

# Specialization in Investor Information and the Diversification Discount

Ronaldo Carpio<sup>\*1</sup> and Meixin Guo<sup>†2</sup>

<sup>1</sup>*School of Banking & Finance, University of International Business & Economics*

<sup>2</sup>*Department of Economics, Tsinghua University*

July 28, 2018

## Abstract

We present a theory of the diversification discount based on investor specialization in information; this is modeled by different investors having a lower belief variance for different assets. We show that a discount exists in a general multi-asset, multi-investor setting; the discount increases with the degree of information specialization among investors. We empirically test our model using corporate spinoffs from the USA, using a novel measure of industry similarity based on sell-side analyst behavior. We find that higher abnormal returns on the spinoff announcement date are associated with higher industry dissimilarity between parent and child firms.

**Keywords:** Diversification discount; Heterogeneous beliefs

**JEL classification:** G14; G34

---

\*Corresponding author. 9th Floor, Boxue Bldg., 10 Huixindongjie, Beijing, 100029, China. Email: rncarpio@yahoo.com. Tel: +86 18811577940, Fax: +86 1064495059.

†Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China, 100084. Email: guomx@sem.tsinghua.edu.cn. Tel: (8610) 6279-5839.

# 1 Introduction

In corporate finance, the "diversification discount" refers to the empirical regularity that a diversified firm seems to be valued less than a portfolio of comparable single-segment firms. A closely related phenomenon is a corporate spin-off, in which one or more divisions are split off into a separate entity with own stock price, but is still owned by the shareholders of original firm. Managers and activist investors frequently justify spinoffs by arguing that components of a firm may be undervalued due to poor visibility; "pure play" stocks, on the other hand, are said to be rewarded by the market with higher valuations. For example, in 2014, the activist investor Carl Icahn invoked the diversification discount when arguing for the breakup of Ebay and PayPal (La Roche 2014):

"We believe that the separation of the traditional eBay and PayPal businesses will: (1) *highlight the significant value of the disparate businesses currently shrouded by a conglomerate discount the market has afforded eBay*; (2) focus and empower independent management teams to most effectively build two very different business platforms, make economic decisions independent of each other and, most importantly, foster innovation; and (3) provide an even more valuable currency for future bolt-on acquisition opportunities..."

Empirically, spin-offs seem to generate positive abnormal returns (Veld and Veld-Merkoulova 2009). We present a theoretical explanation for this phenomenon based on heterogeneity of investor beliefs; specifically, in the CARA-Gaussian framework, we show that heterogeneity among investors' belief covariance matrices can result in a diversification discount. We interpret this heterogeneity as *specialization in information*; investors differ in the market

sectors that they pay attention to, which is captured by a lower variance for specific assets in their belief covariance matrix. When different assets are combined in a single stock, all investors are forced to hold these assets in the same fixed ratio, regardless of each investor's ideal portfolio. This must be compensated for with a higher risk premium and hence lower prices. We first establish our result in a setting with exogenous beliefs; then, we show that our result is compatible with learning from prices in a noisy rational expectations equilibrium (REE) framework when expected market value is considered.

In contrast to other models of the diversification discount, our result depends only on heterogeneity of beliefs, and does not assume "noise traders", overconfidence, or bounded rationality by investors. In some models of the diversification discount, a spinoff always results in increased value for the firm; if taken to its logical conclusion, this would imply that firms should split into very many parts. In contrast, in our model there is a natural limit to the types of spinoffs that increase market value; a spinoff when there is no heterogeneity in investor specialization does not result in an increase in market value.

We provide a proof for the general multi-asset, multi-investor case with arbitrary belief variances among investors, by applying results from the analysis of positive definite matrices. This reassures us that the commonly considered two-asset, two-investor model is not a special case that can be reversed in higher dimensions, and that we can treat multi-firm spinoffs (which, though rare, exist in the data) similarly to one-firm spinoffs. We also show that the discount is minimized (and hence the incentive for a spinoff is maximized) when information specialization among investors is maximized in a certain sense. We then empirically test our model using data on US corporate spinoffs from 2001-2015. We develop a novel continuous measure of industry similarity based on the behavior of sell-side analysts; we propose that

industries that have more analysts in common should have investor communities that are less specialized, and vice versa. We find that our measure of industry similarity has a positive, strongly significant association with a spinoff's abnormal returns on its announcement date.

The rest of the paper is organized as follows: Section 2 reviews the theoretical literature and shows how investor specialization can generate a diversification discount in a simple two-investor, two-asset model. Section 3 extends this result to any number of investors and assets, and examines the configuration of investor beliefs that would result in the largest possible diversification discount. Section 4 presents our empirical tests of our model's predictions. Section 5 concludes.

## **2 A Two-Investor, Two-Asset Model**

### **2.1 Theoretical Literature**

The literature on the diversification or "conglomerate" discount originates with Lang and Stulz (1994) and Berger and Ofek (1995)'s findings that a conglomerate seems to have a lower market value than a comparable collection of single-segment firms. Maksimovic and Phillips (2013) provides a recent survey of this literature, including the possibility that previous empirical results may be the result of measurement artifacts. A corporate spinoff provides the opportunity to directly observe whether a discount exists, by comparing total market value before and after the spinoff.

A recent paper, Dai (2018), presents a model similar to ours in many respects; this paper builds on Van Nieuwerburgh and Veldkamp (2009)'s model of investors with limited attention

and analyzes the pricing implications of asset bundling in this setting. It examines three channels through which the total market value of a collection of assets may depend on whether they are bundled or not; one of these, the "trade-restriction channel", is essentially the same as the mechanism presented in our paper. Our model differs from Dai (2018) in the following ways: (i) that paper assumes a continuum of agents with independent private signals, as is standard in this literature, while we focus on the case with a finite number of agents. Under the continuum assumption, a premium<sup>1</sup> does not exist almost surely (see Section 3.2.1 for an elaboration). (ii) Also as is standard, it assumes all agents have a common prior belief about the mean of asset returns, which ensures that a premium cannot exist; and (iii) it assumes that agents's belief variances are simultaneously diagonalizable (as in Lemma 3.4), while our result Prop. 3.1 allows for arbitrary belief variances.

Our model's key assumption is that there are different investor groups, or clienteles, that specialize in specific sectors or industries. Bhandari (2013) presents a model based on Miller (1977)'s notion of optimists and pessimists who disagree about the prospects of a firm; a discount arises when different investor groups disagree about the prospects of different segments of a firm.

Another strand of the literature develops models in which the informativeness of stock prices about asset returns has an effect on firm valuation. Habib, Johnsen, and Naik (1995) present a two-asset model based on the Grossman and Stiglitz (1980) framework of informed and uninformed investors. When a firm is split into separate stocks, uninformed investors receive multiple price signals, which results in a reduced belief variance and hence a higher demand

---

<sup>1</sup>As Dai (2018) notes, a diversification premium (instead of a discount) is sometimes observed (Villalonga 2004), so a complete theory of diversification should be able to explain both a discount and a premium; the "trade-restriction" channel in that paper cannot generate a premium, but the "discipline" channel can.

for the stock, compared to when they receive a single price signal. Liu and Qi (2008) presents a model in which firms' managers rely on the informativeness of stock prices in order to direct investment; hence, more price signals increases the productivity of the firm. As is common in these types of models, noise in the asset supply ("noise traders") is necessary to ensure prices do not become fully revealing. However, these results depend on the specific assumptions made about the magnitude of asset supply noise before and after the split. Furthermore, if taken to its logical conclusion, these models imply that a firm should keep on splitting itself into more pieces, as long as the separate signals are not perfectly correlated (e.g. if the asset supply noises for the separate stocks are independent). In contrast, our model will generate an increase in market value only if the separate stocks correspond to existing specialization in the investor population.

Another theoretical approach is to assume investors are boundedly rational in some way. In Cao, Wang, and Zhang (2005), investors are heterogeneous in their degree of *uncertainty aversion*; in equilibrium, more uncertainty-averse investors do not invest at all when there are two separate stocks, but will invest in a single stock. In Hirshleifer and Teoh (2003), there are two types of investors: *attentive* investors, who process all available financial information on a firm, and *inattentive* investors who do not. Inattentive investors only pay attention to the average growth rate of a firm, which leads to undervaluation of a firm with a rapidly growing segment. Scheinkman and Xiong (2003) present a dynamic asset-pricing model in which two investors observe signals from two subsidiaries of the same firm; both investors are overconfident in that they believe their signal is more precise than it actually is.

Two game-theoretic models of the discount are presented in Chemmanur and Yan (2004), in which spinoffs are more likely to be taken over by more productive managers, and Nanda

and Narayanan (1999), in which a spinoff enables investors to distinguish between divisions of differing quality in a separating equilibrium.

Finally, the decision of whether to create a spinoff or not can be seen as a type of security design problem. There are several papers examining this problem from different viewpoints; Allen and Gale (1988) use risk-allocation considerations, while Boot and Thakor (1993) use informational considerations. Note that the security design literature generally tries to explain how to maximize investors' welfare, while this paper and the other papers mentioned above try to explain how to maximize the firm's total market value, which is not necessarily the same thing.

In the CARA-Gaussian framework, investor specialization occurs when different investors have a lower belief variance for different assets. Our approach is inspired by Van Nieuwerburgh and Veldkamp (2009)'s model of investors with limited attention, in which investors have a finite capacity to reduce the variance of signals they receive, and must prioritize which industries to learn more about. They develop a multi-asset noisy REE model where investors specialize in learning exclusively about a single asset type, and apply this to the home bias puzzle. Our paper does not explicitly incorporate a signal choice for investors, but we can interpret the exogenously given variances in our model as the result of such specialization.

## 2.2 The Model

We present a standard asset pricing model with finite heterogeneous agents, CARA utility, and Gaussian asset returns. Initially, we will assume that agents' beliefs are given exogenously; later, we will allow agents to learn from prices and private signals. There are  $I$

agents and two periods; agents trade in the first period and consume in the second. Each agent  $i = 1, \dots, I$  invests his initial wealth  $w_{0i}$  in a riskless asset and  $N$  risky assets. Agent  $i$  has population mass  $\lambda_i$ , where  $\lambda_i > 0$  and  $\sum_i \lambda_i = 1$ . The riskless rate of return is assumed to be 1. Let  $\tilde{F}$  denote the random vector of risky asset returns. Agent  $i$ 's final wealth is

$$\tilde{w}_{1i} = w_{0i}R + D_i^T(\tilde{F} - PR) \quad (2.1)$$

$P$  is the vector of market prices, and  $D_i(P)$  is the vector of holdings of risky assets. Each agent  $i$  has exponential utility  $u_i(w) = -\exp(w/\rho_i)$ . Agent  $i$ 's belief about the asset payoff  $\tilde{F}$  is Gaussian, with mean  $\mu^i$  and variance  $\Sigma_i = E_i[(\tilde{F} - \mu^i)^T(\tilde{F} - \mu^i)]$ . For now, we will take these parameters as exogenously given.

Given beliefs characterized by  $\mu^i$  and  $\Sigma_i$ , optimal demand is given by

$$D_i(P) = \frac{1}{\rho_i} \Sigma_i^{-1}(\mu^i - P) \quad (2.2)$$

The aggregate supply of assets is given by the vector  $Z$ , which we will assume is a nonrandom vector of ones for now. Let  $V_i = \frac{\rho_i}{\lambda_i} \Sigma_i$ ; this is the "effective" variance of agent  $i$ 's belief, incorporating the agent's risk tolerance and population mass. In the rest of the paper, we will simply refer to  $V_i$  as the variance. In equilibrium, the market clearing condition is

$$\sum_{i=1}^I \lambda_i D_i(P) = \sum_{i=1}^I V_i^{-1}(\mu^i - P) = Z \quad (2.3)$$

Equilibrium prices are given by:

$$P = \left(\sum_{i=1}^I V_i^{-1}\right)^{-1} \left(\sum_{i=1}^I V_i^{-1} \mu^i\right) - \left(\sum_{i=1}^I V_i^{-1}\right)^{-1} Z \quad (2.4)$$



The first term in this expression,  $(\sum_{i=1}^I V_i^{-1})^{-1}(\sum_{i=1}^I V_i^{-1}\mu^i)$  is equivalent to the Bayesian posterior mean that results from observing  $I$  normally distributed signals with realized values  $\mu^1, \dots, \mu^I$ , and where the  $i$ th signal is known to have a variance of  $V_i$ . The quantity  $(\sum_{i=1}^I V_i^{-1})^{-1}$  in the second term is the Bayesian posterior variance. Thus, we have:

**Remark 2.1** (Markets as a Bayesian aggregator of information). Equilibrium prices are identical to an economy with a representative agent who has "observed" the (scaled) beliefs of the  $I$  agents and combined them using Bayesian updating.

