \bar{\lambda}/\gamma$, $\lambda_{0,2}^{1} = \underline{\lambda}$, $\lambda_{0,1}^{2} = \underline{\lambda}/\gamma$, and $\lambda_{0,2}^{2} = \bar{\lambda}$, where $\gamma \in [1, \bar{\lambda}/\underline{\lambda})$.*

$\gamma > 1$ reduces the prior precision of both types of f_1. The learning technology featuring increasing return to scale implies that one unit of capacity spent on f_1 produces a lower gain in signal precision than on f_2. Besides $\lambda_{0,2}^{2} > \lambda_{0,1}^{2}$ as before, $\gamma < \bar{\lambda}/\underline{\lambda}$ guarantees that $\lambda_{0,1}^{1} > \lambda_{0,2}^{1}$, so that the expertise of each type is preserved.[19]

We first use the polar case in Section 3 — $X_1 + X_2 = \sqrt{2} f_1$ — to study the impact of asymmetric prior precision on the discipline channel. Intuitively, the asymmetric prior assumption does not alter the channel that asset bundling lowers risk premium by washing out idiosyncratic risks and channeling investors' learning capacity to non-diversifiable risks. Proposition 5.6 shows that the discipline channel introduced in Section 3 still works under Assumption 5.1, or under the alternative Assumption 5.2, where the reduction of prior precision is introduced to the idiosyncratic risk f_2.

Assumption 5.2. *$\lambda_{0,1}^{1} = \bar{\lambda}$, $\lambda_{0,2}^{1} = \underline{\lambda}/\gamma$, $\lambda_{0,1}^{2} = \underline{\lambda}$ and $\lambda_{0,2}^{2} = \bar{\lambda}/\gamma$, where $\gamma \in [1, \bar{\lambda}/\underline{\lambda})$.*

Proposition 5.6. *Propositions 3.1–3.4 still hold under Assumption 5.1 or Assumption 5.2.*

Now we use the polar case in Section 4 — $X_1 + X_2 = f_1 + f_2$ — to study the impact of asymmetric prior precision as formulated by Assumption 5.1[20] on the trade-restriction channel and the specialization-destruction channel. Unfortunately, results from analytical derivation are not available in this case.[21] Yet, numerical analysis shows similar patterns for different combinations of parameter values explored, as exemplified by the following one: $\bar{\lambda} = 10$, $\underline{\lambda} = 1$, $K = 20$, $\rho = 1$, $\sigma = 2$.

To make sense of the results, we decompose the difference of the originator's payoffs of not bundling and bundling the assets into the two channels:

$$\pi_{T=\mathbf{I}} - \pi_{T=\mathbf{1}'} = \left(\pi_{T=\mathbf{1}',T=\mathbf{I}} - \pi_{T=\mathbf{1}'}\right) + \left(\pi_{T=\mathbf{I}} - \pi_{T=\mathbf{1}',T=\mathbf{I}}\right),$$

[19] The author thanks an anonymous referee for suggesting this direction.

[20] As $w_1 = w_2$ in this case, the analysis under Assumption 5.2 is completely symmetric.

[21] This is because, for generic γ, if $T = \mathbf{I}$, the proportion of type 2 players specializing in learning about f_2 does not have a closed-form solution. Subsequently, neither does the payoff of the originator if $T = \mathbf{I}$.

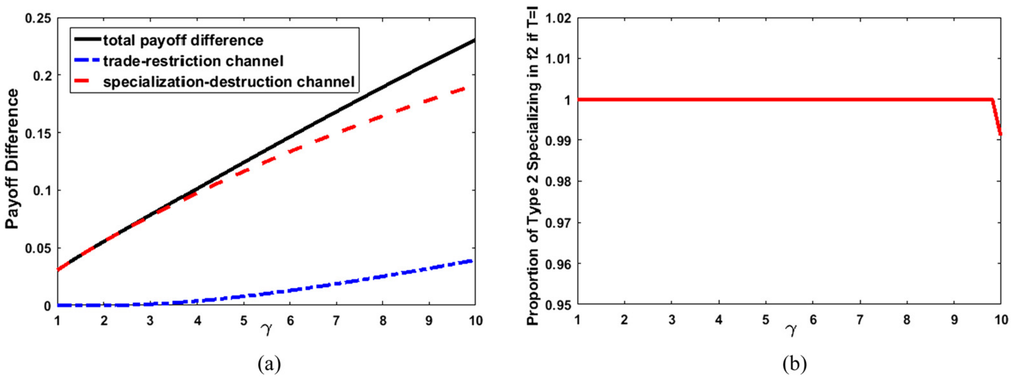


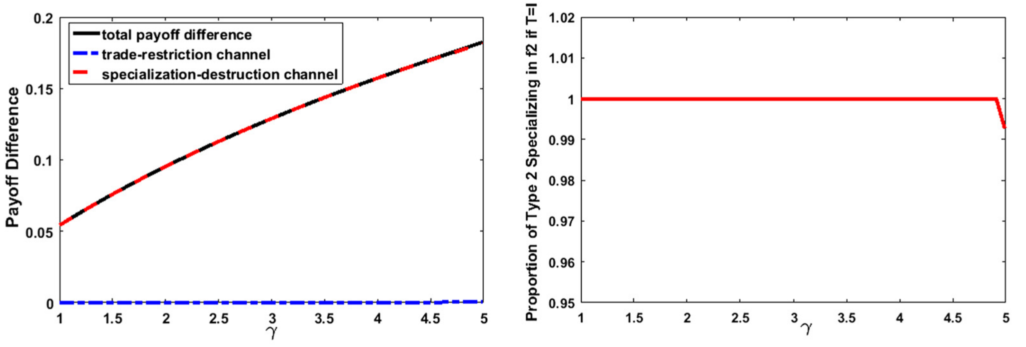
Fig. 1. $\bar{\lambda} = 10$, $\underline{\lambda} = 1$, $K = 20$, $\rho = 1$, $\sigma = 2$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

where $\pi_{T=\mathbf{1}',T=\mathbf{I}}$ is the hypothetical payoff of the originator if learning choices of investors follow bundling ($T = \mathbf{1}'$) while the tradable asset span follows not bundling ($T = \mathbf{I}$). The first difference $\left(\pi_{T=\mathbf{1}',T=\mathbf{I}} - \pi_{T=\mathbf{1}'}\right)$ refers to the trade-restriction channel, as both payoffs are derived from the same learning choices, with the only difference being tradable asset spans. In contrast, the second difference $\left(\pi_{T=\mathbf{I}} - \pi_{T=\mathbf{1}',T=\mathbf{I}}\right)$ refers to the specialization-destruction channel, as both payoffs are derived from the same tradable asset span, with the only difference being the learning choices induced.

To make sense of the pattern of the trade-restriction channel marked with the blue dotted line in Fig. 1a, first note that when $\gamma = 1$, both types of investors are able to perfectly equalize knowledge of the two risks with their learning capacity, as $K > \bar{\lambda}/\underline{\lambda}$. Type 1's prior precision of the less known risk relative to the more known one, $\lambda^1_{0,2}/\lambda^1_{0,1} = \gamma/\left(\bar{\lambda}/\underline{\lambda}\right)$, increases with γ. So their ability to equalize knowledge across risks increases with γ, and thus they can always achieve perfect knowledge equalization. In contrast, type 2's ability for such knowledge equalization decreases with γ, and can no longer achieve it perfectly when γ reaches a threshold. Hence, as γ increases from 1 to $\bar{\lambda}/\underline{\lambda}$, the trade-restriction channel first stays at zero and then starts to increase when γ reaches that threshold.[22]

To make sense of the pattern of the specialization-destruction channel marked with the red broken line in Fig. 1a, we need to understand how investors' learning choices vary with γ. When assets are not bundled (i.e., $T = \mathbf{I}$), under Assumption 5.1, each investor must still specialize in learning about one risk as before. As when $\gamma = 1$, type 1 investors must still specialize in f_1, the risk that is equally important, less known on average and type 1 have expertise in learning about. As shown in Fig. 1b, the proportion of type 2 investors specializing in f_2 remains 1 as when $\gamma = 1$ unless γ takes extremely large values, and the impact is negligible even in those cases. When assets are bundled (i.e., $T = \mathbf{1}'$), an increase in γ makes type 1 investors' prior knowledge of f_1 (which is $\lambda^1_{0,1} = \bar{\lambda}/\gamma$) and f_2 (which is $\lambda^1_{0,2} = \underline{\lambda}$) closer and reduces the capacity they allocate to f_2, the risk they do not have expertise in learning about, in knowledge equalization across risks. Meanwhile, an increase in γ widens the difference between type 2 investors' prior knowledge of f_1 (which is $\lambda^2_{0,1} = \underline{\lambda}/\gamma$) and f_2 (which is $\lambda^2_{0,2} = \bar{\lambda}$), and increases

[22] For other combinations of parameter values explored, the constant zero range of the trade-restriction channel does not exist if $K < \bar{\lambda}/\underline{\lambda}$. The strictly increasing range does not exist if K is so large that both types of investors can always achieve perfect knowledge equalization across risks.



