Essays on Asset Management

Tianyi (Robert) Yang

Submitted to the Tepper School of Business in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Dissertation Committee:

Bryan Routledge (Chair)

Matthew Denes

Brent Glover

Burton Hollifield

Austin Sudbury

May 2020

A significant portion of household wealth is managed by financial professionals and investment companies. How these entities create value for their investors and whether their contributions are sufficient to justify the compensation are of great interest to both researchers and the general public. In this dissertation, we seek to further our understanding of the asset management industry.

Chapter 1 explores the relationship between the fee of an investment product and how it is described by the money manager. To control for variations in dynamic strategies of different funds, we focus on Unit Investment Trust (UIT), a class of investment product with buy-and-hold portfolios and fixed life spans. In order to compare different portfolios, we adopt the method of collaborative filtering to transform the portfolio to a latent representation. A significant portion of dispersion in fees of UITs is correlated with how the investment objectives are stated, even after controlling for the variation in the portfolios. UITs whose investment statements include more persuasive marketing languages and more discussions on stock selection procedures, diversification, and stock weightings charge higher sales fees.

In Chapter 2, we consider a model of money management that reconciles the investment expertise of institutional investors and the lack of outperformance on their in-house managed capital. Due to the synergy between institutions' knowledge and external asset managers' skill, institutions never trade on their information directly. Instead, informed institutions delegate investment decisions to external asset managers while passing on their investment knowledge. Only uninformed institutions invest in-house due to the search cost of finding a skilled asset manager. We discuss how management fees and the size of the money management industry are related to the number of institutions with expertise and the value of institution expertise. Price is more informative when knowledge of institutions is more diffused.

Chapter 3 is based on a joint paper with Bryan Routledge and Matthew Denes. We model the dynamic decisions of a Venture Capital (VC) fund during fundraising and capital deployment. VCs with better access and network receive greater deal flows, raise larger funds, and deploy capital more frequently. On the contrary, VCs with better screening abilities deploy capital at a slower rate due to the higher expected payoffs of future prospective projects. We estimate the model using

VC deal data at each fund level. Funds differ mainly in access and are homogeneous in terms of screening ability. Differences in deal flows alone account for almost all the dispersion in fund values. By optimally matching estimated deal flow and screening, we can improve the average fund value by 2.5 percent. This corresponds to the unrealized gain from complementarity.

I am grateful to all my dissertation committee members Bryan Routledge, Matthew Denes, Brent Glover, Burton Hollifield, and Austin Sudbury for their advice and guidance on this dissertation. My gratitude also goes to Chris Telmer, Chester Spatt, and Stefano Sacchetto for insightful feedback and support on my various projects. I am especially indebted to Bryan Routledge for his help and patience throughout my PhD career. There were numerous occasions where the obstacles seem insurmountable at the time and I would have given up had Bryan not been there constantly offering encouragement and perspective.

I am also grateful to my peers and friends. Siyu Lu, Jiayi Li, and Wenjie Xue have many times offered discussions that lead to insightful feedback. They have also made My life as a PhD student much more enjoyable outside of research.

Lastly, I would like to thank my family. To my mom Shen Lee and my dad George Lee, thank you for always believing in me and never letting me forget how proud you are of me. To my amazing wife and best friend Xiaowen Yin, thank you for all your sacrifices over the past six years. Your companion and positivity have gotten me through some of my darkest times. I would not have gotten to this point without you by my side.

TABLE OF CONTENTS

CHAPTER 1. Fees and Language of Unit Investment Trusts

1.1 Introduction

We explore how the language money managers use in the marketing material may affect the fees that they charge. Specifically, can one fund with the same strategy as another charge a higher fee by describing its strategy differently? If so, what kind of languages are associated with higher fees?