### 2.3 A Two-Asset Example

Suppose there are two assets  $n = 1, 2$ , and two agents  $i = 1, 2$ , each with the same size:  $\lambda_1 = \lambda_2 = \frac{1}{2}$ . We assume both types have the same risk aversion  $\rho_1 = \rho_2 = \rho$ . Investor  $i$ 's belief has mean  $\mu^i = (\mu_1^i, \mu_2^i)$  and variance  $\Sigma^i$ , where for  $\beta \in (0, 1)$ :

$$\Sigma_1 = \begin{bmatrix} \beta\sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \beta\sigma^2 \end{bmatrix}$$

We interpret this as specialization in information by agents: agent 1 has relatively more expertise about asset 1 than asset 2, and vice versa. This specialization may arise if agent 1 has chosen to observe signals about the industry of asset 1 but not asset 2 (and vice-versa for agent 2). We will compare two regimes: a "combined-firm" regime where the two assets are held by a single firm with a single stock, and a "separate-stocks" regime where each asset has its own stock. In each case, we will assume that the supply of all stocks is normalized to 1.

First, consider the combined-firm regime. We assume there are no production efficiencies or inefficiencies caused by combining the two assets; the return of the single stock is simply the sum of the returns of the individual assets. The belief means of each agent for the combined firm are  $\mu_1^1 + \mu_2^1$  and  $\mu_1^2 + \mu_2^2$ , respectively, and the variances of both agents are  $(1 + \beta)\sigma^2$ . The equilibrium price of the single stock is given by

$$P_{combined} = \left( \frac{1}{2\rho(1+\beta)\sigma^2} + \frac{1}{2\rho(1+\beta)\sigma^2} \right)^{-1} \left( \frac{1}{2\rho(1+\beta)\sigma^2}(\mu_1^1 + \mu_2^1) + \frac{1}{2\rho(1+\beta)\sigma^2}(\mu_1^2 + \mu_2^2) - 1 \right) \quad (2.5)$$

$$= \frac{\mu_1^1 + \mu_2^1 + \mu_1^2 + \mu_2^2}{2} - \rho(1+\beta)\sigma^2 \quad (2.6)$$

Total market value is given by  $P \cdot Z$ , which is simply  $P_{combined}$ .

Now, consider the separate-stocks regime. The price vector is given by

$$P = \left( \frac{1}{2\rho}\Sigma_1^{-1} + \frac{1}{2\rho}\Sigma_2^{-1} \right)^{-1} \left( \frac{1}{2\rho}\Sigma_1^{-1} \begin{bmatrix} \mu_1^1 \\ \mu_2^1 \end{bmatrix} + \frac{1}{2\rho}\Sigma_2^{-1} \begin{bmatrix} \mu_1^2 \\ \mu_2^2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \quad (2.7)$$

$$= \frac{1}{1+\beta} \begin{bmatrix} \mu_1^1 + \beta\mu_1^2 - 2\beta\rho\sigma^2 \\ \beta\mu_1^1 + \mu_2^2 - 2\beta\rho\sigma^2 \end{bmatrix} \quad (2.8)$$

Combined market value is given by  $P \cdot Z = p_1 + p_2$ , which is equal to:

$$P_{separate} = \frac{\mu_1^1 + \beta\mu_1^2 + \beta\mu_1^1 + \mu_2^2 - 4\rho\beta\sigma^2}{1+\beta} \quad (2.9)$$

The difference between the two market values is:

$$P_{separate} - P_{combined} = \frac{(1-\beta)(\mu_1^2 - \mu_1^1 + \mu_2^1 - \mu_2^2 + 2(1-\beta)\rho\sigma^2)}{2(1+\beta)} \quad (2.10)$$

In general, if there are no restrictions on  $\mu^1$  and  $\mu^2$ , the discount may be positive or negative.

### 2.3.1 Equal Belief Means

Consider the case where both agents have the same belief mean:  $\mu^1 = \mu^2 = (\mu_1, \mu_2)$ . Then we have

$$P_{combined} = (\mu_1 + \mu_2) - \rho(1 + \beta)\sigma^2 \quad (2.11)$$

$$P_{separate} = (\mu_1 + \mu_2) - \rho \frac{4\beta\sigma^2}{1 + \beta} \quad (2.12)$$

$$P_{separate} - P_{combined} = \rho \frac{(1 - \beta)^2\sigma^2}{1 + \beta} \quad (2.13)$$

This expression is positive for  $\beta \in (0, 1)$ ; its derivative with respect to  $\beta$  is  $\rho\sigma^2 \left(1 - \frac{4}{(1+\beta)^2}\right)$ , which is negative. The diversification discount increases with  $\sigma^2$  and with a greater degree of specialization (that is, a lower  $\beta$ ) among agents.

Suppose  $\mu_1 = \mu_2 = \bar{\mu}$ . We can express the discount (as a fraction of  $P_{combined}$ ) as a function of  $\frac{\sigma^2}{\bar{\mu}}$ :

$$\frac{P_{separate} - P_{combined}}{P_{combined}} = \left( \frac{1 + \beta}{(1 - \beta)^2} \left( \frac{2}{\rho \frac{\sigma^2}{\bar{\mu}}} - (1 + \beta) \right) \right)^{-1} \quad (2.14)$$

Figure 1 plots the discount fraction for the parameter values  $\rho = 1, \frac{\sigma^2}{\bar{\mu}} = \{0.1, 0.25, 0.5\}$ . When the variance of payoffs is high compared to its mean and the degree of specialization is high, the discount can be economically significant.

Intuitively, this phenomenon can be understood as the result of a restriction of the investor's

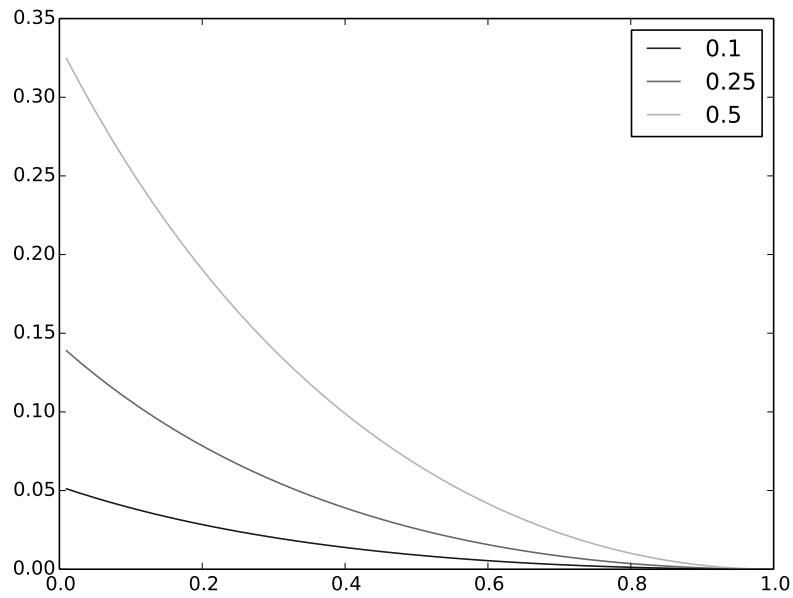


Figure 1: Discount as a fraction of the combined price for  $\rho = 1$ ,  $\frac{\sigma^2}{\mu} = \{0.1, 0.25, 0.5\}$

choice set. Under the single-firm regime, the ratio of asset 1 to asset 2 in every investor's portfolio is fixed, which prevents investors from achieving their ideal portfolio. In equilibrium, investors will demand a higher risk premium (and hence lower prices) to compensate.<sup>2</sup> In the separate-stocks regime, agent  $i$ 's demand for each asset is given by:

$$D_i(P) = \frac{1}{\rho_i} \Sigma_i^{-1} (\mu^i - P) = \begin{bmatrix} D_{i,1} \\ D_{i,2} \end{bmatrix} = \frac{1}{\rho_i} \begin{bmatrix} (\beta\sigma^2)^{-1}(\mu_1^i - P_1) \\ (\sigma^2)^{-1}(\mu_2^i - P_2) \end{bmatrix} \quad (2.15)$$

Suppose  $\mu_1^i = \mu_2^i$  and  $P_1 = P_2$ ; then the ratio  $D_{i,1}/D_{i,2}$  is equal to  $\beta$  (each agent has a greater demand for the asset in which he has more expertise); in the single-firm regime, both investors must hold assets 1 and 2 in a 1:1 ratio.

### 3 General Multi-Investor, Multi-Asset Model

We generalize this result to any number of assets and agent types, with agent beliefs that can have arbitrary correlations between asset returns, including non-diagonal covariance matrices. In order to do this, we will introduce tools from the analysis of positive definite matrices; while this increases the technical burden for the reader, it allows us to give results that hold for any number of arbitrary covariance matrices.

---

<sup>2</sup>This is essentially the "trade-restriction" channel of Dai (2018).

### 3.1 Preliminaries

For any matrix  $A$ , let  $\mathbf{sum}(A)$  denote the sum of all elements of  $A$ . Suppose we have an  $N$ -dimensional random vector  $x = (x_1, \dots, x_N)^T$  with covariance matrix  $\Sigma$ ; then the variance of  $x_1 + \dots + x_N$  is given by  $\mathbf{sum}(\Sigma)$ . Suppose an agent's belief covariance matrix for the separate-stocks regime is  $V$ ; then, this agent's belief about returns for the combined firm must have variance  $\mathbf{sum}(V)$ . If  $V$  is diagonal, this is also equal to the *trace* of  $V$ , denoted  $\mathbf{tr}(V)$ . We will utilize the following definitions and results on positive definite matrices (proofs are given in the Appendix). In what follows,  $A_1, \dots, A_n, A'_1, \dots, A'_n, A$ , and  $B$  are arbitrary real-valued, symmetric, positive definite matrices of the same dimension.

**Definition 3.1.** We define the Loewner partial ordering:  $A \geq_L (>_L) B$  whenever  $A - B$  is positive semidefinite (positive definite).

The Loewner ordering has the following statistical interpretation (Horn (1990), p.141): suppose  $\tilde{X}, \tilde{Y}$  are  $\mathbb{R}^n$ -valued random variables, with  $Var(\tilde{X}) = A$  and  $Var(\tilde{Y}) = B$ . Then  $A \geq_L (>_L) B$  iff for any nonzero  $c \in \mathbb{R}^n$ ,  $Var(c \cdot \tilde{X}) \geq (>) Var(c \cdot \tilde{Y})$ . It is a natural way to order variances based on conditioning on events: if  $\tilde{X}, \tilde{Y}$  are Gaussian, then  $Var(\tilde{X}) \geq_L Var(\tilde{X}|\tilde{Y})$ , i.e. uncertainty must weakly decrease after observing new information.

**Definition 3.2.** The *parallel sum* of  $A_1, \dots, A_n$  is denoted  $A_1 : \dots : A_n = (A_1^{-1} + \dots + A_n^{-1})^{-1}$ . The *harmonic mean* of  $A_1, \dots, A_n$  is  $n(A_1 : \dots : A_n)$ .

The Bayesian posterior after observing normally distributed signals with realized values  $\mu^1, \dots, \mu^I$  and known variances  $V_1, \dots, V_I$  has mean  $(V_1 : \dots : V_I)(\sum_{i=1}^I V_i^{-1} \mu^i)$  and variance  $V_1 : \dots : V_I$ .

**Lemma 3.1.** (*monotonicity and joint concavity of parallel sum*):  $A_1 : \dots : A_n$  is monotoni-

cally increasing in  $A_1, \dots, A_n$ :

$$A_1 \geq_L A'_1, \dots, A_n \geq_L A'_n \Rightarrow A_1 : \dots : A_n \geq_L A'_1 : \dots : A'_n \quad (3.1)$$

and jointly concave in  $A_1, \dots, A_n$ . For  $t \in [0, 1]$ :

$$(tA_1 + (1-t)A'_1) : \dots : (tA_n + (1-t)A'_n) \geq_L t(A_1 : \dots : A_n) + (1-t)(A'_1 : \dots : A'_n) \quad (3.2)$$

**Lemma 3.2.** (inequality between parallel sum of sum/trace and sum/trace of parallel sum):

1.  $\mathbf{sum}(A_1 : \dots : A_n) \leq \mathbf{sum}(A_1) : \dots : \mathbf{sum}(A_n)$ .
2.  $\mathbf{tr}(A_1 : \dots : A_n) \leq \mathbf{tr}(A_1) : \dots : \mathbf{tr}(A_n)$ .