(a) (b)

Fig. 2. $\bar{\lambda} = 5$, $\underline{\lambda} = 1$, $K = 20$, $\rho = 1$, $\sigma = 2$.

the capacity they allocate to f_1, the risk they do not have expertise in learning about. The latter effect dominates the former, resulting in greater waste of learning expertise as γ increases, because $\left(\lambda^2_{0,1}/\lambda^2_{0,2}\right) / \left(\lambda^1_{0,2}/\lambda^1_{0,1}\right) = 1/\gamma^2 < 1$ and decreases with γ. In addition, despite the fact that one unit of capacity yields a greater gain of signal precision if allocated to f_2, bundling the assets channels more capacity to f_1. This makes the originator even worse off pooling the assets than when $\gamma = 1$.

How does the magnitude of expertise $\bar{\lambda}/\underline{\lambda}$ affect the results concerning the trade-restriction channel and the specialization-destruction channel? Figs. 2a and 2b illustrate it by analyzing the following case: $\bar{\lambda} = 5$, $\underline{\lambda} = 1$, $K = 20$, $\rho = 1$, $\sigma = 2$. That is, everything is the same as the previous case except for $\bar{\lambda}$, which is now 5 instead of 10. Given the same learning capacity, the lower magnitude of expertise $\bar{\lambda}/\underline{\lambda}$ makes it easier for each investor to equalize his knowledge across risks, and thus weakens the trade-restriction channel and strengthens the specialization-destruction channel. As shown in Fig. 2a, the trade-restriction channel is constantly zero except for extremely large values of γ, and its magnitude is negligible even then. Meanwhile, the specialization-destruction channel still increases with γ as in Fig. 1a, and coincides with total payoff difference when the trade-restriction channel is zero. Fig. 2b shows the similar pattern to Fig. 1b: The proportion of type 2 investors specializing in f_2 remains 1 unless γ takes extremely large values, and the impact is negligible even then.

5.4. *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, 2015) 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.

5.5. *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.[23]

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

In Section 3, it was shown that, if there is only one non-diversifiable risk, the originator prefers pooling the assets ($T = \mathbf{1}'$) to selling the assets as they are ($T = \mathbf{I}$). 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.[24] Second, investors profit from the liquidity trader's demand, because the liquidity trader always moves price(s) against herself.

[23] The author thanks Markus Brunnermeier for suggesting this direction.

[24] The proof of Proposition A.1 in the appendix shows that an investor's expected utility is $\frac{1}{2}E\{(\mathbf{Y} - \mathbf{p})'[Var(\mathbf{Y})]^{-1}(\mathbf{Y} - \mathbf{p})\}$; i.e., expected return enters quadratically in the numerator, while variance enters linearly in the denominator.

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 fulfilling her "true" demand for the original assets, and causing her to lose less to investors.[25]

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.[26]

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.[27]

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?' " (Maksimovic and Phillips, 2007)

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

[25] 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.

[26] The author thanks Shang–Jin Wei for suggesting this direction.

[27] 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 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).[28] Only a 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 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.

[28] 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 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).

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.

Appendix A

A.1. Problems and propositions in the n-risk setup

The derivations in this subsection are in the n-risk setup introduced in Subsection 5.2. The 2-risk setup of the baseline model introduced in Section 2 is a special case, and the results here also apply. Since the learning stage of the model is based on Van Nieuwerburgh and Veldkamp (2009), the proofs also borrow largely from their work.

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

A.1.1. Price(s) of tradable assets and the originator's payoff

The posterior of investor α 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 $N(\boldsymbol{\mu}^{\alpha,i}, (\Omega^{\alpha,i})^{-1})$, where $\boldsymbol{\mu}^{\alpha,i} = T(\Lambda^{\alpha,i})^{-1}(\Lambda_0^{(i)}\mathbf{s}^{\alpha,i} + \Lambda_\eta^{\alpha,i}\boldsymbol{\eta}^{\alpha,i})$, and $\Omega^{\alpha,i} = (T\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T')^{-1}$.

Define $\Omega^a \equiv \int_{\alpha,i} \Omega^{\alpha,i}$, 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),[29] which gives the equilibrium prices as a function of asset payoffs $\mathbf{Y}$ and supply from liquidity traders $\boldsymbol{\varepsilon}$:

$$\begin{aligned} &\mathbf{p}_T = \mathbf{A}_T + \mathbf{Y} + C_T\boldsymbol{\varepsilon}_T, \text{ where} \\ &\mathbf{A}_T = -\rho[\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a + \Omega^a]^{-1}\mathbf{1} \\ &C_T = \rho(\Omega^a)^{-1} \end{aligned} \tag{6}$$

Note that $(\mathbf{p}_T - \mathbf{A}_T) \sim N(\mathbf{Y}, \Omega_{T,p}^{-1})$, where $\Omega_{T,p} = [C_T Var(\boldsymbol{\varepsilon}_T)C_T']^{-1} = \frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a$. Thus, we have:

Proposition A.1. *The originator's payoff is* $E_0[\mathbf{1}'\mathbf{p}_T] = \mathbf{E}_0[\mathbf{1}'(\mathbf{A}_T + \mathbf{Y} + C_T\boldsymbol{\varepsilon}_T)] = \mathbf{E}_0[\sum X_i] + \mathbf{1}'\mathbf{A}_T$, *where* $\mathbf{A}_T = -\rho[\Omega_{T,p} + \Omega^a]^{-1}\mathbf{1} = -\rho[\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a + \Omega^a]^{-1}\mathbf{1}$.

A.1.2. Portfolio choice and information acquisition of investors

The following proposition articulates the objective function of an investor.

Proposition A.2. *The information acquisition problem of investor* (α, i) *is*

$$\begin{aligned} &\max_{\Lambda^{\alpha,i}} Tr[\Omega^{\alpha,i}\Omega_{T,p}^{-1}] + A_T'\Omega^{\alpha,i}A_T \\ &\text{s.t. (2) and (3)} \end{aligned} \tag{7}$$

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

29 Investors in Admati (1985) have common priors, while we treat priors as though they were private signals. This is also the case in Van Nieuwerburgh and Veldkamp (2009), and thus the formula for $\mathbf{p}_T$ here is similar to theirs.

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

$$\mathbf{q}^{\alpha,i} = \frac{1}{\rho}\hat{\Omega}^{\alpha,i}(\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T) \tag{8}$$

Thus, ex ante, his expected utility is

$$U = E_0[\frac{1}{2}(\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T)'\hat{\Omega}^{\alpha,i}(\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T)]$$

He knows the distribution of his exogenous signal $\mathbf{s}^{\alpha,i}$ ex ante. Conditional on it, $(\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T)$ is a normal vector, with mean $-\mathbf{A}_T$, and variance $\Omega_{T,p}^{-1} - (\hat{\Omega}^{\alpha,i})^{-1}$. To derive this variance, note that $Var(\hat{\boldsymbol{\mu}}^{\alpha,i}|\Lambda_0^i) = T\Gamma(\Lambda_0^i)^{-1}\Gamma'T' - (\hat{\Omega}^{\alpha,i})^{-1}$, $Var(\mathbf{p}_T|\Lambda_0^i) = T\Gamma(\Lambda_0^i)^{-1}\Gamma'T' + \Omega_{T,p}^{-1}$, and $cov(\mathbf{p}_T, \hat{\boldsymbol{\mu}}^{\alpha,i}|\Lambda_0^i) = T\Gamma(\Lambda_0^i)^{-1}\Gamma'T'$.

If a generic random vector $\mathbf{z} \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $E[\mathbf{z}'\mathbf{z}] = \boldsymbol{\mu}'\boldsymbol{\mu} + Tr(\boldsymbol{\Sigma})$. Hence,

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

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

Now, we can prove Proposition 2.1.

Proof of Proposition 2.1. For 2 $m \times n$ bundling strategy T_1 and T_2, if they lead to the same asset span, then ∃ a full-rank $m \times m$ matrix M such that $\mathbf{1}'M = \mathbf{1}'$, and that $T_2 = MT_1$, then

$$\begin{aligned} \mathbf{1}'A_{T_2} &= \mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_2')^{-1}(T_2T_2')\int_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_2')^{-1} \\ &\quad + \int_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_2')^{-1}]^{-1}\mathbf{1} \\ &= \mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int_{\alpha,i}(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1'M')^{-1}(MT_1T_1'M')\int_{\alpha,i}(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1'M')^{-1} \\ &\quad + \int_{\alpha,i}(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1'M)^{-1}]^{-1}\mathbf{1} \\ &= \mathbf{1}'M[\frac{1}{\rho^2\sigma^2}\int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1}(T_1T_1')\int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1} \\ &\quad + \int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1}]^{-1}M'\mathbf{1} \\ &= \mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1}(T_1T_1')\int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1} \\ &\quad + \int_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma'T_1')^{-1}]^{-1}\mathbf{1} = \mathbf{1}'A_{T_1} \end{aligned}$$

So the originator has the same payoff.
Concerning the investor's information acquisition problem,
The first term in (7):

$$
\begin{aligned}
Tr[\Omega_{T_2}^{\alpha,i}\Omega_{T_2,p}^{-1}] &= Tr\{(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1} \\
&\quad \times [\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}(T_2T_2')\int\limits_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}]^{-1}\} \\
&= Tr\{(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1} \\
&\quad \times [\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}(MT_1T_1'M') \\
&\quad \times \int\limits_{\alpha,i}(MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1}]^{-1}\} \\
&= Tr\{M'^{-1}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1} \\
&\quad \times [\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1T_1') \\
&\quad \int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}]^{-1}M'\} \\
&= Tr[M'^{-1}\Omega_{T_1}^{\alpha,i}\Omega_{T_1,p}^{-1}M'] = Tr[\Omega_{T_1}^{\alpha,i}\Omega_{T_1,p}^{-1}].
\end{aligned}
$$

And because

$$
\begin{aligned}
&\mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}(T_2T_2')\int\limits_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1} \\
&\quad + \int\limits_{\alpha,i}(T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1}]^{-1} \\
&= \mathbf{1}'M[\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1T_1')\int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1} \\
&\quad + \int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}]M' \\
&= \mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}(T_1T_1')\int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1} \\
&\quad + \int\limits_{\alpha,i}(T_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1')^{-1}]M',
\end{aligned}
$$

and $\Omega_{T_2}^{\alpha,i} = (T_2\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_2')^{-1} = (MT_1\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma' T_1'M')^{-1} = M'^{-1}\Omega_{T_1}^{\alpha,i}M^{-1}$, the second term in (7) $A_{T_2}'\Omega_{T_2}^{\alpha,i}A_{T_2} = A_{T_1}'\Omega_{T_1}^{\alpha,i}A_{T_1}$.