Understanding the role language plays in how investors shop for funds can provide insight into the frictions and rationality within the market of asset management. If indeed investors pay higher fees based on the marketing language alone, this could suggest a novel source of behavior bias among investors. Alternatively, we can interpret the effect of the language on fees through the lens of a search model. It is possible that the high-fee funds with well-written documents are flagship funds with effectively low search costs to the investors. Both explanations shed light on the nature of interactions between money managers and investors.

Ideally, we would like to compare the fees of two otherwise identical funds that only differ in their descriptions of the strategies. However, since the strategies money managers employ are dynamic, no two funds are truly identical. Even if the portfolios of two funds happen to align at some moment, they will likely diverge once one fund makes a trade. In addition, the dynamic actions of the money managers are usually unobservable to the economists. As such, it is difficult to control for the underlying strategies of the funds and isolate the effect of the language on fees.

We address this issue by exploiting the unique features of Unit Investment Trusts(UITs), a class of investment vehicles legally required to hold fixed portfolios. Consequently, when comparing two UITs, we only need to look at their initial portfolios. There is no need to control for the difference in market timing since it is muted for all UITs.

However, even static portfolios of UITs are hard to compare, thus controlled for. Each portfolio can be seen as a high dimensional object encompassing which companies to hold and how much of each to hold. Even if we only consider the binary choice of whether to include any given company into the portfolio, with 5000 potential companies, there are 2^{5000} possible portfolio configurations.

To address the dimensionality problem, we employ collaborative filtering, a machine learning technique that projects the portfolio onto a lower-dimensional space. Intuitively, we leverage the collection of all UIT portfolios. By examining which UITs tend to hold similar companies, and which companies are generally held by similar UITs, the algorithm picks up the latent structure between UITs and companies. The resulting latent representations are essentially the most efficient way of recovering the observed similarity between UITs and companies. We then cluster UITs based on the latent representations to control for the variation in asset allocation.

To extract information in the investment objectives, we adopt Latent Dirichlet Allocation (LDA), a generative language model. The outputs of LDA include a set of topics commonly discussed in the investment objectives, as well as how much the investment objective of each UIT is devoted to each topic. We have thus transformed each unstructured text into a discrete distribution over topics, which we can much more easily analyze with the fees.

We interpret the meanings of topics by examining the words most commonly used within each topic. Upon inspection, certain topics correspond to specific fields such as infrastructure and technology. Other topics focus on the criterion of stock selection, diversification, and stock weighting. One topic stands out for the purpose of our paper, which we refer to as "marketing". The representative words under this topic include *best*, *rank*, *prospect*, *peer*, and *favor*.

We then regress the fees of UITs on the latent representations of portfolios and topics distribution of the investment objective. By introducing topic distributions, the adjusted R^2 increase from 19% to 31%, even after controlling for the portfolio and other features of UITs. We conclude the description of the portfolios is significantly correlated to fees, even conditioning on the underlying portfolios.

Among topics covered in the investment objective, more discussions on the stock selection process, benefits of diversification and stock weightings are associated with higher fees. UITs whose investment statements include more marketing languages also have higher fees. These results are consistent with the aforementioned behavioral bias of investors and the role language plays in the context of search frictions.

This paper also contributes to the literature as the first systematic review of UIT portfolios. In addition to providing main institutional backgrounds on the industry, we document that the typical UIT holds a 1/N portfolios with around 30 stocks, has a lifespan of 2 years, and charges an annualized sales load of 2%. Equity UITs tend to invest in blue-chip companies and there is no evidence of outperformance relative to common benchmarks.

Section 1.2 summarizes the basic institutional backgrounds of the UIT industry. In section 1.3 we discuss the dataset and show some high-level statistics. In section 1.4 we examine the relationship between the language and the fee. Section 1.5 discusses the results. Section 1.6 reviews related literature. Section 1.7 concludes.

1.2 Overview of the UIT Industry

A Unit Investment Trust (UIT) is a unique type of investment product with a buy-and-hold portfolio and predetermined termination date. Like mutual funds, UITs offer redeemable shares that allow investors to cash out. However, as required by law, there is no trading during the life of the trusts[1].