**Lemma 3.3.** (joint concavity of sum/trace of parallel sum):  $\mathbf{sum}(A_1 : \dots : A_n)$  and  $\mathbf{tr}(A_1 : \dots : A_n)$  are jointly concave over  $A_1, \dots, A_n$ .

**Lemma 3.4.** (equality of sum/trace of parallel sum): Suppose  $A_1, \dots, A_n$  are simultaneously diagonalizable: there exists an orthornormal  $P$  such that for each  $i = 1, \dots, n$ ,  $A_i = P^T D_i P$ , where  $D_i$  is the diagonal matrix containing the eigenvalues of  $A_i$ , and the rows of  $P$  are the shared eigenvectors of  $A_1, \dots, A_n$ . If  $D_i = \alpha_i D$  where  $\alpha_i \neq 0$  for  $i = 1, \dots, n$ , that is, each  $D_i$  matrix is some matrix  $D$  multiplied by a nonzero scalar, then  $\mathbf{tr}(A_1 : \dots : A_n) = \mathbf{tr}(A_1) : \dots : \mathbf{tr}(A_n)$  and  $\mathbf{sum}(A_1 : \dots : A_n) = \mathbf{sum}(A_1) : \dots : \mathbf{sum}(A_n)$ .

The economic interpretation of the conditions of Lemma 3.4 is that all agents agree on what the statistically independent risk factors affecting asset returns are (the eigenvectors), though they may disagree on how much each risk factor contributes to a specific asset's return (the eigenvalues). Furthermore, all agents agree on the relative contribution of each risk factor to every asset, though they may disagree on the absolute contribution.

Let  $I$ , the number of agent types, and  $N$ , the number of assets, be positive integers.

**Assumption 3.1.**

1. The risk-free interest rate  $R$  is 1.
2. The vector of asset supplies  $Z$  is a vector of ones.
3. Every agent's belief covariance matrix  $V_i$  is positive definite.

In the combined-firm regime, the equilibrium price is given by

$$P_{combined} = (\mathbf{sum}(V_1) : \dots : \mathbf{sum}(V_I)) \left( \sum_i^I \mathbf{sum}(V_i)^{-1} \mathbf{sum}(\mu^i) - 1 \right) \quad (3.3)$$

In the separate-stocks regime, the equilibrium price vector is given by

$$\begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} = \left( \sum_{i=1}^I V_i^{-1} \right)^{-1} \left( \sum_{i=1}^I V_i^{-1} \mu^i \right) - \left( \sum_{i=1}^I V_i^{-1} \right)^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (3.4)$$

The combined value of all stocks is given by

$$P_{separate} = \sum_{n=1}^N p_n = \mathbf{sum} \left( (V_1 : \dots : V_I) \left( \sum_{i=1}^I V_i^{-1} \mu^i \right) \right) - \mathbf{sum}(V_1 : \dots : V_I) \quad (3.5)$$

Now, we can compare prices in the two regimes under different assumptions about agent beliefs.

**Proposition 3.1.** (*diversification discount under equal belief means*): *Suppose all agents have the same belief means:  $\mu^i = \mu$  for  $i = 1, \dots, I$ ; then  $P_{combined} \leq P_{separate}$ .*



*Proof.*

$$P_{combined} = \mathbf{sum}(\mu) - [\mathbf{sum}(V_1) : \dots : \mathbf{sum}(V_I)] \quad (3.6)$$

$$P_{separate} = \mathbf{sum}(\mu) - \mathbf{sum}(V_1 : \dots : V_I) \quad (3.7)$$

By Lemma 3.2,  $\mathbf{sum}(V_1 : \dots : V_I) \leq \mathbf{sum}(V_1) : \dots : \mathbf{sum}(V_I)$ , therefore  $P_{combined} \leq P_{separate}$ . □

**Proposition 3.2.** *(no discount with common eigenvectors and eigenvalues are equal up to scale factor): Suppose  $V_1, \dots, V_i$  satisfy the conditions in Lemma 3.4 (common eigenvectors and eigenvalues are equal up to a scale factor). Then  $P_{combined} = P_{separate}$ .*

Under the equal belief mean assumption, there is never a diversification premium. A zero discount holds if agents have homogeneous beliefs, but can also occur if all agents agree on the independent risk factors and their relative contributions to asset prices. Changes in the distribution of risk tolerance or population mass among agent types, therefore, do not lead to a discount. This also ensures that if a spinoff occurs when there is no heterogeneity of beliefs, there will be no diversification discount and hence no increase in value for the firm. This is in contrast to models where the informativeness of stock prices always increases with a spinoff.

If both belief means and variances are arbitrary, a discount, premium, or neither may exist. To see this, consider the equilibrium price equation 2.4 with  $I = 2$ . The first term is a Bayesian posterior mean after observing two signals (or alternatively, updating a prior with one signal); in the combined-firm regime and separate-stocks regime, it is given by the first term in Equations 3.3 and 3.5, respectively. It is known that in Bayesian inference with

Gaussian signals, if only the prior and observation locations are known while the variances are free variables, then the posterior mean may lie virtually anywhere (Chamberlain and Leamer (1976), Theorem 1). Therefore, the element sum of the posterior mean in the separate-stocks regime may be arbitrarily higher or lower than in the combined-firm regime, as was demonstrated by Eqn 2.10 for the 2-agent, 2-asset case.

### 3.2 Noisy REE Model

We have seen that a discount exists if belief means are equal, but may or may not exist if belief means are allowed to vary arbitrarily. In this section we consider the multi-asset, noisy REE model of Admati (1985). Noisy REE models, beginning with Grossman and Stiglitz (1980) and Hellwig (1980), relax the strong assumption of equal belief means in a structured way, by assuming agents share a common prior belief about (unobserved) asset returns, then endowing agents with a private signal that is correlated with asset returns. Furthermore, agents are allowed to learn from prices, which have dual roles of clearing markets while also serving as a signal that reveals some part of other agents' private information. Hellwig (1980) developed a multi-agent REE model with one risky asset; Admati (1985) extended this to the case with multiple risky assets.

Assume two time periods; at  $t=0$ , the economy receives a Gaussian shock  $\tilde{y} \sim N(\bar{y}, \Sigma_y)$ , which includes the asset return realization  $\tilde{F}$ , signals observed by agents, and asset supply noise. At  $t=1$ , agents update their beliefs, trade, and reach equilibrium. We assume an equilibrium exists, and at equilibrium:

- agent  $i$ 's updated belief mean is a fixed linear function of  $\tilde{y}$ :  $\tilde{\mu}^i = \gamma^i(\tilde{y})$ , where  $\gamma^i(\cdot)$  is

linear;

- agent  $i$ 's updated belief variance is a deterministic constant  $W_i$ ;
- the asset supply vector is a fixed linear function of  $\tilde{y}$ :  $\tilde{Z} = Z(\tilde{y})$ , with  $E[Z(\tilde{y})]$  being a vector of ones.

Agents' demand vectors and the market clearing condition then become linear functions of  $\tilde{y}$ , given by substituting  $(\gamma^i(\tilde{y}), Z(\tilde{y}), W_i)$  for  $(\mu^i, Z, V_i)$  in Eqns. 2.2 and 2.4. Since Eqns. 2.4, 3.3 and 3.5 are linear in  $(\mu^i, Z)$  when  $V_i$  is taken as a constant, we can take time-0 expectations w.r.t.  $\tilde{y}$ :

$$E[PR] = \left( \sum_{i=1}^I W_i^{-1} \right)^{-1} \left( \sum_{i=1}^I W_i^{-1} E[\gamma^i(\tilde{y})] \right) - \left( \sum_{i=1}^I W_i^{-1} \right)^{-1} E[Z(\tilde{y})] \quad (3.8)$$

$$E[P_{combined}] = (\mathbf{sum}(W_1) : \dots : \mathbf{sum}(W_I)) \left( \sum_i^I \mathbf{sum}(W_i)^{-1} \mathbf{sum}(E[\gamma^i(\tilde{y})]) \right) - (\mathbf{sum}(W_1) : \dots : \mathbf{sum}(W_I)) \quad (3.9)$$

$$E[P_{separate}] = \mathbf{sum} \left( (W_1 : \dots : W_I) \left( \sum_{i=1}^I W_i^{-1} E[\gamma^i(\tilde{y})] \right) \right) - \mathbf{sum}(W_1 : \dots : W_I) \quad (3.10)$$

If  $E[\gamma^i(\tilde{y})]$  is the same for  $i = 1, \dots, I$ , then the proof of Proposition 3.1 goes through, applied to  $E[P_{combined}]$  and  $E[P_{separate}]$ .

**Proposition 3.3.** (*diversification discount in expectation*) *Suppose that the economy is driven by a Gaussian shock  $\tilde{y}$ , and a linear equilibrium exists where:*

- agent  $i$ 's belief mean is linear in  $\tilde{y}$ :  $\tilde{\mu}^i = \gamma^i(\tilde{y})$ , where  $\gamma^i(\cdot)$  is linear
- agent  $i$ 's belief variance is a deterministic constant  $W_i$ ;
- the asset supply  $\tilde{Z} = Z(\tilde{y})$  is linear in  $\tilde{y}$ , with  $E[Z(\tilde{y})]$  equal to a vector of ones

- the market price is linear in  $\tilde{y}$ , satisfying the condition  $P(\tilde{y}) = (\sum_{i=1}^I W_i^{-1})^{-1} (\sum_{i=1}^I W_i^{-1} \gamma^i(\tilde{y})) - (\sum_{i=1}^I W_i^{-1})^{-1} Z(\tilde{y})$

Then  $E[P_{combined}] \leq E[P_{separate}]$ . If  $W_1, \dots, W_I$  have common eigenvectors and their eigenvalues are equal up to a scale factor, then equality holds.

If a discount holds in expectation, then a risk-neutral firm would optimally choose to break itself up into the separate-stocks regime rather than stay in the combined-firm regime. Note that we do not make any assumption on how equilibrium beliefs are reached, so long as belief means are linear in  $\tilde{y}$  (and therefore Gaussian). We show that this result holds in the noisy REE model of Admati (1985), which we briefly summarize below. Let  $\tilde{F}$  denote the unobserved vector of asset returns, and let  $\tilde{Z}$  denote the asset supply; both are Gaussian random variables with means  $\bar{F}, \bar{Z}$  respectively. Each agent  $i$  receives a private signal  $\tilde{Y}_i = \tilde{F} + \tilde{\epsilon}_i$ , where  $\tilde{\epsilon}_i$  is iid zero-mean Gaussian noise. Thus, the Gaussian shock  $\tilde{y}$  is given by  $\tilde{y} = (\tilde{F}, \tilde{Z}, \tilde{\epsilon}_1, \dots, \tilde{\epsilon}_I)$ . If we assume that a linear equilibrium pricing function exists, then  $\tilde{P}$  is a linear function of  $\tilde{y}$ , and  $(\tilde{F}, \tilde{P}, \tilde{Z}, \tilde{Y}_1, \dots, \tilde{Y}_I)$  are jointly Gaussian with a deterministic variance matrix. In a noisy REE, each agent is assumed to know the true joint distribution of  $(\tilde{F}, \tilde{Y}_i, \tilde{P})$ ; after observing his private signal  $\tilde{Y}_i$  and the equilibrium price  $\tilde{P}$ , his belief is the conditional distribution of  $\tilde{F}|\tilde{y}_i, \tilde{P}$ , which is also Gaussian. Agent  $i$ 's updated belief mean is  $\tilde{\mu}^i = E[\tilde{F}|\tilde{y}_i, \tilde{P}]$ , which is a linear function of  $\tilde{y}$ ; the belief variance is  $Var(\tilde{F}|\tilde{y}_i, \tilde{P})$ , which is a deterministic constant. Applying the law of iterated expectations, we get  $E[\tilde{\mu}^i] = E[E[\tilde{F}|\tilde{y}_i, \tilde{P}]] = E[\tilde{F}] = \bar{F}$ ; thus, the expectation of all agents' belief means are equal. Therefore, the conditions of Proposition 3.3 are satisfied, and a diversification discount weakly holds for expected market value. Note, however, that we cannot say anything about the relative size of the discount when compared to the setting without learning from

prices.