Therefore, the information acquisition problem is also invariant. This concludes the proof. □

A.1.3. Information acquisition of investors and the originator's payoff in the subgame induced by $T = \mathbf{1}'$

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

Proof. If $T = \mathbf{1}'$, then all matrices in (7) are scalars, and the investor can only affect $\Omega^{\alpha,j}$. The objective function is a decreasing function of $(\Omega^{\alpha,i})^{-1} = \mathbf{1}'\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma\mathbf{1} = \sum_i \{w_i^2(\lambda_i^{\alpha,j})^{-1}\}$, and thus the investor chooses to minimize it. This is his dominant strategy, so the subgame equilibrium is unique. □

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 A.4. *In the unique subgame equilibrium induced by $T = \mathbf{1}'$, the originator's payoff is*

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

Proof. By Proposition A.1, the originator's payoff is

$$\begin{aligned}
\mathbf{E}_0[\sum X_i] + \mathbf{1}'\mathbf{A}_T &= \mathbf{E}_0[\sum X_i] - \rho\mathbf{1}'[\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a + \Omega^a]^{-1}\mathbf{1} \\
&= \mathbf{E}_0[\sum X_i] - \rho[\frac{1}{\rho^2\sigma^2}\int_{\alpha,j} (\mathbf{1}'\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma'\mathbf{1})^{-1}(\mathbf{1}'\mathbf{1}) \\
&\quad \times \int_{\alpha,j} (\mathbf{1}'\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma'\mathbf{1})^{-1} \\
&\quad + \int_{\alpha,j} (\mathbf{1}'\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma'\mathbf{1})^{-1}]^{-1} \\
&= \mathbf{E}_0[\sum X_i] - \rho\{\frac{n}{\rho^2\sigma^2}[\int_{\alpha,j} (\sum_i w_i^2(\lambda_i^{\alpha,j})^{-1})^{-1}]^2 \\
&\quad + \int_{\alpha,j} (\sum_i w_i^2(\lambda_i^{\alpha,j})^{-1})^{-1}\}^{-1}.
\end{aligned}$$

The second equality is due to the fact that both $\mathbf{1}'$ and $\mathbf{A}_T$ are scalers. □

A.1.4. Information acquisition of investors and the originator's payoff in the subgame induced by $T = \mathbf{I}$

Proposition A.5. *In any equilibrium induced by $T = \mathbf{I}$, investor (α, j) specializes in risk i_0, where $i_0 \in \arg\max_i\{L_i \lambda_{0,i}^j\}$, where*

$$L_i = \{\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2\}^{-1} + \{\rho w_i[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}\}^2$$

Proof. If $T = \mathbf{I}$, using the fact that Γ is an orthogonal matrix, the first term in (7) is

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

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

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

Since

$$A_T' = -\rho\mathbf{1}'[\frac{1}{\rho^2\sigma^2}\int_{\alpha,j}(\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma')^{-1}\int_{\alpha,j}(\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma')^{-1} + \int_{\alpha,j}(\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma')^{-1}]^{-1}$$

$$= -\rho\mathbf{1}'\Gamma[\frac{1}{\rho^2\sigma^2}(\int_{\alpha,j}\Lambda^{\alpha,j})^2 + \int_{\alpha,j}\Lambda^{\alpha,j}]^{-1}\Gamma', \text{ and}$$

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

the second term in (7) is

$$A_T'\Omega^{\alpha,j}A_T$$
$$= \rho^2\mathbf{1}'\Gamma[\frac{1}{\rho^2\sigma^2}(\int_{\alpha,j}\Lambda^{\alpha,j})^2 + \int_{\alpha,j}\Lambda^{\alpha,j}]^{-1}\Lambda^{\alpha,j}[\frac{1}{\rho^2\sigma^2}(\int_{\alpha,j}\Lambda^{\alpha,j})^2 + \int_{\alpha,j}\Lambda^{\alpha,j}]^{-1}\Gamma'\mathbf{1}$$
$$= \sum_{i=1}^{n}\{\rho w_i[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}\}^2\lambda_i^{\alpha,j}, \text{ since } \Lambda^{\alpha,j} \text{ is diagonal.}$$

Thus, let $L_i = \{\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2\}^{-1} + \{\rho w_i[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}\}^2$. The objective function (7) 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 (α, j) would specialize in risk i_0, where $i_0 \in \arg\max_i\{L_i\lambda_{0,i}^j\}$. □

Proposition A.6. *In a subgame equilibrium induced by $T = \mathbf{I}$, 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 A.1, the originator's payoff is

$$\begin{aligned}
\mathbf{E}_0[\sum X_i] + \mathbf{1}'\mathbf{A}_T &= \mathbf{E}_0[\sum X_i] - \rho \mathbf{1}'[\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a + \Omega^a]^{-1}\mathbf{1} \\
&= \mathbf{E}_0[\sum X_i] - \rho \mathbf{1}'[\frac{1}{\rho^2\sigma^2}[\int_{\alpha,i}(\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma')^{-1}]^2 \\
&\quad + \int_{\alpha,i}(\Gamma(\Lambda^{\alpha,i})^{-1}\Gamma')^{-1}]^{-1}\mathbf{1} \\
&= \mathbf{E}_0[\sum X_i] - \rho \mathbf{1}'\Gamma[\frac{1}{\rho^2\sigma^2}(\int_{\alpha,i}\Lambda^{\alpha,i})^2 + \int_{\alpha,i}\Lambda^{\alpha,i}]^{-1}\Gamma'\mathbf{1} \\
&= \mathbf{E}_0[\sum X_i] - \rho \sum_i w_i^2[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}. \qquad \square
\end{aligned}$$

A.1.5. Proof of propositions in Section 5.2

By Proposition A.5, the optimality benchmark (5) is

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

subject to capacity constraint (2) and no-forgetting constraint (3) $\forall \alpha, j$

In the intermediate case we considered: $\exists\ 1 \le i^* \le n$, such that $\forall i \le i^*$, $w_i = w > 0$, and $w_i = 0\ \forall i > i^*$, the benchmark is equivalent to:

$$\min_{\{\Lambda^{\alpha,j}\}} \sum_{i=1}^{i^*}[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}$$

subject to capacity constraint (2) and no-forgetting constraint (3) $\forall \alpha, j$

Proof of Proposition 5.4.

Step 1: In the solution to the optimality benchmark, no investor should learn about any idiosyncratic risk $i > i^*$, because the objective function is independent of λ_i^a, $\forall i > i^*$.

Step 2: "In the solution to the optimality benchmark, each investor (α, j) would specialize in one risk $i \le i^*$."

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 $\int_D \lambda_{i_1}^{\alpha,j} = K_1\lambda_{0,i_1}^j$, $\int_D \lambda_{i_2}^{\alpha,j} = K_2\lambda_{0,i_2}^j$, where $K_1 > 1$, $K_2 > 1$, and $K_1K_2 \le K$.

If instead we let $d \in (\frac{K_1-1}{K_1K_2-1}, \frac{K_1-1}{K_1K_2-1}(1-\frac{1}{K_1}))$ 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:

$$\int_D \tilde{\lambda}_{i_1}^{\alpha,j} = [dK_1K_2 + 1 - d]\lambda_{0,i_1}^j > K_1\lambda_{0,i_1}^j = \int_D \lambda_{i_1}^{\alpha,j}$$

$$\int_D \tilde{\lambda}_{i_2}^{\alpha,j} = [d + (1-d)K_1K_2]\lambda_{0,i_2}^j > K_2\lambda_{0,i_2}^j = \int_D \lambda_{i_2}^{\alpha,j}$$

$$\int_D \tilde{\lambda}_{i}^{\alpha,j} = \int_D \lambda_{i}^{\alpha,j} \ \forall i \neq i_1, i_2, \text{and}$$

$$\int_{D^C} \tilde{\lambda}_{i}^{\alpha,j} = \int_{D^C} \lambda_{i}^{\alpha,j} \ \forall i$$

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 $\{b_i^j : b_i^j$ is the proportion of type j investors who specialize in risk $i\}$.