### 3.2.1 Infinite vs. Finite Agents

The REE literature typically assumes that there are infinitely many agents, each of which observes an i.i.d. Gaussian private signal. It is further assumed that a law of large numbers applies, and the *realized* average signal across agents is then assumed to be equal to the *expectation* of the signal distribution, a.s. This makes an analytic solution tractable, but also guarantees that the conditions for Lemma 3.1 hold a.s. To see this, suppose that there are a continuum of agents, indexed by  $i \in [0, 1]$ . After observing their private signal, each agent's belief mean is an i.i.d. Gaussian random variable:  $\tilde{\mu}^i \sim N(\bar{\mu}, \Sigma_\mu)$ . If we assume that the realized average belief across the agent population is equal to its expectation, a.s., then the first term in Eq 2.4 becomes (Admati (1985), Eq 16):

$$\left( \int_0^1 V_i^{-1} \right)^{-1} \int_0^1 V_i^{-1} \tilde{\mu}^i di = \left( \int_0^1 V_i^{-1} \right)^{-1} \int_0^1 V_i^{-1} \bar{\mu} = \bar{\mu} \quad \text{a.s.} \quad (3.11)$$

This term is equal in the combined and separate stocks regime a.s., therefore Lemma 3.1 holds a.s. The assumption of infinite or finite agents results in different behavior at equilibrium; with infinite agents, a diversification premium will not occur as a result of agents' realized beliefs a.s., but may occur with finite agents.<sup>3</sup>

---

<sup>3</sup>Hellwig (1980) proved existence of a linear REE in the single-asset case with infinite and finite agents; Admati (1985) only proved existence in the multi-asset case with infinite agents. Existence of equilibrium with finite agents was proven in Carpio and Guo (2017).

### 3.3 Beliefs That Maximize Combined Market Value

We can ask the question: what distribution of belief variances  $V_1, \dots, V_I$  among agents will result in a larger discount (and hence a larger incentive to create a spinoff)? The diversification discount is larger when  $\mathbf{sum}(V_1 : \dots : V_I)$  is smaller. The joint concavity of  $\mathbf{sum}(V_1 : \dots : V_I)$  (by Lemma 3.3) has two immediate implications. First, the discount is smaller when  $V_1, \dots, V_I$  is a convex combination of other possible collections of variances. Second, a concave function over a convex region is minimized (and therefore the discount is maximized) at an extreme point (i.e. a point that is not a convex combination of other points); if we can formulate an appropriate convex feasible region, we can find the variances that maximize the discount.

### 3.4 Simultaneously diagonalizable variances

A tractable case is to assume all agents' variances have the same eigenvectors: there exists an orthonormal  $P$  such that  $V_i = P^T D_i P$ , where  $D_i$  is a diagonal matrix; the rows of  $P$  are the shared eigenvectors and the diagonal entries of  $D_i$  are the associated eigenvalues of  $V_i$ . As before, we interpret this to mean that all agents agree on the independent risk factors underlying asset returns. The next results show that we can reduce the feasible set to the set of diagonal positive definite matrices.

**Lemma 3.5.** *(parallel sum for simultaneously diagonalizable matrices) Suppose  $A_1, \dots, A_n$  are positive definite, and  $A_i = P^T D_i P$  for  $i = 1, \dots, n$ , where  $P$  is orthogonal and  $D_1, \dots, D_n$  are diagonal. Then*

$$A_1 : \dots : A_n = P^T (D_1 : \dots : D_n) P \tag{3.12}$$

**Lemma 3.6.** (*joint concavity for simultaneously diagonalizable matrices*) Let  $D_1, \dots, D_n, D'_1, \dots, D'_n$  be diagonal, positive definite matrices, and let  $P$  be an orthogonal matrix of the same dimension. For  $t \in [0, 1]$ :

$$\begin{aligned} P^T(tD_1 + (1-t)D'_1)P : \dots : P^T(tD_n + (1-t)D'_n)P \geq \\ tP^T(D_1 : \dots : D_n)P + (1-t)P^T(D'_1 : \dots : D'_n)P \end{aligned} \quad (3.13)$$

and  $\mathbf{sum}(P^T D_1 P : \dots : P^T D_n P)$  is jointly concave over  $D_1, \dots, D_n$ .

Therefore, we will assume that each  $V_i$  is diagonal, and we will seek an extreme point of a suitably defined convex subset of the positive definite diagonal matrices. Suppose that all agents have the same beliefs about the returns of the combined firm, which implies that the trace of each  $V_i$  is equal. We can set up an additive constraint on the diagonal entries by setting a lower bound on each diagonal entry of  $V_i$ ; we interpret this as setting a minimum amount of uncertainty for each asset. Let  $[V]_n$  denote the  $n$ -th diagonal entry of  $V$ .

Suppose  $\mathbf{tr}(V_i) = N\underline{\sigma}^2 + T$  for  $i = 1, \dots, I$ , where  $\underline{\sigma}^2 \geq 0, T \geq 0$ , and  $[V_i]_n \geq \underline{\sigma}^2$  for  $i = 1, \dots, I, n = 1, \dots, N$ . We say that  $V$  is *specialized in variance* in asset  $n$  if all diagonal entries of  $V$  are equal to  $\underline{\sigma}^2$ , except for  $[V]_n$  which is equal to  $\underline{\sigma}^2 + T$ .

**Proposition 3.4.** (*maximization of dividend discount under trace constraint*): *The dividend discount will be maximized when each  $V_i$  is specialized in variance for some asset, and each asset has at least  $\lfloor N/I \rfloor$   $V_i$ s that specialize in it, and no asset has more than  $\lfloor N/I \rfloor + 1$ .*

These results imply that the incentive to generate spinoffs depends on the degree of specialization among the investor community. When there is less overlap of information among investor groups, the discount is larger.

### 3.5 Discussion and Empirical Implications

We have shown that a diversification discount exists when different investor types have more precise information about different assets, and that the discount is maximized (and therefore the incentive for a spinoff is largest) when the population of investors "specializes" in the sense that there is minimum overlap in the types of assets that each investor type knows more about. The results in this paper make predictions that can be tested empirically; we list some of these predictions and how they might be tested.

First, spinoff and merger patterns are affected by the interests of the investor/analyst community. If investors form into groups focusing on a specific sector, industry, country, etc, then a firm splitting up is more likely when it contains divisions that align with these groupings (and conversely, a merger is less likely). Furthermore, spinoffs whose businesses are aligned with this grouping will provide greater abnormal returns than those that do not. On the other hand, if there is no investor/analyst group following a particular sector, then there will be little incentive to spin off business segments in this sector, even if the segment is relatively independent of the rest of the firm.

Second, spinoffs of unrelated divisions should be more likely, and should return a higher abnormal return, since divisions in unrelated industries or market sectors will likely be the focus of different investor groups. Most other theoretical models of the diversification discount also make this prediction; one possible way to distinguish the predictions of this model from others is by comparing returns of spinoffs of unrelated divisions when there is little investor attention in a particular sector, to when it is high.

Third, in contrast with some other models of spinoffs, the incentive to split comes solely



from investor behavior, with no necessary connection to management behavior or even firm productivity. Thus, our model predicts spinoffs of divisions or assets that seem to be currently adequately managed, if there is some investor group that would pay attention to it. For example, in recent years it has become popular for companies to spin off their real estate holdings into REITs, under the hypothesis that these assets were undervalued when held in the original firm. Furthermore, spinoffs may generate abnormal returns, even if there is no improvement in the actual productivity of the spinoff.

Fourth, if there is a time trend in the amount of investor interest in a specific sector, or if there is a trend towards more investor specialization in general, then there should be a corresponding trend in spinoffs. Interest in some specific sectors has varied over time (for example; Internet stocks, "social media" stocks, China stocks, etc). We can try to measure the degree of investor or analyst specialization into distinct sectors, and compare with the number and valuation of spinoffs in these sectors over time. Also, it seems that specialization of investors in general has been steadily increasing over time, due to improvements in information technology and financial sophistication. We could check if there has been a corresponding rise in spinoffs and valuations of all types over time.

In the next section, we empirically test our model's prediction that the diversification discount for a multi-segment firm is related to the degree of specialization among the investor groups following that firm.

## 4 Data and Empirical Methodology

Our model predicts a larger diversification discount when a firm is composed of two or more segments, each followed by its own investor group, and different investor groups are more specialized in their information (i.e. they pay more attention to their specific industry, and less attention to other industries). In this section we present an empirical test of this prediction, using the behavior of sell-side analysts as a way to measure investor specialization. We collect data on US corporate spinoffs from 2001-2015, and compute the abnormal return associated with each spinoff. Then, we regress abnormal return on measures of similarity between industry categories. These measures include simple dummy variables indicating when a parent and child firm are in different industries, but also a novel continuous measure that is constructed from patterns of analyst reports. The idea is that if two industries have many analysts in common, then there is little specialization between the investor communities of these two industries, since the same group of people will be observing signals from both industries. However, if two industries have few analysts in common, then specialization between the investor communities is high, since signals from one industry will not be observed by analysts following the other industry.

The basic timeline of a spinoff is as follows: first is the announcement date (*AD*), which is generally the earliest date that the firm's board of directors publicly announces that they intend to spin off a segment of the firm into a separate, publicly traded company. If the firm proceeds with the spinoff decision, then some time after the announcement date (usually a period of several months), after regulatory approval has been obtained, there is a *declaration date* on which the parent firm publicly announces that it will distribute stock of the child company to the parent's shareholders, to take place on the *ex-date* (*ED*). The

ex-date is usually the first date on which the child firm's stock is publicly traded. In this paper we will focus on abnormal returns over four periods: 1)  $(AD - 1, AD + 1)$ , called the announcement date (AnncDate) effect; 2)  $(AD - 1, ED - 1)$ , called the AnncDate-ExDate effect; 3)  $(ED - 1, ED)$ , called the ExDate effect; and 4)  $(ED - 1, ED + 30)$ , called the ExDate 30-day effect.

## 4.1 Empirical Literature

Previous empirical papers examining the excess return from spinoffs include Chemmanur, Krishnan, and Nandy (2014), Bhandari (2013), Burch and Nanda (2003), Krishnaswami and Subramaniam (1999), Daley, Mehotra, and Sivakumar (1997), and Vijh (1994). Veld and Veld-Merkoulova (2009) review 26 studies of spinoffs; they find that there is a consistent and large positive effect on the announcement date. However, there is a mixed record for long-run effects after the announcement date. Krishnaswami and Subramaniam (1999) and Daley, Mehotra, and Sivakumar (1997) find that "cross-industry" spinoffs, where the child has a different 2-digit SIC code than its parent, have a significantly higher AnncDate effect; Desai and Jain (1999) find higher AnncDate and ExDate effects for "focus-increasing" spinoffs, one measure of which is a different SIC code between parent and child. Bhandari (2013) finds a larger ExDate effect for spinoffs where there is more disagreement among shareholders about the prospects of the child firm, as measured by institutional holdings before and after the spinoff.

## 4.2 Data Sources

For data on US corporate spinoffs, we search the CRSP database for distributions with a distribution code of 3763 or 3764 (indicating a tax-free equity distribution) between 2001 and 2015. For each observation, we do a manual search of news sources and company press releases to obtain the announcement date; we also note whether the spinoff was simultaneous with a merger or a combination of assets from other firms. From CRSP's STOCKNAMES database, we obtain the PERMCO and SIC code of both the parent and child firm. If the SIC code is missing or has a value of 9999 (indicating an unknown industry), we next turn to Compustat, and finally to a manual search of news and SEC filings. The SIC code then determines the firm's Fama-French 48-industry (henceforth FF48) classification (Fama and French (1997)); we choose the 48-industry scheme because it is the most fine-grained.

We exclude the following cases from our spinoff dataset:

- if we cannot obtain the child's SIC code from CRSP, Compustat, or a manual search.
- if the spinoff is simultaneous with a merger or combination with assets of other firms, as determined by manual searching of news and press releases.
- if the parent and child firm have the same PERMCO. Thus, we exclude issues of tracking stocks or new share classes.
- if the first trading date of the child firm's PERMCO is more than 10 days before the ex-date, or after the ex-date. This will exclude cases where the child firm was already publicly traded, and hence not a true spinoff.
- if there is more than one parent PERMCO associated with the spinoff. This excludes

cases where multiple parent firms distribute stock of the same child firm on the same date.

- we allow observations with multiple spinoffs (more than one child PERMCO); however, each of the child firms must satisfy the conditions above. The regressors are calculated as a weighted average over each child firm, with the child’s market value relative to the total market value of all child firms on the ex-date as the weight.<sup>4</sup>

In contrast with some of the literature, we do not exclude foreign firms or REITs.

Daily returns for firms with multiple PERMNOs are computed as the market value-weighted average over all PERMNOs associated with the firm’s PERMCO. We note the parent PERMCO’s total market value on the day before the ex-date, and the child PERMCO’s total market value on the first trading day where it is nonzero. This will be used to calculate the parent/child market value ratio of the spinoff.

Data on analyst forecasts from 2001-2016 comes from the IBES detail file. We attempt to link CRSP PERMNO to IBES ticker, if possible, and use that to determine the SIC and FF48 code for each firm in the IBES data.

### **4.3 An analyst-based measure of industry similarity**

Previous papers on spinoffs have used a simple dummy variable (e.g. same 2-digit or 3-digit SIC code) to indicate whether a child firm is in a different industry from its parent; in this paper we go beyond a binary indicator to construct a novel continuous measure

---

<sup>4</sup>Our dataset has four observations of multiple spinoffs: IAC Interactive in 2008, Cendant in 2006, Tyco in 2007, and Temple Inland, Inc. in 2007. Excluding these observations has little effect on our results.

of distance between industry categories (FF48 classifications), based on patterns of sell-side analyst behavior. The idea of using analyst behavior to classify firms was used in Kaustia and Rantala (2015) to identify "peer firms" of a company; the hope is that this can be more accurate than relying on pre-determined categories, since analysts are able to continually evaluate firms' actual behavior in real time. In the corporate diversification literature, Hoberg and Phillips (2010) constructed a continuous measure of similarity between firms using text-based analysis of product descriptions in 10-K statements. This measure of similarity can be seen as reflecting the company's point of view, while our measure is oriented towards the beliefs of investors.

Our measure is an application of a graph-theoretic method used in network and link analysis (Fouss, Pirotte, Renders, and Saerens (2007), Fouss, Saerens, and Shimbo (2016) Ch. 2)<sup>5</sup> Suppose there are  $I$  firms and  $N$  analysts in the economy, and the set of firms  $1, \dots, I$  is partitioned into  $J$  disjoint sets,  $S_1, \dots, S_J$ , each representing an "industry" classification. We say that an analyst  $a \in 1, \dots, N$  "follows" firm  $i$  if  $a$  issues at least one forecast for firm  $i$ , and "follows" an industry  $j \in 1, \dots, J$  if  $a$  issues at least one forecast for a firm in industry  $j$ . Let  $A_j$  denote the set of all analysts who follow industry  $j$ ; we will call this the "analyst community" of industry  $j$ . We construct a graph that encodes the relationships between industries and analysts; then, we compute a measure of similarity between industry nodes. The nodes in our graph are the industries  $1, \dots, J$  and analysts  $1, \dots, N$ . We add an edge between industry  $j$  and analyst  $a$  if  $a$  follows  $j$ ; the weight of the edge is the number of firms in industry  $j$  that analyst  $a$  follows. Then, we compute the *commute-time distance* (henceforth CTD) between each pair of nodes  $i, j$  in the graph; this is defined as the average

---

<sup>5</sup>For example, given a database of movies, people, and information on who has watched what, a set of recommended movies for a given person  $i$  can be constructed as the set of unwatched movies that are most similar to  $i$ , in some sense.

number of steps that a random walker, starting in node  $i$ , takes until it enters  $j$  for the first time, and returns back to  $i$ .<sup>6</sup> This provides a measure of similarity between any node in the graph; the advantage of this method is that it uses global information, and not just information local to  $i$  or  $j$ . We extract the pairwise distances between industry nodes  $1...J$ ; a lower distance indicates a higher degree of similarity. See Appendix 6.1 for an example of calculating these similarity measures.

Consider two analyst communities  $A_1, A_2$  that follow industries  $I_1, I_2$ . If  $A_1$  and  $A_2$  have significant overlap (i.e. there are many analysts who follow both  $I_1$  and  $I_2$ ), then it is more likely that a signal about the future prospects of  $I_1$  will be observed by analysts in  $A_2$ , and vice versa; thus,  $A_1$  and  $A_2$  would have a low degree of information specialization, in the framework of our model. In the graph encoding relationships between industries and analysts, industries with many analysts in common will have many paths connecting them, since each "follow" by an individual analyst  $a$  of industry  $I$  results in an edge connecting nodes  $a$  and  $I$ . The CTD measures the "bandwidth" of the total set of connections, direct and indirect, between two nodes in a graph. If we assume that more connections between two industries via common analysts is correlated with more common signals being observed by, and more information being transmitted between, investor groups that follow those two industries, then the CTD is one way to measure the information specialization between investor groups.

We calculate this distance between each pair of FF48 industry categories (except code 48) using IBES analyst data for each year from 2001 to 2015. Firms with an erroneous or missing SIC code will be coded as 48, which is the "Other" category used for firms that don't fit into

---

<sup>6</sup>The commute-time distance can also be thought of as the resistance between two nodes in an electrical network, where each edge has a resistance inversely proportional to its weight.

any other categorization; this results in a lot of unrelated firms ending up in this category. To avoid erroneously grouping these firms together, we drop IBES analyst reports for firms classified as code 48. It may be possible for the CTD between two nodes to be infinite, if they are unconnected in the graph representation. If this occurs, we replace it with two times the largest CTD value for that year. Finally, we transform the distance into its z-score, calculated using all spinoff observations, to facilitate interpretation.

## 4.4 Empirical Methodology

Our main regression specification is

$$R_i = \alpha + \beta_1 d_i + \beta_2 Year_i + \beta_3 X_i + e_i \quad (4.1)$$

where  $R_i$  is the abnormal return,  $d_i$  is a measure of dissimilarity between parent and child industries,  $year_i$  is the year of the ex-date, and  $X$  is a vector of control variables. We include year as a regressor to test the hypothesis that spinoff effects have been decreasing over time. Abnormal returns are calculated using a market model estimated over a 155-day period ending 45 days before the announcement date; the CRSP value-weighted index is used for the market portfolio.<sup>7</sup>

---

<sup>7</sup>Let  $R_{it}$  denote the daily simple return of security  $i$  at time  $t$ . We estimate the market model with the equation

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}$$

The abnormal return at time  $t$  is estimated as  $\hat{\epsilon}_{it} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt}$ , and the cumulative abnormal over a period  $t = 1, \dots, T$  is the  $\sum \hat{\epsilon}_{it}$ . Alternative methods for computing the abnormal return include: (i) estimating the market model with log returns instead of simple returns; (ii) calculating buy-and-hold returns rather than cumulative abnormal returns. We found very little difference between these methods for the short-horizon AnncDate and ExDate effects, but larger differences for the longer-horizon ExDate 30-day effect and especially the AnncDate-ExDate effect, which may have a horizon longer than a year.



For measures of industry dissimilarity, we use: dummy variables indicating that parent and child have the same 2-digit SIC, 3-digit SIC, or same FF48, and our continuous distance measure presented above. Following Bhandari (2013), for the control variables, we include a dummy indicating whether the parent was in the S&P 500. We calculate the excess share turnover for the parent firm during the observed trading period relative to a reference period, which is either  $(AD - 90, AD - 31)$  for the AnncDate and AnncDate-ExDate effects, or  $(ED + 31, ED + 90)$  for the ExDate and ExDate 30-day effects. If the parent firm's PERMCO has multiple PERMNOs, we calculate the value-weighted turnover using all the parent firm's PERMNOs. We do not include variables related to productivity or investment; since we are only examining returns over a very short time period, these factors should not come into play. Following Bhandari (2013), we also examine a subset of spinoffs where the parent/child market value ratio is less than 25 and 10; we would expect any spinoff effects to be larger when the child firm is relatively larger in market value compared to the parent.

## 4.5 Empirical Results

Tables 1, 2, and 3 show summary statistics for the entire sample of spinoffs, and for two restricted samples where the parent/child market value ratio is limited to be at most 25 and 10, respectively. Consistent with previous studies, there is a positive mean AnncDate and ExDate effect. Our AnncDate and ExDate effects are 2.73% and 1.54%, respectively, which are somewhat less than reported in previous studies (for comparison, Vijh (1994) reports effects of 2.90% and 3.03%, while Bhandari (2013) reports effects of 3.28% and 2.38%, respectively). As expected, when the sample is restricted to spinoffs with relatively larger child firms, the average spinoff effects become larger. The AnncDate-ExDate effect

goes from 3.45% in the full sample, to 5.49% in the most restricted sample. The ExDate 30-day effect goes from negative and close to zero in full sample, to 0.542% (about one third of the ExDate effect) in the most restricted sample. Note, however, that our estimates of abnormal return become more sensitive to how the market model is estimated and how holding period returns are calculated, as the holding period increases. Therefore, we devote the most attention to the AnncDate and ExDate effects.

Table 4 reports regression results for the AnncDate effect ( $AD - 1$  to  $AD + 1$ ). Consistent with previous studies, there is a positive, significant abnormal return on the announcement date; however, in contrast to those results, this effect is not associated with different SIC or FF48 codes between parent and child. Our CTD measure is significant at the 1% level in all samples; a 1 standard deviation increase in CTD is associated with 0.785 to 1.0 more percentage points of abnormal return, with larger coefficients found in the more restricted samples. Year has no significant effect, indicating that in our data at least, there does not seem to be any reduction in the AnncDate effect over time.

Table 5 reports results for AnncDate-ExDate effect ( $AD - 1$  to  $ED - 1$ ). None of our regressors are significant.

Table 6 reports results for the ExDate effect ( $ED - 1$  to  $ED$ ). Once again, consistent with previous studies, the effect is positive and significant, although somewhat smaller in magnitude. Year is negative and significant at the 1% level in all specifications, indicating that the ExDate effect has clearly decreased over time.

Table 7 reports results for the ExDate 30-day effect ( $ED - 1$  to  $ED + 30$ ). Year is significant at the 10% level in the most restricted sample; none of our industry similarity measures are significant.

## 4.6 Discussion

Our regression results for the `AnncDate` and `ExDate` effects are broadly consistent with previous studies, finding a positive and significant mean effect. Our measure of industry similarity has the expected sign and is significant at the 1% level for the `AnncDate` effect. The coefficient estimates of our similarity measure is economically significant, with a 1 standard deviation change in the similarity measure associated with 0.785–1.0 more percentage points in abnormal return, which is a large fraction of the total spinoff effect. We can conceive of a spinoff effect that require separate stocks to be actually traded, and an effect that does not. For example, if an investor believes that an announced spinoff will likely be completed, and is willing to wait until the spinoff is completed, then he can buy the parent stock on the announcement date, resulting in the `AnncDate` effect. On the other hand, if an investor is interested in only one of the parent or child stocks, and is not willing to buy the parent stock before the spinoff takes place (perhaps due to transaction costs, or institutional investor "style" guidelines), then this will result in an effect that only manifests at the ex-date (see Bhandari (2013) for an example of one such mechanism). Our measure appears to capture an aspect of industry similarity that investors act upon when a spinoff is announced. Finally, as a separate result, we find that spinoff effects on or after the ex-date have been decreasing over time. This may be due to decreasing transaction costs, since transaction costs make it more costly to buy the parent stock before the ex-date, and then sell one of the parent or child stocks after the spinoff is completed.

## 5 Conclusion

We have shown that in the CARA-Gaussian framework, a diversification discount can arise with a specific type of heterogeneity in beliefs: when different investors have more precise information about different assets. We interpret this as specialization in information about specific sectors or industries. In noisy REE models with learning from prices, an expected discount exists. The discount (and therefore the incentive for a spinoff) increases with the degree of specialization among the population of investors. Our result depends only on heterogeneity of investor beliefs, and does not require "noise traders" or bounded rationality by investors. Furthermore, we have derived results in a general multi-investor, multi-asset setting. To empirically test our model, we develop a novel continuous measure of industry similarity based on sell-side analyst behavior, using the assumption that if two industries have many analysts in common, then the investor groups following those two industries are likely to be less specialized in information. We find that our measure of industry similarity has a positive, strongly significant association with a spinoff's abnormal returns on its announcement date, providing empirical support for our theory.

For future research, we can ask what kind of financial structure will maximize the total market value for the firm, given its collection of assets and given the beliefs in the population. We can also examine the effect of investor heterogeneity in risk aversion and in the means of beliefs about asset returns.