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

If $b_{i_0}^{j_0} = d_1 > 0$ and $\exists k_0$ s.t. $b_{j_0}^{k_0} = d_2 > 0$. Let $d = \min\{d_1, d_2\}$ Consider a different allocation $\{\tilde{b}_i^j : \tilde{b}_{j_0}^{j_0} = b_{j_0}^{j_0} + d, \tilde{b}_{j_0}^{j_0} = d_1 - d, \tilde{b}_{i_0}^{k_0} = b_{i_0}^{k_0} + d, \tilde{b}_{j_0}^{k_0} = d_2 - d$, and $\tilde{b}_i^j = b_i^j$ otherwise$\}$. We have $\tilde{\lambda}_{j_0}^a = \lambda_{j_0}^a + \frac{dK(\bar{\lambda}-\underline{\lambda})}{n} > \lambda_{j_0}^a$, $\tilde{\lambda}_{i_0}^a = \lambda_{i_0}^a + \frac{dK[(\lambda_{0,i_0}^{k_0})^{-1}-\underline{\lambda}]}{n} \geq \lambda_{i_0}^a$, and $\tilde{\lambda}_{i}^a = \lambda_i^a \ \forall i \neq i_0, j_0$. This yields a strict improvement of the originator's payoff, which is a contradiction.

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

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 $b_{i_0}^{j_0} > 0$. This means $b_{j_0}^{j_0} \leq 1 - b_{i_0}^{j_0} < 1$. By step 3, $b_{j_0}^k = 0 \ \forall k \neq j_0$, and $b_{i_0}^{i_o} = 1$, and thus $\tilde{\lambda}_{j_0}^a = \frac{\bar{\lambda}+(n-1)\underline{\lambda}}{n} + \frac{b_{j_0}^{j_0}(K-1)\bar{\lambda}}{n} < \frac{\bar{\lambda}+(n-1)\underline{\lambda}}{n} + \frac{b_{i_0}^{i_o}(K-1)\bar{\lambda}}{n} < \frac{\bar{\lambda}+(n-1)\underline{\lambda}}{n} + \frac{b_{i_0}^{i_o}(K-1)\bar{\lambda}+(K-1)\underline{\lambda}\sum_{k\neq i_0} b_{i_0}^k}{n} = \lambda_{i_0}^a$.

Consider a different allocation $\{\tilde{b}_i^j : \tilde{b}_{j_0}^{j_0} = b_{j_0}^{j_0} + b_{i_0}^{j_0}, \tilde{b}_{i_0}^{j_0} = 0$, and $\tilde{b}_i^j = b_i^j$ otherwise$\}$.

$$\begin{aligned}
&\sum_{i=1}^{i^*}[\tilde{\lambda}_i^a + \frac{1}{\rho^2\sigma^2}(\tilde{\lambda}_i^a)^2]^{-1} - \sum_{i=1}^{i^*}[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1} \\
&= f[\lambda_{j_0}^a + \frac{b_{i_0}^{j_0}(K-1)\bar{\lambda}}{n}] + f[\lambda_{i_0}^a - \frac{b_{i_0}^{j_0}(K-1)\underline{\lambda}}{n}] - [f(\lambda_{j_0}^a) + f(\lambda_{i_0}^a)] \\
&< f[\lambda_{j_0}^a + \frac{b_{i_0}^{j_0}(K-1)\underline{\lambda}}{n}] + f[\lambda_{i_0}^a - \frac{b_{i_0}^{j_0}(K-1)\underline{\lambda}}{n}] - [f(\lambda_{j_0}^a) + f(\lambda_{i_0}^a)] < 0.
\end{aligned}$$

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

Suppose $\exists i_1, i_2 \le i^*$, s.t. $\sum_{i=1}^n b_{i_1}^j - \sum_{i=1}^n b_{i_2.}^j \equiv 2b > 0$. By step 4, $\sum_{i=1}^{i*} b_{i_1}^j = b_{i_1}^{i_1} = 1 = b_{i_2}^{i_2} = \sum_{i=1}^{i*} b_{i_2.}^j$, and thus $\sum_{i=1}^n b_{i_1}^j - \sum_{i=1}^n b_{i_2.}^j = \sum_{i=i^*}^n b_{i_1}^j - \sum_{i=i^*}^n b_{i_2.}^j = 2b > 0$. This implies $\lambda_{i_1}^a - \lambda_{i_2}^a = \frac{2b(K-1)\underline{\lambda}}{n} > 0$.

Consider an alternative allocation $\{\tilde{b}_i^j\}$, such that $\sum_{i=i^*}^n \tilde{b}_{i_1}^j = \sum_{i=i^*}^n b_{i_1}^j - b$, $\sum_{i=i^*}^n \tilde{b}_{i_2}^j = \sum_{i=i^*}^n b_{i_2}^j + b$, and $\tilde{b}_i^j = b_i^j$ otherwise. Then

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

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. □

Proof of Proposition 5.5. We first prove that the categorization strategy induces the capacity allocation of the solution to the optimality benchmark.

Note that

$$(\Omega^{\alpha,j})^{-1} = T\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma' T'$$
$$= \begin{pmatrix} w\mathbf{I}_{i^*} & \mathbf{0}_{(n-i^*)*i^*} \end{pmatrix} \begin{pmatrix} (\lambda_1^{\alpha,j})^{-1} & 0 & \cdots & 0 \\ 0 & (\lambda_2^{\alpha,j})^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & (\lambda_n^{\alpha,j})^{-1} \end{pmatrix} \begin{pmatrix} w\mathbf{I}_{i^*} \\ \mathbf{0}_{(n-i^*)*i^*} \end{pmatrix}$$
$$= w^2 \begin{pmatrix} (\lambda_1^{\alpha,j})^{-1} & 0 & \cdots & 0 \\ 0 & (\lambda_2^{\alpha,j})^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & (\lambda_{i^*}^{\alpha,j})^{-1} \end{pmatrix},$$

$$TT' = T\Gamma\Gamma' T' = \begin{pmatrix} w\mathbf{I}_{i^*} & \mathbf{0}_{(n-i^*)*i^*} \end{pmatrix} \begin{pmatrix} w\mathbf{I}_{i^*} \\ \mathbf{0}_{(n-i^*)*i^*} \end{pmatrix} = w^2\mathbf{I}_{i^*},$$

$$\Omega_{T,p} = \frac{1}{\rho^2\sigma^2}\Omega^a TT'\Omega^a = \frac{1}{\rho^2\sigma^2}\int\limits_{\alpha,j}(T\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma' T')^{-1} TT' \int\limits_{\alpha,j}(T\Gamma(\Lambda^{\alpha,j})^{-1}\Gamma' T')^{-1}$$
$$= \frac{1}{\rho^2\sigma^2} w^{-2} \begin{pmatrix} \lambda_1^a & 0 & \cdots & 0 \\ 0 & \lambda_2^a & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{i^*}^a \end{pmatrix} w^2 w^2 \begin{pmatrix} \lambda_1^a & 0 & \cdots & 0 \\ 0 & \lambda_2^a & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{i^*}^a \end{pmatrix}$$

$$= \frac{1}{\rho^2\sigma^2} w^{-2} \begin{pmatrix} (\lambda_1^a)^2 & 0 & \cdots & 0 \\ 0 & (\lambda_2^a)^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & (\lambda_{i^*}^a)^2 \end{pmatrix}.$$

And

$$\begin{aligned} A_T &= -\rho[\Omega^a + \Omega_{T,p}]^{-1}\mathbf{1} \\ &= -\rho[w^{-2} \begin{pmatrix} \lambda_1^a & 0 & \cdots & 0 \\ 0 & \lambda_2^a & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{i^*}^a \end{pmatrix} \\ &\quad + \frac{1}{\rho^2\sigma^2} w^{-2} \begin{pmatrix} (\lambda_1^a)^2 & 0 & \cdots & 0 \\ 0 & (\lambda_2^a)^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & (\lambda_{i^*}^a)^2 \end{pmatrix}]^{-1}\mathbf{1} \\ &= -\rho w^2 \begin{pmatrix} [\lambda_1^a + \frac{1}{\rho^2\sigma^2}(\lambda_1^a)^2]^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & [\lambda_{i^*}^a + \frac{1}{\rho^2\sigma^2}(\lambda_{i^*}^a)^2]^{-1} \end{pmatrix} \mathbf{1} \end{aligned}$$

So the first term in the objective function (7) is

$$\begin{aligned} Tr[\Omega^{\alpha,j}\Omega_{T,p}^{-1}] &= Tr\{w^{-2} \begin{pmatrix} \lambda_1^{\alpha,j} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_{i^*}^{\alpha,j} \end{pmatrix} \\ &\quad \times [\frac{1}{\rho^2\sigma^2} w^{-2} \begin{pmatrix} (\lambda_1^a)^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & (\lambda_{i^*}^a)^2 \end{pmatrix}]^{-1}\} \\ &= \sum_{i=1}^{i^*} \{\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2\}^{-1} \lambda_i^{\alpha,j} \end{aligned}$$

And the second term in (7) is

$$\begin{aligned} A_T'\Omega^{\alpha,j}A_T &= \mathbf{1}'(-\rho w^2) \begin{pmatrix} [\lambda_1^a + \frac{1}{\rho^2\sigma^2}(\lambda_1^a)^2]^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & [\lambda_{i^*}^a + \frac{1}{\rho^2\sigma^2}(\lambda_{i^*}^a)^2]^{-1} \end{pmatrix} \\ &\quad \times w^{-2} \begin{pmatrix} \lambda_1^{\alpha,j} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_{i^*}^{\alpha,j} \end{pmatrix} \end{aligned}$$

$$\times(-\rho w^2)\begin{pmatrix} [\lambda_1^a+\frac{1}{\rho^2\sigma^2}(\lambda_1^a)^2]^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & [\lambda_{i^*}^a+\frac{1}{\rho^2\sigma^2}(\lambda_{i^*}^a)^2]^{-1}\end{pmatrix}\mathbf{1}$$

$$=\sum_{i=1}^{i^*}\rho^2 w^2[\lambda_1^a+\frac{1}{\rho^2\sigma^2}(\lambda_1^a)^2]^{-2}\lambda_i^{\alpha,j}.$$

Let $V_i \equiv \{\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2\}^{-1}+\rho^2 w^2[\lambda_i^a+\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-2}$, $i=1,2,...,i^*$, and $V_i\equiv 0$ if $i>i^*$. With an argument analogous to the proof of Proposition A.4, it can be proved that investor (α, j) devotes all his capacity to a single risk i_0, where $i_0\in\arg\max_{i\le i^*}\{V_i\lambda_{0,i}^j\}$.

$V_k\equiv 0\ \forall k>i^*$ implies that no investor learns about diversifiable risks. $V_i=V_j>0\ \forall i,j\le i^*$ implies that $V_i\lambda_{0,i}^i=V_i\bar{\lambda}>V_i\underline{\lambda}=V_j\lambda_{0,j}^i>0=V_k\lambda_{0,k}^i\ \forall i,j\le i^*, i\ne j, \forall k>i^*$, which verifies that every type $i\le i^*$ investor would specialize in learning about his own risk f_i. And since $V_j\lambda_{0,j}^k=V_i\lambda_{0,i}^k\ \forall i,j\le i^*$, 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 $j\le i^*$ investor would specialize in his own risk j in any subgame equilibrium. Suppose among type $j_1, j_2, ..., j_m\le i^*$ investors, there are respectively a strictly positive proportion $b_{j_1}, b_{j_2}, ..., b_{j_m}$ who are not learning about their own risks, and WLOG $b_{j_1}\ge b_{j_2}\ge ...\ge b_{j_m}$. By Proposition A.4, for each of such risk $j_k, k=1,2,...,m$, there exists a different risk $\tilde{j}_k\ne j_k$ such that $V_{j_k}\bar{\lambda}\le V_{\tilde{j}_k}\underline{\lambda}$. This implies $V_{j_k}\underline{\lambda}<V_{\tilde{j}_k}\underline{\lambda}<V_{\tilde{j}_k}\bar{\lambda}$, i.e. none of non-type j_k investors would specialize in risk j_k. This implies $\Lambda_{j_1}^a\ge\Lambda_j^a\ \forall j\le i^*$, and thus $V_{j_1}\ge V_j\ \forall j\le i^*$. Now we have $V_{j_1}\bar{\lambda}\ge V_{\tilde{j}_k}\bar{\lambda}>V_{\tilde{j}_k}\underline{\lambda}$, 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 $V_{i_0}>V_i$ $\forall i\ne i_0$, 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, $\mathbf{E}_0[\sum X_i]+\mathbf{1}'A_T$, is identical to that of the optimality benchmark. This concludes the proof. □

A.2. Propositions in the 2-risk setup

Proposition 3.1 is a special case of Proposition A.4 and A.6, where $n=2$, $w_1^2=2$ and $w_2^2=0$.

Proof of Proposition 3.2 is straightforward from Proposition A.3, as $\sum_i\{w_i^2(\lambda_i^{\alpha,j})^{-1}\}=2(\lambda_1^{\alpha,j})^{-1}$.

Proof of Proposition 3.3. We use Proposition A.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 $L_1\bar{\lambda}\le L_2\underline{\lambda}$. This implies $L_1\underline{\lambda}<L_2\bar{\lambda}$, 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 $L_1\bar{\lambda}>L_2\underline{\lambda}$, a contradiction. So all type 1 investors learn only about f_1.

Suppose all type 2 investors also learn only about f_1. Then $L_2\lambda_{0,2}^2=\{\frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\}^{-1}\bar{\lambda}$, and $L_1\lambda_{0,1}^2\to 0$ as $K\to\infty$. Thus, $\exists K_0<\infty$ such that a positive proportion of type 2 investors learn about f_2 if $K>K_0$. □

Proof of Proposition 4.5. Again by Proposition A.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 $L_1\bar{\lambda} > L_2\underline{\lambda}$ and $L_1\underline{\lambda} < L_2\bar{\lambda}$, justifying each investor's choice.

The equilibrium is unique. Otherwise, say WLOG if some type 2 investors learn about f_1 in equilibrium, then $L_1\underline{\lambda} \geq L_2\bar{\lambda}$, which implies $L_1\bar{\lambda} > L_2\underline{\lambda}$; 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 $L_1\underline{\lambda} < L_2\bar{\lambda}$, a contradiction. □

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

By Proposition A.3, the optimization problem for investor (α, j) is $\min_{\{(\lambda_i^{\alpha,j})^{-1}\}} \sum_i (\lambda_i^{\alpha,j})^{-1}$ s.t. $(\lambda_i^{\alpha,j})^{-1} \leq (\lambda_{0,i}^{j})^{-1}$ and $\prod_i (\lambda_i^{\alpha,j})^{-1} \geq \frac{1}{K}(\bar{\lambda}\underline{\lambda})^{-1}$, $\forall i$. The first order condition for this problem is $1 - \frac{v}{(\lambda_i^{\alpha,j})^{-1}} \prod_l (\lambda_l^{\alpha,j})^{-1} + z_i = 0$, 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 $(\lambda_i^{\alpha,j})^{-1} = \frac{v}{K}(\bar{\lambda}\underline{\lambda})^{-1}$. Taking a product on both sides and imposing the capacity constraint again yields $v = (\frac{1}{K}(\bar{\lambda}\underline{\lambda})^{-1})^{-\frac{1}{2}}$, and thus $(\lambda_i^{\alpha,j})^{-1} = (\frac{1}{K}(\bar{\lambda}\underline{\lambda})^{-1})^{\frac{1}{2}}$ which strictly decreases in K, verifying the guess.

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

Proof of Proposition 4.1. By Proposition 4.5, in the subgame equilibrium induced by $T = \mathbf{I}$, $\lambda_1^a = \lambda_2^a = \frac{K\bar{\lambda}+\underline{\lambda}}{2}$. So by Proposition A.6, the originator's payoff is

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

By Proposition 4.6, in the subgame equilibrium induced by $T = \mathbf{1}'$, if $K \geq \bar{\lambda}/\underline{\lambda}$, $(\lambda_1^{\alpha,j})^{-1} + (\lambda_2^{\alpha,j})^{-1} = 2(\sqrt{K\bar{\lambda}\underline{\lambda}})^{-1}$ $\forall \alpha, j$. So by Proposition A.4, the originator's payoff is

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

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

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

Proof of Proposition 4.3. By (8), the expected portfolio holdings of investor (α, i) are $E[\mathbf{q}_T^{\alpha,i}] = \frac{1}{\rho}\hat{\Omega}^{\alpha,i} E[\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T]$. As discussed in the proof of Proposition A.2, $E[\hat{\boldsymbol{\mu}}^{\alpha,i} - \mathbf{p}_T] = -A_T = \rho[\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a + \Omega^a]^{-1}\mathbf{1}$, and $\hat{\Omega}^{\alpha,i} = \Omega^{\alpha,i} + \frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a$.

If $T = \mathbf{I}$, by orthogonality of Γ, $\Omega^{\alpha,i} = [\Gamma(\Lambda_0^i)^{-1}\Gamma]^{-1} = \Gamma\Lambda_0^i\Gamma'$. $\Omega^a = \int_{\alpha,i}\Omega^{\alpha,i} = \Gamma\Lambda_0^a\Gamma'$, where $\Lambda_0^a = \int_i \Lambda_0^i = \frac{\bar{\lambda}+\underline{\lambda}}{2}\mathbf{I}$. $\frac{1}{\rho^2\sigma^2}\Omega^a(TT')\Omega^a = \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\mathbf{I}$.

Hence,

$$E[\mathbf{q}_T^{\alpha,i}] = \Gamma(\Lambda_0^i + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\mathbf{I})\Gamma'\Gamma(\frac{\bar{\lambda}+\underline{\lambda}}{2}\mathbf{I} + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\mathbf{I})^{-1}\Gamma'\mathbf{1}$$
$$= [\frac{\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1}\Gamma(\Lambda_0^i + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\mathbf{I})\mathbf{1}.$$

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 $\Gamma' E[\mathbf{q}_T^{\alpha,i}] = [\frac{\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1}(\Lambda_0^i + \frac{1}{\rho^2\sigma^2}(\frac{\bar{\lambda}+\underline{\lambda}}{2})^2\mathbf{I})\mathbf{1}$. Since $\Lambda_0^1 = diag(\bar{\lambda}, \underline{\lambda})$ and $\Lambda_0^2 = diag(\underline{\lambda}, \bar{\lambda})$, this proves the first statement.

If $T = \mathbf{1}'$, $\forall(\alpha, i)$, $\Omega^{\alpha,i} = (\bar{\lambda} + \underline{\lambda})^{-1} = \Omega^a$. 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. □

Proof of Proposition 4.4 is straightforward from Proposition 4.1.