Table 1: Summary Statistics for all spinoffs

	count	mean	std	min	25%	50%	75%	max
$AD - 1$ to $AD + 1$	178	2.731	5.799	-28.042	-0.239	1.531	5.165	25.299
$AD - 1$ to $ED - 1$	177	3.458	43.509	-180.763	-14.904	2.150	17.351	274.544
$ED - 1$ to $ED$	188	1.543	4.190	-12.017	-0.639	1.086	2.950	32.270
$ED - 1$ to $ED + 30$	188	-0.001	13.150	-48.091	-6.411	0.138	5.922	60.083
Parent/child market value ratio	188	7.330	12.129	0.000	1.315	3.002	7.966	83.267
Same 2-digit SIC	188	0.465	0.499	0	0	0	1	1
Same 3-digit SIC	188	0.338	0.473	0	0	0	1	1
Same FF48	188	0.504	0.499	0	0	0.746995	1	1
Commute-time distance	185	262.930	609.868	0	0	0.000	317.585	6036.994
Parent in SP500	188	0.356	0.480	0	0	0	1	1
Parent's average turnover %, $AD - 90$ to $AD - 31$	180	0.008	0.005	0.000	0.004	0.007	0.009	0.033
Parent's average turnover %, $ED + 31$ to $ED + 90$	185	0.009	0.007	0.000	0.005	0.008	0.011	0.045
Parent AnncDate Excess Turnover Ratio	180	2.935	2.764	0.013	1.078	2.016	3.643	17.867
Parent AnncDate-ExDate Excess Turnover Ratio	174	1.192	0.452	0.338	0.896	1.101	1.421	3.137
Parent ExDate Excess Turnover Ratio	183	2.955	2.500	0.400	1.382	2.089	3.766	18.095
Parent ExDate 10-day Excess Turnover Ratio	183	1.670	0.944	0.544	1.113	1.449	2.002	8.713
Parent ExDate 30-day Excess Turnover Ratio	183	1.360	0.540	0.408	1.031	1.253	1.608	4.704

This table reports descriptive statistics for the full sample of spinoffs.

Table 2: Summary Statistics for spinoffs with parent/child market value ratio  $\leq 25$

	count	mean	std	min	25%	50%	75%	max
<i>AD</i> - 1 to <i>AD</i> + 1	170	2.792	5.916	-28.042	-0.256	1.609	5.203	25.299
<i>AD</i> - 1 to <i>ED</i> - 1	169	3.848	44.441	-180.763	-14.904	2.677	19.704	274.544
<i>ED</i> - 1 to <i>ED</i>	180	1.570	4.262	-12.017	-0.639	1.129	2.968	32.270
<i>ED</i> - 1 to <i>ED</i> + 30	180	0.070	13.321	-48.091	-6.330	0.269	5.922	60.083
Parent/child market value ratio	180	5.289	6.052	0.000	1.251	2.832	6.808	24.755
Same 2-digit SIC	180	0.464	0.499	0	0	0	1	1
Same 3-digit SIC	180	0.331	0.470	0	0	0	1	1
Same FF48	180	0.499	0.499	0	0	0.3969563	1	1
Commute-time distance	177	265.713	620.137	0	0	0.000	317.585	6036.994
Parent in SP500	180	0.339	0.475	0	0	0	1	1
Parent's average turnover %, <i>AD</i> - 90 to <i>AD</i> - 31	172	0.008	0.005	0.000	0.004	0.007	0.010	0.033
Parent's average turnover %, <i>ED</i> + 31 to <i>ED</i> + 90	177	0.010	0.007	0.000	0.005	0.008	0.012	0.045
Parent AnncDate Excess Turnover Ratio	172	3.003	2.803	0.013	1.150	2.117	3.810	17.867
Parent AnncDate-ExDate Excess Turnover Ratio	166	1.194	0.459	0.338	0.891	1.103	1.430	3.137
Parent ExDate Excess Turnover Ratio	175	3.023	2.531	0.400	1.435	2.106	3.831	18.095
Parent ExDate 10-day Excess Turnover Ratio	175	1.688	0.960	0.544	1.113	1.459	2.015	8.713
Parent ExDate 30-day Excess Turnover Ratio	175	1.368	0.550	0.408	1.028	1.266	1.617	4.704

This table reports descriptive statistics for the sample of spinoffs where the parent/child market value is less than 25.

Table 3: Summary Statistics for spinoffs with parent/child market value ratio  $\leq 10$

	count	mean	std	min	25%	50%	75%	max
$AD - 1$ to $AD + 1$	147	3.149	6.271	-28.042	-0.195	2.000	6.207	25.299
$AD - 1$ to $ED - 1$	146	5.498	46.762	-180.763	-14.048	3.944	21.812	274.544
$ED - 1$ to $ED$	156	1.760	4.497	-12.017	-0.582	1.267	3.425	32.270
$ED - 1$ to $ED + 30$	156	0.542	13.346	-48.091	-6.009	0.704	6.603	60.083
Parent/child market value ratio	156	3.219	2.714	0.000	1.178	2.327	4.886	9.998
Same 2-digit SIC	156	0.478	0.499	0	0	0	1	1
Same 3-digit SIC	156	0.336	0.472	0	0	0	1	1
Same FF48	156	0.505	0.499	0	0	0.7469947	1	1
Commute-time distance	154	251.182	606.813	0	0	0.000	317.407	6036.994
Parent in SP500	156	0.288	0.455	0	0	0	1	1
Parent's average turnover %, $AD - 90$ to $AD - 31$	149	0.008	0.006	0.000	0.004	0.007	0.010	0.033
Parent's average turnover %, $ED + 31$ to $ED + 90$	153	0.010	0.007	0.000	0.005	0.008	0.012	0.045
Parent AnncDate Excess Turnover Ratio	149	3.221	2.924	0.013	1.184	2.262	4.326	17.867
Parent AnncDate-ExDate Excess Turnover Ratio	143	1.219	0.482	0.338	0.889	1.112	1.463	3.137
Parent ExDate Excess Turnover Ratio	151	3.290	2.618	0.400	1.528	2.406	4.034	18.095
Parent ExDate 10-day Excess Turnover Ratio	151	1.783	0.993	0.544	1.201	1.536	2.062	8.713
Parent ExDate 30-day Excess Turnover Ratio	151	1.423	0.566	0.408	1.055	1.318	1.665	4.704

This table reports descriptive statistics for the sample of spinoffs where the parent/child market value is less than 10.

Table 4: AnncDate effect ( $AD - 1$  to  $AD + 1$ )

	All spinoffs				Mktval ratio $j= 25$	Mktval ratio $j= 10$
Year	0.049 (0.114)	0.049 (0.114)	0.043 (0.111)	0.091 (0.119)	0.114 (0.129)	0.114 (0.149)
Parent in SP500	-1.837 ** (0.814)	-1.839 ** (0.807)	-1.860 ** (0.825)	-2.017 ** (0.840)	-2.246 ** (0.900)	-2.278 ** (1.108)
Parent Excess Turnover Ratio	0.642 ** (0.265)	0.635 ** (0.266)	0.632 ** (0.264)	0.639 ** (0.264)	0.640 ** (0.263)	0.635 ** (0.270)
Same 2-digit SIC	0.388 (0.750)					
Same 3-digit SIC		0.132 (0.783)				
Same FF48			-0.746 (0.793)			
Commute-time dis- tance z-score				0.785 *** (0.301)	0.820 *** (0.304)	1.000 *** (0.289)
N	178	178	178	175	167	145
Adjusted $R^2$	0.091	0.090	0.094	0.109	0.109	0.098

Robust standard deviations are shown in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively. Parent Excess Turnover Ratio is the ratio of the average value-weighted turnover in the given time period, to the average value-weighted turnover in a reference period of  $AD - 90$  to  $AD - 31$ .



Table 5: AnncDate-ExDate effect ( $AD - 1$  to  $ED - 1$ )

	All spinoffs				Mktval ratio $\leq 25$	Mktval ratio $\leq 10$
Year	-0.718 (0.785)	-0.692 (0.783)	-0.733 (0.788)	-0.578 (0.779)	-0.606 (0.838)	-0.663 (0.944)
Parent in SP500	-1.816 (6.075)	-1.133 (5.977)	-1.880 (6.098)	-2.605 (6.199)	-2.264 (6.554)	0.952 (7.325)
Parent Excess Turnover Ratio	-5.005 (6.592)	-4.673 (6.653)	-5.086 (6.591)	-5.193 (6.795)	-5.541 (6.941)	-5.795 (7.283)
Same 2-digit SIC	0.140 (6.363)					
Same 3-digit SIC		6.972 (6.967)				
Same FF48			-2.054 (6.499)			
Commute-time dis- tance z-score				2.290 (3.024)	2.132 (3.108)	3.305 (2.910)
N	174	174	174	172	164	142
Adjusted $R^2$	-0.016	-0.009	-0.015	-0.012	-0.014	-0.013

Robust standard deviations are shown in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively. Parent Excess Turnover Ratio is the ratio of the average value-weighted turnover in the given time period, to the average value-weighted turnover in a reference period of  $AD - 90$  to  $AD - 31$ .

Table 6: ExDate effect ( $ED - 1$  to  $ED$ )

	All spinoffs				Mktval ratio $\leq 25$	Mktval ratio $\leq 10$
Year	-0.257 *** (0.072)	-0.257 *** (0.072)	-0.259 *** (0.072)	-0.251 *** (0.073)	-0.269 *** (0.077)	-0.283 *** (0.087)
Parent in SP500	-0.537 (0.558)	-0.532 (0.553)	-0.552 (0.554)	-0.553 (0.553)	-0.527 (0.566)	-0.709 (0.643)
Parent Excess Turnover Ratio	0.228 * (0.126)	0.236 * (0.126)	0.220 * (0.126)	0.227 * (0.126)	0.217 * (0.127)	0.185 (0.130)
Same 2-digit SIC	-0.525 (0.600)					
Same 3-digit SIC	-0.403 (0.618)					
Same FF48	-0.907 (0.605)					
Commute-time dis- tance z-score					0.085 (0.211)	0.090 (0.223)
N	183	183	183	180	172	149
Adjusted $R^2$	0.087	0.085	0.095	0.081	0.086	0.083

Robust standard deviations are shown in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively. Parent Excess Turnover Ratio is the ratio of the average value-weighted turnover in the given time period, to the average value-weighted turnover in a reference period of  $ED + 31$  to  $ED + 90$ .

Table 7: ExDate 30-day effect ( $ED - 1$  to  $ED + 30$ )

	All spinoffs				Mktval ratio $\leq 25$	Mktval ratio $\leq 10$
Year	-0.325 (0.244)	-0.324 (0.244)	-0.323 (0.244)	-0.319 (0.249)	-0.385 (0.266)	-0.459 (0.292)
Parent in SP500	3.085 (2.033)	3.123 (1.987)	3.028 (2.029)	2.954 (2.051)	3.322 (2.183)	3.757 (2.399)
Parent Excess Turnover Ratio	3.283 ** (1.481)	3.191 ** (1.495)	3.221 ** (1.495)	3.148 ** (1.551)	3.321 ** (1.569)	2.663 (1.634)
Same 2-digit SIC	0.988 (1.922)					
Same 3-digit SIC		1.312 (2.049)				
Same FF48			0.219 (1.900)			
Commute-time distance z-score				0.023 (1.116)	0.055 (1.130)	0.912 * (0.539)
N	183	183	183	180	172	149
Adjusted $R^2$	0.014	0.015	0.013	0.010	0.016	0.024

Robust standard deviations are shown in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively. Parent Excess Turnover Ratio is the ratio of the average value-weighted turnover in the given time period, to the average value-weighted turnover in a reference period of  $ED + 31$  to  $ED + 90$ .

## 6 Appendix

### 6.1 Example of Analyst-Based Industry Similarity Measure

Suppose there are two industries  $I_1, I_2$  and four firms  $f_1, f_2, f_3, f_4$ , with  $I_1 = \{f_1, f_2\}$  and  $I_2 = \{f_3, f_4\}$ , and three analysts  $a_1, a_2, a_3$ . Analyst  $a_1, a_2, a_3$  each follow two firms:  $\{f_1, f_2\}$ ,  $\{f_2, f_3\}$ , and  $\{f_3, f_4\}$  respectively. Then, the "analyst community" following industries  $I_1, I_2$  are  $A_1 = \{a_1, a_2\}$  and  $A_2 = \{a_2, a_3\}$  respectively. Both  $A_1$  and  $A_2$  have a total of four firms "followed", aggregated over all analysts in each "analyst community". This market is shown in Figure 2.