Proof of Proposition 5.7. It is shown in the proof of Proposition A.2 that $2U^{\alpha,i} = Tr[\Omega^{\alpha,i}\Omega_{T,p}^{-1}] + A_T'\Omega^{\alpha,i}A_T + A_T'\Omega_{T,p}A_T$. Thus,

$$2U^a \equiv \int_{\alpha,i} 2U^{\alpha,i} = \int_{\alpha,i} Tr[\Omega^{\alpha,i}\Omega_{T,p}^{-1}] + \int_{\alpha,i}(A_T'\Omega^{\alpha,i}A_T + A_T'\Omega_{T,p}A_T)$$
$$= Tr[\Omega^a\Omega_{T,p}^{-1}] + A_T'(\Omega^a + \Omega_{T,p})A_T$$
$$= Tr[\rho^2\sigma^2(TT')^{-1}(\Omega^a)^{-1}] + \rho^2\mathbf{1}'(\Omega^a + \Omega_{T,p})\mathbf{1}.$$

Note that the second term $= -\rho\mathbf{1}'A_T$. We know from Proposition A.1 that $\mathbf{1}'A_T$ is the greatest for the originator's favored bundling strategy. Hence, the second term is smaller for $T = \mathbf{1}'$ than for $T = \mathbf{I}$. This comparison corresponds to the first reason in the explanation of Proposition 5.7 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 $w_2^2 = 0$,

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

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

The first inequality corresponds to the second reason in the explanation of Proposition 5.7 in the text, and the second inequality to the third reason. □

Proof of Proposition 5.1. 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 $T = \mathbf{1}'$, to show that the originator's payoff changes continuously with w_2, by Proposition A.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 4.6, 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 $w_1^2\bar{\lambda}^{-1} \geq w_2^2\underline{\lambda}^{-1}$, he goes for f_1 first before learning about f_2:

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

if $w_1^2(K\bar{\lambda})^{-1} \leq w_2^2\underline{\lambda}^{-1}$, then $w_1^2(\lambda_1^{\alpha,1})^{-1} = w_2^2(\lambda_2^{\alpha,1})^{-1} = w_1 w_2\sqrt{(K\bar{\lambda}\underline{\lambda})^{-1}}$;

if $w_1^2\bar{\lambda}^{-1} < w_2^2\underline{\lambda}^{-1}$, he goes for f_2 first before learning about f_1:

if $w_1^2\bar{\lambda}^{-1} < w_2^2(K\underline{\lambda})^{-1}$, then $\lambda_1^{\alpha,1} = \bar{\lambda}$, $\lambda_2^{\alpha,1} = K\underline{\lambda}$;

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

Therefore, $$w_1^2(\lambda_1^{\alpha,1})^{-1} + w_2^2(\lambda_2^{\alpha,1})^{-1} = \begin{cases} w_1^2(K\bar{\lambda})^{-1} + w_2^2\underline{\lambda}^{-1}, \text{ if } \frac{w_2^2}{w_1^2} < \frac{\underline{\lambda}}{K\bar{\lambda}} \\ 2w_1w_2\sqrt{(K\bar{\lambda}\underline{\lambda})^{-1}}, \text{ if } \frac{\underline{\lambda}}{K\bar{\lambda}} \leq \frac{w_2^2}{w_1^2} \leq \frac{K\underline{\lambda}}{\bar{\lambda}} \\ w_1^2\bar{\lambda}^{-1} + w_2^2(K\underline{\lambda})^{-1}, \text{ if } \frac{w_2^2}{w_1^2} > \frac{K\underline{\lambda}}{\bar{\lambda}} \end{cases}.$$

Since $\frac{w_2^2}{w_1^2}$ strictly increases with w_2 in the range of interest, $w_1^2(\lambda_1^{\alpha,1})^{-1} + w_2^2(\lambda_2^{\alpha,1})^{-1}$ is continuous in w_2 in each segment, and its value coincides at the two thresholds.

For a type 2 investor, since $w_1^2\underline{\lambda}^{-1} > w_2^2\bar{\lambda}^{-1}$, he always goes for f_1 first before learning about f_2:

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

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

Therefore, $$w_1^2(\lambda_1^{\alpha,2})^{-1} + w_2^2(\lambda_2^{\alpha,2})^{-1} = \begin{cases} w_1^2(K\underline{\lambda})^{-1} + w_2^2\bar{\lambda}^{-1}, \text{ if } \frac{w_2^2}{w_1^2} < \frac{\bar{\lambda}}{K\underline{\lambda}} \\ 2w_1w_2\sqrt{(K\bar{\lambda}\underline{\lambda})^{-1}}, \text{ if } \frac{w_2^2}{w_1^2} \geq \frac{\bar{\lambda}}{K\underline{\lambda}} \end{cases},$$ which is again continuous in w_2 in each segment, and its value coincides at the threshold.

Thus, we have proved that if $T = \mathbf{1}'$, then the originator's payoff changes continuously with w_2.

Now we prove that the same conclusion holds if $T = \mathbf{I}$.

As in Proposition A.5, let $L_i = \{\frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2\}^{-1} + \{\rho w_i[\lambda_i^a + \frac{1}{\rho^2\sigma^2}(\lambda_i^a)^2]^{-1}\}^2$, $i = 1, 2$.

Since $w_1 \geq w_2 \geq 0$, by an argument analogous to the proof of Proposition 4.5, 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 $\lambda_1^a = \frac{K(\bar{\lambda}+\underline{\lambda})}{2} - b\frac{(K-1)\underline{\lambda}}{2}$, which strictly decreases in b, and $\lambda_2^a = \frac{\bar{\lambda}+\underline{\lambda}}{2} + b\frac{(K-1)\bar{\lambda}}{2}$, which strictly increases in b.

L_1 and L_2 can be viewed as functions of b, w_2, K, ρ and σ. L_1 strictly increases in b, decreases in w_2 and K. L_2 strictly decreases in b, increases in w_2, and decreases in K.

Let $\underline{w}$ be such that $\underline{\lambda}L_1(0, \underline{w}; K, \rho, \sigma) = \bar{\lambda}L_2(0, \underline{w}; K, \rho, \sigma)$, and $\bar{w}$ be such that $\underline{\lambda}L_1(1, \bar{w}; K, \rho, \sigma) = \bar{\lambda}L_2(1, \bar{w}; K, \rho, \sigma)$. By the monotonicity of L_1 and L_2 in the first and the second arguments, $\underline{w} < \bar{w}$.

The differentiability of L_1 and L_2 and monotonicity of them in the first and the second arguments also implies that there is a differentiable and strictly increasing function $\hat{b}(w_2)$ defined implicitly by $\underline{\lambda} L_1(\hat{b}, w_2; K, \rho, \sigma) = \bar{\lambda} L_2(\hat{b}, w_2; K, \rho, \sigma)$.

If $w_2 \leq \underline{w}$, we have $\underline{\lambda} L_1(0, w_2; K, \rho, \sigma) \geq \bar{\lambda} L_2(0, w_2; K, \rho, \sigma)$. This means that every type 2 investor weakly prefers to learn about f_1 even if so does everyone else. This implies that equilibrium $b = 0$. A symmetric argument implies that equilibrium $b = 1$ if $w_2 \geq \bar{w}$.

If $w_2 \in (\underline{w}, \bar{w})$, we have $\hat{b}(w_2) \in (0, 1)$ and $\underline{\lambda} L_1(\hat{b}(w_2), w_2; K, \rho, \sigma) = \bar{\lambda} L_2(\hat{b}(w_2), w_2; K, \rho, \sigma)$. This means that every type 2 investor is indifferent between learning about f_1 and f_2 if exactly $\hat{b}(w_2)$ proportion of them learns about f_2. This implies that equilibrium $b = \hat{b}(w_2)$.

Thus, equilibrium $b = \begin{cases} 0, \text{ if } w_2 \leq \underline{w} \\ \hat{b}(w_2) \text{ if } w_2 \in (\underline{w}, \bar{w}) \\ 1 \text{ if } w_2 \geq \bar{w} \end{cases}$, which is a continuous and weakly increasing function of w_2. This concludes the proof. □

To prove Proposition 5.2, we first provide a formal derivation for the analysis of investors' learning behavior and payoff monotonicity in the text when $T = \mathbf{1}'$ and when $T = \mathbf{I}$.

To facilitate discussion, let $\pi_{T=\mathbf{1}'}$ and $\pi_{T=\mathbf{I}}$ be the originator's payoff of choosing bundling ($T = \mathbf{1}'$) and not bundling ($T = \mathbf{I}$), respectively.

Lemma A.1. *If $T = \mathbf{I}$, then $\exists\, 0 < K^{**} < \infty$ and $0 < \rho^* < \infty$ such that if $K \geq K^{**}$ or $\rho \leq \rho^*$, all type i ($i = 1, 2$) investors specialize in learning about f_i in the subgame equilibrium induced, and $\pi_{T=\mathbf{1}'} = E_0[\sum X_i] - 2\rho[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1}$ $\forall w_2 \in [0, 1]$.*

Proof. It is shown in the proof of Proposition 5.1 that all type 1 investors specialize in learning about f_1 in the subgame equilibrium induced.

To see the equilibrium learning behavior of type 2 investors, $\forall w_2 \in [0, 1]$,

$$\begin{aligned}
\frac{\underline{\lambda} L_1(1, w_2; K, \rho, \sigma)}{\bar{\lambda} L_2(1, w_2; K, \rho, \sigma)} &\leq \frac{\underline{\lambda} L_1(1, 0; K, \rho, \sigma)}{\bar{\lambda} L_2(1, 0; K, \rho, \sigma)} \\
&= \frac{\underline{\lambda}}{\bar{\lambda}} \frac{\{\frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2\}^{-1} + 2\rho^2[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-2}}{\{\frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2\}^{-1}} \\
&= \frac{\underline{\lambda}}{\bar{\lambda}} \left[1 + \frac{2}{\sigma^2}\left(1 + \frac{1}{\rho^2\sigma^2}\frac{K\bar{\lambda}+\underline{\lambda}}{2}\right)^{-2}\right].
\end{aligned}$$

The last expression is strictly decreasing in K and increasing in ρ, and converges to $\frac{\underline{\lambda}}{\bar{\lambda}} < 1$ as $K \to \infty$ or $\rho \to 0$. This implies that when K is large enough or when ρ is small enough, all type 2 investors strictly prefer to learn about f_2 if so do all others of the same type. This proves the first claim.

By Proposition A.6, $\forall w_2 \in [0, 1]$,

$$\begin{aligned}
\pi_{T=\mathbf{I}} &= E_0[\sum X_i] - \rho \sum_{i=1,2} w_i^2 \left[\lambda_i^a + \frac{1}{\rho^2\sigma^2}\left(\lambda_i^a\right)^2\right]^{-1} \\
&= E_0[\sum X_i] - \rho\left(w_1^2 + w_2^2\right)[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1} \\
&= E_0[\sum X_i] - 2\rho[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1} \qquad \square
\end{aligned}$$

Lemma A.2. *Let*

$$\Omega_{\lambda_1,\lambda_2} = \left[w_1^2\lambda_1^{-1} + w_2^2\lambda_2^{-1}\right]^{-1}. \tag{9}$$

If $K \geq \bar{\lambda}/\underline{\lambda}$, then

$$\pi_{T=\mathbf{1}'} = \begin{cases} E_0[\sum X_i] - 2\rho\left\{\Omega_{K\bar{\lambda},\underline{\lambda}} + \Omega_{K\underline{\lambda},\bar{\lambda}} + \frac{1}{\rho^2\sigma^2}(\Omega_{K\bar{\lambda},\underline{\lambda}} + \Omega_{K\underline{\lambda},\bar{\lambda}})^2\right\}^{-1}, \\ \quad \text{if } w_2^2/w_1^2 \leq \frac{\underline{\lambda}}{k\bar{\lambda}} \\ E_0[\sum X_i] - 2\rho\left\{\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \Omega_{K\underline{\lambda},\bar{\lambda}} + \frac{1}{\rho^2\sigma^2}(\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \Omega_{K\underline{\lambda},\bar{\lambda}})^2\right\}^{-1}, \\ \quad \text{if } w_2^2/w_1^2 \in [\frac{\underline{\lambda}}{k\bar{\lambda}}, \frac{\bar{\lambda}}{k\underline{\lambda}}] \\ E_0[\sum X_i] - 2\rho\left\{\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{w_1w_2} + \frac{1}{\rho^2\sigma^2}(\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{w_1w_2})^2\right\}^{-1}, \\ \quad \text{if } w_2^2/w_1^2 \geq \frac{\bar{\lambda}}{k\underline{\lambda}}, \end{cases}$$

and $\pi_{T=\mathbf{1}'}$ is continuous and decreasing in w_2/w_1.

Proof. If $K \geq \bar{\lambda}/\underline{\lambda}$, $\frac{K\underline{\lambda}}{\bar{\lambda}} \geq 1$. As shown in the proof of Proposition 5.1,

$$\left[w_1^2(\lambda_1^{\alpha,1})^{-1} + w_2^2(\lambda_2^{\alpha,1})^{-1}\right]^{-1} = \begin{cases} \Omega_{K\bar{\lambda},\underline{\lambda}}, \text{ if } \frac{w_2^2}{w_1^2} < \frac{\underline{\lambda}}{K\bar{\lambda}} \\ \frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2}, \text{ if } \frac{\underline{\lambda}}{K\bar{\lambda}} \leq \frac{w_2^2}{w_1^2} \leq 1 \end{cases}$$

and is continuous at $\frac{w_2^2}{w_1^2} = \frac{\underline{\lambda}}{K\bar{\lambda}}$, and

$$\left[w_1^2(\lambda_1^{\alpha,2})^{-1} + w_2^2(\lambda_2^{\alpha,2})^{-1}\right]^{-1} = \begin{cases} \Omega_{K\underline{\lambda},\bar{\lambda}}, \text{ if } \frac{w_2^2}{w_1^2} < \frac{\bar{\lambda}}{K\underline{\lambda}} \\ \frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2}, \text{ if } \frac{w_2^2}{w_1^2} \geq \frac{\bar{\lambda}}{K\underline{\lambda}} \end{cases}$$

and is continuous at $\frac{w_2^2}{w_1^2} = \frac{\bar{\lambda}}{K\underline{\lambda}}$.

The formula of $\pi_{T=\mathbf{1}'}$ and its continuity with respect to w_2/w_1 then follows from Proposition A.4. To show the monotonicity, it suffices to show that $\Omega_{K\bar{\lambda},\underline{\lambda}}$ and $\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{w_1w_2}$ strictly decrease with w_2/w_1, and that $\Omega_{K\underline{\lambda},\bar{\lambda}}$ weakly decreases with w_2/w_1. This is the case for $\Omega_{K\bar{\lambda},\underline{\lambda}}$ and $\Omega_{K\underline{\lambda},\bar{\lambda}}$, since $K\bar{\lambda} > K\underline{\lambda} \geq \bar{\lambda} > \underline{\lambda}$. To see the monotonicity of $\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{w_1w_2}$ with respect to w_2/w_1, note that $w_1w_2 = \sqrt{2-w_2^2}w_2$, $\frac{d}{dw_2}\left(\sqrt{2-w_2^2}w_2\right) = \frac{2-w_2^2-w_2}{\sqrt{2-w_2^2}} > 0$ for $w_2 \in [0,1)$, and that $w_2/w_1 = w_2/\sqrt{2-w_2^2}$ strictly increases with w_2 for $w_2 \in [0,1]$. This concludes the proof. □

Now we can prove Proposition 5.2. Since we fix all parameter values other than w_2 (or equivalently w_2/w_1, since $w_1^2 + w_2^2 = 2$), for notational convenience, we treat $\pi_{T=\mathbf{1}'}$ and $\pi_{T=\mathbf{I}}$ as functions of w_2 and suppress all other parameters.

Proof of Proposition 5.2. Let $K^* = \max\left\{K^{**}, \bar{\lambda}/\underline{\lambda}\right\}$, where K^{**} is the one in Lemma A.1. If $K > K^*$, or if $K > \bar{\lambda}/\underline{\lambda}$ and $\rho < \rho^*$, where ρ^* is the one in Lemma A.1, by Lemma A.1, $\pi_{T=\mathbf{I}}$ is a constant. And by Lemma A.2, $\pi_{T=\mathbf{1}'}$ is continuous in and strictly decreases with w_2. If in addition we have $\pi_{T=\mathbf{1}'}(0) > \pi_{T=\mathbf{I}}$. and $\pi_{T=\mathbf{1}'}(1) < \pi_{T=\mathbf{I}}$, then $\pi_{T=\mathbf{1}'}$ crosses $\pi_{T=\mathbf{I}}$ exactly once from top, which proves the existence of the threshold w_2^*.

$\pi_{T=\mathbf{1}'}(1) < \pi_{T=\mathbf{I}}$ directly results from Proposition 4.2. To see why $\pi_{T=\mathbf{1}'}(0) > \pi_{T=\mathbf{I}}(0)$, recall from Proposition 3.1 and 3.2 that it suffices to show that if $T = \mathbf{I}$ and $w_2 = 0$, a positive mass of type 2 investors learn about f_2. And Lemma A.1 shows that none of type 2 investors learn about f_1 in that case.

To characterize the threshold w_2^*, we first show that $w_2^* \in \left(0, \sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}\right)$, or equivalently $0 < (w_2^*/w_1^*)^2 < \frac{\underline{\lambda}}{K\bar{\lambda}}$. Because $\pi_{T=\mathbf{1}'}$ crosses $\pi_{T=\mathbf{I}}$ exactly once from top, it suffices to show that $\pi_{T=\mathbf{1}'}\left(\sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}\right) < \pi_{T=\mathbf{I}}$. When $w_2 = \sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}$, $w_1^2 = \frac{2K\bar{\lambda}}{K\bar{\lambda}+\underline{\lambda}}$, and $w_2^2 = \frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}$.