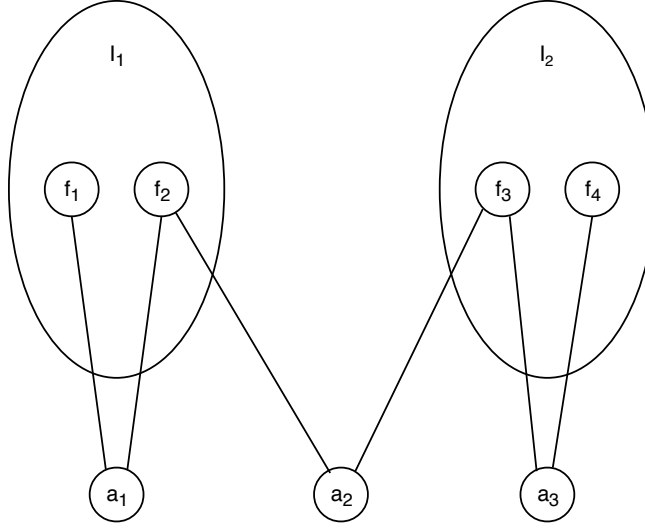


Figure 2: Example of industries, firms, and analysts.

Now, consider the graph representation of the same market, shown in Figure 3. There are two "industry" nodes  $I_1, I_2$  and three "analyst" nodes,  $a_1, a_2, a_3$ . Edges  $(I_1, a_1)$  and  $(I_2, a_3)$  have a weight of 2; edges  $(I_1, a_2)$  and  $(I_2, a_2)$  have a weight of 1. Let  $L$  denote the Laplacian matrix of this graph, and  $L^+$  denote the Moore-Penrose pseudoinverse of  $L$ ; then the CTD between nodes  $i, j$  is given by Eq. 5 in Fouss, Pirotte, Renders, and Saerens (2007):

$$n(i, j) = V_G(e_i - e_j)^T L^+(e_i - e_j) \quad (6.1)$$

where  $V_G$ , the "volume" of the graph, is the sum of the degrees of all nodes, and  $e_i$  is the  $i$ th column of the identity matrix  $I$  with a dimension equal to the total number of nodes. For this example, let the nodes of the graph be arranged as follows:  $I_1, a_1, a_2, I_2, a_3$ . Then the matrix of CTDs is given in Table 8. The distance between nodes  $I_1, I_2$  is 24; the distance between nodes  $a_1, a_3$  is 36.

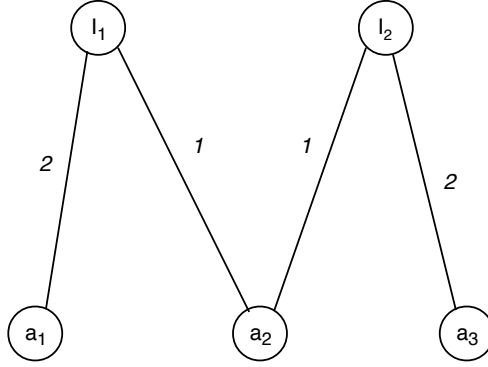


Figure 3: Graph for the market shown in Figure 2

0	6	12	24	30
6	0	18	30	36
12	18	0	12	18
24	30	12	0	6
30	36	18	6	0

Table 8: Commute-time distances for the graph in Figure 3. The order of the nodes is:  $l_1$ ,  $a_1$ ,  $a_2$ ,  $l_2$ ,  $a_3$ .

## 6.2 Proofs

Lemma 3.1 (monotonicity and joint concavity).

*Proof.* The case  $n = 2$  is proved in Bhatia (2007), Theorem 4.1.1. We will prove the cases  $n > 2$  by induction. Note that parallel sum is associative:  $A_1 : \dots : A_k : A_{k+1} = (A_1 : \dots : A_k) : A_{k+1}$ . Suppose that monotonicity holds for  $n = k$ , and  $A_1 \geq_L A'_1, \dots, A_{k+1} \geq_L A'_{k+1}$ . Then

$$(A_1 : \dots : A_k) : A_{k+1} \geq_L (A'_1 : \dots : A'_k) : A'_{k+1} \Rightarrow \quad (6.2)$$

$$A_1 : \dots : A_k : A_{k+1} \geq_L A'_1 : \dots : A'_k : A'_{k+1} \quad (6.3)$$

Suppose that joint concavity holds for  $n = k$ . Then for  $t \in [0, 1]$ ,

$$(tA_1 + (1-t)A'_1) : \dots : (tA_k + (1-t)A'_k) : (tA_{k+1} + (1-t)A'_{k+1}) \quad (6.4)$$

$$\geq_L (t(A_1 : \dots : A_k) + (1-t)(A'_1 : \dots : A'_k)) : (tA_{k+1} + (1-t)A'_{k+1}) \quad (6.5)$$

$$\geq_L t(A_1 : \dots : A_k) : A_{k+1} + (1-t)(A'_1 : \dots : A'_k) : A'_{k+1} \quad (6.6)$$

$$= t(A_1 : \dots : A_{k+1}) + (1-t)(A'_1 : \dots : A'_{k+1}) \quad (6.7)$$

Therefore, monotonicity and joint concavity hold for all  $n \geq 2$ .  $\square$

**Definition 6.1.** Let  $\mathbb{M}_n$  denote the set of all  $n \times n$  real-valued matrices. A linear map  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_k$  is *positive* if  $\Phi(A)$  is positive semidefinite whenever  $A$  is positive semidefinite.

**Lemma 6.1.** (*inequality of positive linear map of parallel sum*): Let  $\Phi$  be any positive linear map on  $\mathbb{M}_n$ , and let  $A_1, \dots, A_n$  be positive semidefinite. Then  $\Phi(A_1 : \dots : A_n) \leq_L \Phi(A_1) : \dots : \Phi(A_n)$ .

*Proof.* The case  $n = 2$  is proved in Bhatia (2007), Theorem 4.1.5. Suppose it holds for  $n = k$ . Then

$$\Phi(A_1) : \dots : \Phi(A_k) \geq_L \Phi(A_1 : \dots : A_k) \quad (6.8)$$

$$(\Phi(A_1) : \dots : \Phi(A_k)) : \Phi(A_{k+1}) \geq_L \Phi(A_1 : \dots : A_k) : \Phi(A_{k+1}) \quad (6.9)$$

$$\Phi(A_1) : \dots : \Phi(A_k) : \Phi(A_{k+1}) \geq_L \Phi(A_1 : \dots : A_k) : \Phi(A_{k+1}) \quad (6.10)$$

by associativity, and the right hand side is  $\geq_L \Phi((A_1 : \dots : A_k) : A_{k+1}) = \Phi(A_1 : \dots : A_k : A_{k+1})$ .

Therefore, it holds for all  $n \geq 2$ .  $\square$

Lemma 3.2 (inequality for sum/trace).

*Proof.* **sum** and **tr** are positive linear maps (Bhatia (2007) Example 2.2.1). Therefore, Lemma 6.1

applies. □

Lemma 3.3 (joint concavity of sum/trace of parallel sum).

*Proof.* First, we show that **sum** and **tr** are monotone. For **tr**, see Bernstein (2009), Corollary 8.4.10. For **sum**, let  $e$  denote a column vector of ones;  $\mathbf{sum}(A) = e^T A e$ . By the definition of positive semidefinite,  $A \geq_L B \Rightarrow e^T(A - B)e \geq 0 \Rightarrow e^T A e \geq e^T B e \Rightarrow \mathbf{sum}(A) \geq \mathbf{sum}(B)$ .

By joint concavity of the parallel sum, for  $t \in [0, 1]$ :

$$(tA_1 + (1-t)A'_1) : \dots : (tA_n + (1-t)A'_n) \geq_L t(A_1 : \dots : A_n) + (1-t)(A'_1 : \dots : A'_n) \quad (6.11)$$

$$\mathbf{tr}((tA_1 + (1-t)A'_1) : \dots : (tA_n + (1-t)A'_n)) \geq_L \mathbf{tr}(t(A_1 : \dots : A_n) + (1-t)(A'_1 : \dots : A'_n)) \quad (6.12)$$

$$= t\mathbf{tr}(A_1 : \dots : A_n) + (1-t)\mathbf{tr}(A'_1 : \dots : A'_n) \quad (6.13)$$

The proof is identical for **sum**. □

**Lemma 6.2** (variance of sum of variables). *Suppose  $\tilde{X} = (\tilde{x}_1, \dots, \tilde{x}_n)$  is a  $n$ -dimensional random vector with variance  $A$ , diagonalized as  $A = P^T D P$ , with  $P$  an orthornormal matrix whose rows are the eigenvectors of  $A$  and  $D$  a diagonal matrix containing the eigenvalues of  $A$ . Then  $\mathbf{sum}(A) = \text{Var}(\sum_i \tilde{x}_i) = \sum_i (\sum_j P_{i,j})^2 D_{i,i}$ .*

*Proof.* Let  $\tilde{Y} = (\tilde{y}_1, \dots, \tilde{y}_n) \sim N(0, I_n)$ . Then  $\tilde{X} = P^{-1} D^{\frac{1}{2}} \tilde{Y}$  and  $\tilde{x}_i = \sum_{j=1}^n P_{j,i} (D_{j,j})^{\frac{1}{2}} \tilde{y}_j$ , and  $\mathbf{sum}(A) = \text{Var}(\sum_i \tilde{x}_i) = \sum_i (\sum_j P_{i,j})^2 D_{i,i}$ . □

Lemma 3.4 (equality of sum/trace of parallel sum).

*Proof.* For **tr**, applying Lemma 3.5:

$$\mathbf{tr}(A_1 : \dots : A_n) = \mathbf{tr}(P^T D_1 P : \dots : P^T D_n P) = \mathbf{tr}(P^T D_1 : \dots : D_n P) \quad (6.14)$$

$$= \mathbf{tr}(D_1 : \dots : D_n P P^T) = \mathbf{tr}(\alpha_1 D : \dots : \alpha_n D) = (\alpha_1 : \dots : \alpha_n) \mathbf{tr}(D) \quad (6.15)$$

$$\mathbf{tr}(A_1) : \dots : \mathbf{tr}(A_n) = \mathbf{tr}(P^T D_1 P) : \dots : \mathbf{tr}(P^T D_n P) = \mathbf{tr}(D_1 P P^T) : \dots : \mathbf{tr}(D_n P P^T) \quad (6.16)$$

$$= \mathbf{tr}(D_1) : \dots : \mathbf{tr}(D_n) = \mathbf{tr}(\alpha_1 D) : \dots : \mathbf{tr}(\alpha_n D) = (\alpha_1 : \dots : \alpha_n) \mathbf{tr}(D) \quad (6.17)$$

For **sum**, applying Lemma 6.2:

$$\mathbf{sum}(A_1 : \dots : A_n) = \mathbf{sum}(P^T D_1 P : \dots : P^T D_n P) = \mathbf{sum}(P^T D_1 : \dots : D_n P) \quad (6.18)$$

$$= \sum_i \left( \sum_j P_{i,j} \right)^2 (D_1 : \dots : D_n)_{i,i} = \sum_i \left( \sum_j P_{i,j} \right)^2 (\alpha_1 : \dots : \alpha_n) D_{i,i} \quad (6.19)$$

$$\mathbf{sum}(A_1) : \dots : \mathbf{sum}(A_n) = \mathbf{sum}(P^T D_1 P) : \dots : \mathbf{sum}(P^T D_n P) \quad (6.20)$$

$$= \sum_i \left( \sum_j P_{i,j} \right)^2 (D_1)_{i,i} : \dots : \sum_i \left( \sum_j P_{i,j} \right)^2 (D_n)_{i,i} \quad (6.21)$$

$$= \sum_i \left( \sum_j P_{i,j} \right)^2 \alpha_1 D_{i,i} : \dots : \sum_i \left( \sum_j P_{i,j} \right)^2 \alpha_n D_{i,i} \quad (6.22)$$

$$= \sum_i \left( \sum_j P_{i,j} \right)^2 (\alpha_1 : \dots : \alpha_n) D_{i,i} \quad (6.23)$$

□

Proposition 3.2 (no discount with common eigenvectors and eigenvalues are equal up to scale factor).

*Proof.*

$$P_{combined} = (\mathbf{sum}(\alpha_1 P^T D P) : \dots : \mathbf{sum}(\alpha_I P^T D P)) \left( \sum_i^I \mathbf{sum}(\alpha_i P^T D P)^{-1} \mathbf{sum}(\mu^i) - 1 \right) \quad (6.24)$$

$$= (\alpha_1 : \dots : \alpha_I) \mathbf{sum}(P^T D P) \left( \mathbf{sum}(P^T D P)^{-1} \sum_i^I \alpha_i^{-1} \mathbf{sum}(\mu^i) - 1 \right) \quad (6.25)$$

$$= (\alpha_1 : \dots : \alpha_I) \left( \sum_i^I \alpha_i^{-1} \mathbf{sum}(\mu^i) \right) - (\alpha_1 : \dots : \alpha_I) \mathbf{sum}(P^T D P) \quad (6.26)$$



$$P_{separate} = \mathbf{sum} \left( (\alpha_1 P^T D P : \dots : \alpha_I P^T D P) \left( \sum_{i=1}^I \alpha_i P^T D P \mu^i \right) \right) - \mathbf{sum}(\alpha_1 P^T D P : \dots : \alpha_I P^T D P) \quad (6.27)$$

$$= (\alpha_1 : \dots : \alpha_I) \mathbf{sum} \left( \sum_{i=1}^I \alpha_i^{-1} \mu^i \right) - (\alpha_1 : \dots : \alpha_I) \mathbf{sum}(P^T D P) \quad (6.28)$$

These two expressions are equal, since the first term in both expressions is simply  $\sum_i^I \sum_n^N \alpha_i^{-1} \mu_n^i$ .