By Lemma A.2,

$$
\begin{aligned}
&\pi_{T=\mathbf{1}'}\left(\sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}\right) \\
&= E_0[\sum X_i] - 2\rho\left\{\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \Omega_{K\underline{\lambda},\bar{\lambda}} + \frac{1}{\rho^2\sigma^2}\left(\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \Omega_{K\underline{\lambda},\bar{\lambda}}\right)^2\right\}^{-1} \\
&= E_0[\sum X_i] \\
&\quad -2\rho\left\{\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \left[w_1^2\left(K\underline{\lambda}\right)^{-1} + w_2^2\bar{\lambda}^{-1}\right]^{-1}\right. \\
&\quad \left. + \frac{1}{\rho^2\sigma^2}\left(\frac{\sqrt{K\bar{\lambda}\underline{\lambda}}}{2w_1w_2} + \left[w_1^2\left(K\underline{\lambda}\right)^{-1} + w_2^2\bar{\lambda}^{-1}\right]^{-1}\right)^2\right\}^{-1} \\
&= E_0[\sum X_i] - 2\rho\left\{\frac{K\bar{\lambda}+\underline{\lambda}}{4} + \frac{K\bar{\lambda}+\underline{\lambda}}{2\left(\bar{\lambda}/\underline{\lambda}+\underline{\lambda}/\bar{\lambda}\right)}\right. \\
&\quad \left. + \frac{1}{\rho^2\sigma^2}\left(\frac{K\bar{\lambda}+\underline{\lambda}}{4} + \frac{K\bar{\lambda}+\underline{\lambda}}{2\left(\bar{\lambda}/\underline{\lambda}+\underline{\lambda}/\bar{\lambda}\right)}\right)^2\right\}^{-1},
\end{aligned}
$$

where $\Omega_{K\underline{\lambda},\bar{\lambda}}$ is defined by equation (9). The last equality is obtained by plugging in the values of w_1 and w_2. In addition, by Lemma A.1, $\pi_{T=\mathbf{I}} = E_0[\sum X_i] - 2\rho[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1}$.

Recall that the function $g(x) = E_0[\sum X_i] - 2\rho[x + \frac{1}{\rho^2\sigma^2}x^2]^{-1}$ is strictly increasing in $(0, +\infty)$. By the inequality of arithmetic and geometric means, $\bar{\lambda}/\underline{\lambda} + \underline{\lambda}/\bar{\lambda} > 2$, $\frac{K\bar{\lambda}+\underline{\lambda}}{4} + \frac{K\bar{\lambda}+\underline{\lambda}}{2(\bar{\lambda}/\underline{\lambda}+\underline{\lambda}/\bar{\lambda})} < \frac{K\bar{\lambda}+\underline{\lambda}}{4} + \frac{K\bar{\lambda}+\underline{\lambda}}{2\cdot 2} = \frac{K\bar{\lambda}+\underline{\lambda}}{2}$. Therefore, we have $\pi_{T=\mathbf{1}'}\left(\sqrt{\frac{2\underline{\lambda}}{K\bar{\lambda}+\underline{\lambda}}}\right) < \pi_{T=\mathbf{I}}$ as desired.

Now, since $0 < (w_2^*/w_1^*)^2 < \frac{\underline{\lambda}}{K\bar{\lambda}}$,

$$\pi_{T=\mathbf{1}'}\left(w_2^*\right) = E_0[\sum X_i] - 2\rho\left\{\Omega_{K\bar{\lambda},\underline{\lambda}} + \Omega_{K\underline{\lambda},\bar{\lambda}} + \frac{1}{\rho^2\sigma^2}(\Omega_{K\bar{\lambda},\underline{\lambda}} + \Omega_{K\underline{\lambda},\bar{\lambda}})^2\right\}^{-1}|_{w_2=w_2^*}$$

$$= \pi_{T=\mathbf{I}} = E_0[\sum X_i] - 2\rho[\frac{K\bar{\lambda}+\underline{\lambda}}{2} + \frac{1}{\rho^2\sigma^2}(\frac{K\bar{\lambda}+\underline{\lambda}}{2})^2]^{-1},$$

where $\Omega_{K\bar{\lambda},\underline{\lambda}}$ and $\Omega_{K\underline{\lambda},\bar{\lambda}}$ are defined by equation (9). This is equivalent to $\Omega_{K\bar{\lambda},\underline{\lambda}} + \Omega_{K\underline{\lambda},\bar{\lambda}} = \frac{K\bar{\lambda}+\underline{\lambda}}{2}$, which is exactly equation (4). □

Proof of Proposition 5.3. Mechanical manipulation of equation (4) yields $A\left(w_2^{*2}\right)^2 + Bw_2^{*2} + C = 0$, where $A = \frac{(KT+1)(KT-1)(K-T)}{2}$, $B = K^2T^3 - K^2T^2 + KT + K - 2T$, $C = 2T(1-K)$, and $T = \bar{\lambda}/\underline{\lambda}$. By assumption, $K > T > 1$. So $A > 0$ and $C < 0$. As $B^2 - 4AC > 0$, the equation $Ax^2 + Bx + C = 0$ must have two different real roots. In addition, the product of the two roots is $C/A < 0$. So w_2^{*2} is the unique positive root, $\frac{-B+\sqrt{B^2-4AC}}{2A}$. □

Proof of Proposition 5.6 is completely analogous to those of the corresponding propositions in Section 3 and is thus omitted here.

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2018.02.003.

References

Admati, A., 1985. A noisy rational expectations equilibrium for multi-asset securities markets. Econometrica 53 (3), 629–657.

Berger, P., Ofek, E., 1995. Diversification's effect on firm value. J. Financ. Econ. 37 (1), 39–65.

Bond, P., Goldstein, I., 2015. Government intervention and information aggregation by prices. J. Finance 70 (6), 2777–2812.

Comment, R., Jarrell, G., 1995. Corporate focus and stock returns. J. Financ. Econ. 37 (1), 67–87.

Dang, T.V., Gorton, G., Holmstrom, B., 2013. Ignorance, Debt and Financial Crises. Working Paper.

Demarzo, P., 2005. The pooling and tranching of securities: a model of informed intermediation. Rev. Financ. Stud. 18 (1), 1–35.

Demarzo, P., Duffie, D., 1999. A liquidity-based model of security design. Econometrica 67, 65–99.

Duffie, D., Rahi, R., 1995. Financial market innovation and security design. J. Econ. Theory 65 (1), 1–42.

Farhi, E., Tirole, J., 2015. Liquid bundles. J. Econ. Theory 158, 634–655.

Goldstein, I., Yang, L., 2015. Information diversity and complementarities in trading and information acquisition. J. Finance 70 (4), 1723–1765.

Gorton, G., Metrick, A., 2013. Securitization. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance, vol. 2. Elsevier. Chapter 1.

Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70 (3), 393–408.

Hoberg, G., Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: a text-based analysis. Rev. Financ. Stud. 23, 3773–3811.

Hoberg, G., Phillips, G., 2012. Industry Choice and Product Language. Working Paper.

Hund, J., Monk, D., Tice, S., 2010. Uncertainty about average profitability and the diversification discount. J. Financ. Econ. 96, 463–484.

John, K., Ofek, E., 1995. Asset sales and increase in focus. J. Financ. Econ. 37, 105–126.

Krishnaswami, S., Subramaniam, V., 1999. Information assymetry, valuation, and the corporate spin-off decision. J. Finance 53 (1), 73–112.

Lang, L., Stulz, R., 1994. Tobin's q, corporate diversification, and firm performance. J. Polit. Econ. 102 (6), 1248–1280.

Leland, H., 2007. Financial synergies and optimal scope of the firm: implications for mergers, spinoffs, and structured finance. J. Finance 62 (2), 765–807.

Maksimovic, V., Phillips, G., 2002. Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence. J. Finance 57 (2), 721–767.

Maksimovic, V., Phillips, G., 2007. Conglomerate Firms and Internal Capital Markets. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, vol. 1. Elsevier/North-Holland, Amsterdam. Chapter 8.

Maksimovic, V., Phillips, G., 2013. Conglomerate firms, internal capital markets, and the theory of the firm. Annu. Rev. Econ. 5, 225–244.

Marin, J., Rahi, R., 2000. Information revelation and market incompleteness. Rev. Econ. Stud. 67, 455–481.

Matsusaka, J., Nanda, V., 2002. Internal capital markets and corporate refocusing. J. Financ. Intermed. 11 (2), 176–211.

Mondria, J., 2010. Portfolio choice, attention allocation, and price comovement. J. Econ. Theory 145 (5), 1837–1864.

Parlour, C., Plantin, G., 2008. Loan sales and relationship banking. J. Finance 63 (3), 1291–1314.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. J. Financ. Econ. 80, 563–602.

Rajan, R., Servaes, H., Zingales, L., 2000. The cost of diversity: the diversification discount and inefficient investment. J. Finance 55, 35–80.

Shannon, C., 1948. Mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 and 623–656.

Sims, C., 2010. Rational Inattention and Monetary Economics. In: Friedman, B., Woodford, M. (Eds.), Handbook of Monetary Economics, vol. 3. Elsevier. Chapter 4.

Stein, J., 1997. Internal capital markets and the competition for corporate resources. J. Finance 52 (1), 111–133.

Subrahmanyam, A., 1991. A theory of trading in stock index futures. Rev. Financ. Stud. 4 (1), 17–51.

Townsend, R., 1979. Optimal contracts and competitive markets with costly state verification. J. Econ. Theory 21, 265–293.

Van Nieuwerburgh, S., Veldkamp, L., 2009. Information immobility and the home bias puzzle. J. Finance 64 (3), 1187–1215.

Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. Rev. Econ. Stud. 77 (2), 779–805.

Vijh, A., 2002. The positive announcement-period returns of equity carveouts: asymmetric information or divestiture gains? J. Bus. 75, 153–190.

Villalonga, B., 2004. Diversification discount or premium? New evidence from the business information tracking series. J. Finance 59, 479–506.

Yang, M., 2013. Optimality of Debt under Flexible Information Acquisition. Working Paper.