Therefore, there is no discount or premium.  $\square$

Lemma 3.5 (parallel sum for simultaneously diagonalizable matrices).

*Proof.* Since  $P$  is orthogonal,  $P^T = P^{-1}$ . The eigenvectors of  $A^{-1}$  are the same as  $A$ .

$$A_1 : \dots : A_n = ((P^T D_1 P)^{-1} + \dots + (P^T D_n P)^{-1})^{-1} \quad (6.29)$$

$$= (P^T D_1^{-1} P + \dots + P^T D_n^{-1} P)^{-1} = (P^T (D_1^{-1} + \dots + D_n^{-1}) P)^{-1} \quad (6.30)$$

$$= P^T (D_1^{-1} + \dots + D_n^{-1})^{-1} P = P^T (D_1 : \dots : D_n) P \quad (6.31)$$

$\square$

Lemma 3.6 (joint concavity for simultaneously diagonalizable matrices).

*Proof.* The left hand side is equal to  $tP^T D_1 P + (1-t)P^T D'_1 P : \dots : tP^T D_n P + (1-t)P^T D'_n P$ .

The right hand side is equal to  $t(P^T D_1 P : \dots : P^T D_n P) + (1-t)(P^T D'_1 P : \dots : P^T D'_n P)$ . The

results follow from Lemma 3.1 and Lemma 3.3.  $\square$

To prove Proposition 3.4 (maximization of dividend discount under trace constraint), we first establish two lemmas.

**Lemma 6.3.** (*convexity of weighted harmonic mean*): Suppose  $x, y \in \mathbb{R}$ ,  $x > y > 0$ . For  $t \in [0, 1]$ , let  $h(t; x, y)$  denote the weighted harmonic mean of  $x$  and  $y$ :

$$h(t; x, y) = \left( \frac{1-t}{x} + \frac{t}{y} \right)^{-1} \quad (6.32)$$

Then  $h(t; x, y)$  is strictly convex in  $t$ .

*Proof.* As  $t$  goes from 0 to 1,  $h(t; x, y)$  goes from  $y$  to  $x$ . The first and second derivatives of  $h(t; x, y)$  are:

$$\frac{\partial h}{\partial t} = \frac{x(x-y)y}{((1-t)x+ty)^2}, \quad \frac{\partial^2 h}{\partial t^2} = \frac{2x(x-y)y}{((1-t)x+ty)^3} \quad (6.33)$$

The second derivative is always positive, so  $h(t; x, y)$  is strictly convex. □

**Lemma 6.4.** Suppose  $x, y \in \mathbb{R}$ ,  $x > y > 0$ , and let  $N \geq 2$  be an integer. For  $n \in 0 \dots N$ , let  $H(n; x, y)$  denote

$$H(n; N, x, y) = \underbrace{x : \dots : x}_{n \text{ times}} : \underbrace{y : \dots : y}_{N-n \text{ times}} \quad (6.34)$$

$$= \left( \frac{n}{x} + \frac{N-n}{y} \right)^{-1} \quad (6.35)$$

Then  $H(n+1; N, x, y) - H(n; N, x, y)$ , for  $n \in 0 \dots N-1$ , is increasing in  $n$ .

*Proof.* This follows from the fact that  $H(n; N, x, y) = h(n/N; x, y)$  and strict convexity of  $h(t; x, y)$ . □

Proposition 3.4 (maximization of dividend discount under trace constraint).

*Proof.* Let  $v_n^i + \sigma^2 = [V_i]_n$ , the  $n$ -th diagonal entry of  $V_i$ . We wish to find the extreme points of

the feasible set defined by the constraints

$$\sum_{n=1}^N (v_n^i + \underline{\sigma}^2) = N\underline{\sigma}^2 + T, v_n^i \geq 0 \quad \text{for } i = 1, \dots, I, n = 1, \dots, N \quad (6.36)$$

For each  $i$ , this defines a  $N$ -dimensional simplex whose extreme points are where all  $v_n^i$ 's are zero except for one, which is equal to  $T$ . Each of these extreme points corresponds to a diagonal matrix where all diagonal entries are  $\underline{\sigma}^2$  except for one, which is equal to  $\underline{\sigma}^2 + T$ , i.e. a matrix specialized in variance for some asset.

The feasible set of the joint problem is an  $I$ -way Cartesian product of convex sets, therefore convex. Its extreme points are an  $I$ -way Cartesian product of the extreme points of the individual feasible sets. Suppose for asset  $n = 1, \dots, N$ , there are  $k_n$  matrices specialized in that asset. Then

$$\mathbf{sum}(V_1 : \dots : V_I) = \sum_{n=1}^N H(k_n; I, \underline{\sigma}^2 + T, \underline{\sigma}^2) \quad (6.37)$$

Suppose we reorder the assets such that the  $k_n$ 's are in descending order, and suppose there exists an asset  $m$  such that  $k_m < k_1 - 1$ . If we transfer the specialization of one of the matrices from asset 1 to asset  $m$ , this sum changes by the amount

$$\begin{aligned} & H(k_m + 1; I, \underline{\sigma}^2 + T, \underline{\sigma}^2) - H(k_m; I, \underline{\sigma}^2 + T, \underline{\sigma}^2) \\ & - (H(k_1; I, \underline{\sigma}^2 + T, \underline{\sigma}^2) - H(k_1 - 1; I, \underline{\sigma}^2 + T, \underline{\sigma}^2)) \end{aligned} \quad (6.38)$$

which results in a decrease, by Lemma 6.4. Therefore, the sum is minimized when it is no longer possible to decrease the maximum  $k_n$ , and the proposition is proved.  $\square$

## References

- ADMATI, A. (1985): “A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Prices,” *Econometrica*, 53(3), 629–658.
- ALLEN, F., AND D. GALE (1988): “Optimal Security Design,” *Review of Financial Studies*, 1(3), 229–263.
- BERGER, P. G., AND E. OFEK (1995): “Diversification’s effect on firm value,” *Journal of financial economics*, 37(1), 39–65.
- BERNSTEIN, D. S. (2009): *Matrix Mathematics - Theory, Facts, and Formulas*, 2nd ed. Princeton University Press.
- BHANDARI, T. (2013): “Differences of Opinion and Stock Prices: Evidence from Spin-Offs and Mergers,” Discussion paper, DERA Working Paper Series.
- BHATIA, R. (2007): *Positive definite matrices*. Princeton University Press.
- BOOT, A., AND A. THAKOR (1993): “Security Design,” *Journal of Finance*, 48, 1349–1378.
- BURCH, T., AND V. NANDA (2003): “Divisional diversity and the conglomerate discount: evidence from spinoffs,” *Journal of Financial Economics*, 70, 69–98.
- CAO, H., T. WANG, AND H. ZHANG (2005): “Model Uncertainty, Limited Market Participation, and Asset Prices,” *Review of Financial Studies*, 18(4), 1219–1251.
- CARPIO, R., AND M. GUO (2017): “On Equilibrium Existence in a Finite-Agent, Multi-Asset Noisy Rational Expectations Economy,” Discussion paper.

- CHAMBERLAIN, G., AND E. E. LEAMER (1976): “Matrix Weighted Averages and Posterior Bounds,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(1), 73–84.
- CHEMMANUR, T., K. KRISHNAN, AND D. NANDY (2014): “The effects of corporate spin-offs on productivity,” *Journal of Corporate Finance*, 27, 72–98.
- CHEMMANUR, T. J., AND A. YAN (2004): “A theory of corporate spin-offs,” *Journal of Financial Economics*, 72, 259–290.
- DAI, L. (2018): “Asset bundling and information acquisition of investors with different expertise,” *Journal of Economic Theory*, 175, 447–490.
- DALEY, L., V. MEHOTRA, AND R. SIVAKUMAR (1997): “Corporate focus and value creation: Evidence from spinoffs,” *Journal of Financial Economics*, 45, 257–281.
- DESAI, H., AND P. C. JAIN (1999): “Firm performance and focus: long-run stock market performance following spinoffs,” *Journal of financial economics*, 54(1), 75–101.
- FAMA, E., AND K. FRENCH (1997): “Industry Costs of Equity,” *Journal of Financial Economics*, 43, 153–93.
- FOUSS, F., A. PIROTTE, J. RENDERS, AND M. SAERENS (2007): “Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation,” *IEEE Transactions on Knowledge and Data Engineering*, 19, 355–369.
- FOUSS, F., M. SAERENS, AND M. SHIMBO (2016): *Algorithms and Models for Network Data and Link Analysis*. Cambridge University Press.
- GROSSMAN, S., AND J. STIGLITZ (1980): “On the impossibility of informationally efficient markets,” *American Economic Review*, 70, 393–408.

- HABIB, M. A., D. B. JOHNSEN, AND N. Y. NAIK (1995): “Spinoffs and Information,” *Journal of Financial Intermediation*, 6, 153–176.
- HELLWIG, M. F. (1980): “On the Aggregation of Information in Competitive Markets,” *Journal of Economic Theory*, 22, 477–498.
- HIRSHLEIFER, D., AND S. H. TEOH (2003): “Limited attention, information disclosure, and financial reporting,” *Journal of Accounting and Economics*, 36, 337–386.
- HOBERG, G., AND G. PHILLIPS (2010): “Product market synergies and competition in mergers and acquisitions: A text-based analysis,” *The Review of Financial Studies*, 23(10), 3773–3811.
- HORN, R. A. (1990): “The Hadamard Product,” in *Matrix Theory and Applications (Proceedings of Symposia in Applied Mathematics 40)*, pp. 87–169.
- KAUSTIA, M., AND V. RANTALA (2015): “Social learning and corporate peer effects,” *Journal of Financial Economics*, 117, 653–669.
- KRISHNASWAMI, S., AND V. SUBRAMANIAM (1999): “Information asymmetry, valuation, and the corporate spin-off decision,” *Journal of Financial Economics*, 53, 73–112.
- LA ROCHE, J. (2014): “Carl Icahn Shreds The Ebay Board, And Demands That Ebay And PayPal Split,” <http://www.businessinsider.com/carl-icahn-ebay-letter-2014-2>, Accessed: 2015/10/04.
- LANG, L. H., AND R. M. STULZ (1994): “Tobin’s  $q$ , corporate diversification, and firm performance,” *Journal of political economy*, 102(6), 1248–1280.

- LIU, Q., AND R. QI (2008): “Stock trading and diversification discount,” *Economics Letters*, 98, 35–40.
- MAKSIMOVIC, V., AND G. M. PHILLIPS (2013): “Conglomerate firms, internal capital markets, and the theory of the firm,” *Annu. Rev. Financ. Econ.*, 5(1), 225–244.
- MILLER, E. M. (1977): “Risk, uncertainty, and divergence of opinion,” *The Journal of finance*, 32(4), 1151–1168.
- NANDA, V., AND M. NARAYANAN (1999): “Disentangling value: financing needs, firm scope, and divestitures,” *Journal of Financial Intermediation*, 8, 174–204.
- SCHEINKMAN, J., AND W. XIONG (2003): “Overconfidence and Speculative Bubbles,” *Journal of Political Economy*, 111(6), 1183–1220.
- VAN NIEUWERBURGH, S., AND L. VELDKAMP (2009): “Information Immobility and the Home Bias Puzzle,” *Journal of Finance*, 64(3), 1187–1215.
- VELD, C., AND Y. VELD-MERKOULOVA (2009): “Value creation through spin-offs: A review of the empirical evidence,” *International Journal of Management Reviews*, 11(4), 407–420.
- VIJH, A. (1994): “The Spinoff and Merger Ex-Date Effects,” *Journal of Finance*, 49, 581–609.
- VILLALONGA, B. (2004): “Diversification discount or premium? New evidence from the business information tracking series,” *The Journal of Finance*, 59(2), 479–506.