Despite its low name recognition, UIT is actually one of the four types of registered investment companies classified under the Investment Company Act of 1940. The others are mutual funds, exchange-traded funds (ETFs), and closed-end funds (CEFs).

While UITs account for a small fraction of the total market share of investment companies, they have received an even smaller share of scrutiny from both academia and the public. As of

[1]There are special scenarios where trading is allowed, as stated in the prospectus. Specifically, if some underlying company undergoes an M&A or bankruptcy, the issuer of the trust is allowed to liquid the position and substitute with similar investments.

2018, total investments in UIT is \$70B. For comparison, there is \$17T invested in mutual funds, \$3T in ETF, and \$250B in closed-end funds. While smaller than its sibling, the absolute amount of \$70B is still sizable and should warrant UITs some attention. However, a keyword search on SSRN suggests only 6 research papers mention UIT in some capacity (the same numbers are 3421 for mutual fund, 430 for ETF, 150 for CEF).

Given the limited exposure, we highlight in this section some institutional details and common questions regarding UITs.

1.2.1 History and aggregate trends

The first UIT was issued in 1961. Its portfolio consists of solely tax-exempt municipal bonds. Historically, the majority of the asset under management of UITs are in bond securities. During the earlier days, UITs are mainly used for investors to gain exposure while achieving diversification. In particular, UITs are used to gain access to the bond market, both taxable and tax-exempt.

The recent decades have witnessed a shift in the composition of UIT investments, in particular, the rise of equity UITs. Equity UITs have taken over the industry, both in terms of the number of trusts and the total asset under management. As of 2018, equity UITs account for 86% of the total UIT net assets.

The overall size of the UIT industry has fluctuated between \$30B and \$100B for the past two decades. In comparison, the same period has seen a steep increase in the size of both mutual funds and ETFs. Effectively, the market share of UITs has decreased drastically, in large part because of the advancement of information technology and alternative trading venues becoming cheaper and more accessible.

1.2.2 Participants

For each UIT, there are two important parties: the sponsor and the trustee. The sponsor is the issuer of the UIT, responsible for constructing the underlying portfolio. The trustee is the custodian of the UIT, who is in charge of bookkeeping and meeting reporting requirements.

The sponsors are often themselves broker-dealers of the UITs they construct. The entire UIT industry is dominated by 5 players: First Trust, advisors asset management(AAM), Guggenheim, Invesco, and Van Kampen. Each of them constructs their own UITs and sell them directly on their websites. Some also distribute UITs through other verified broker-dealers. As a result, the sponsors are compensated by the organization costs and the sales fee.

Sponsors are not required to maintain a secondary exchange for the UITs they issue. However, all of the major sponsors do.

The trustees are usually banks such as BNY Mellon and Chase. Trustees are compensated for operating expenses. In addition to overseeing the trust, the trustee sometimes serves as the provider of the seeding capital. In this case, the sponsor first applies for a line of credit from the trustee. The sponsor then uses the capital to purchase the underlying securities of the UIT portfolio. Finally, the securities are deposited to the trustees in exchange for the certificates representing the units of the UIT, which are then sold to the investors.

Since UITs, as required by law, possess buy-and-hold portfolios, there is no trading during the life of the trusts. As such, there is no board of directors or investment advisers.

The identities of the investors for UIT can not be explicitly found. However, from the minimum purchasing requirements stated on the prospectus, retail investors are the target demographic. There is a minimum purchase of $1,000 for each UIT. For investment under the IRA, the amount drops to $500.

1.2.3 UIT vs. alternatives

The most distinct features of UITs compared with other types of investment vehicles are the predetermined termination date the absence of trading. According to UIT sponsors' marketing materials, investors should invest in a UIT if they value: (a) diversification (b) transparency, and (c) capital utilization. Investors of UITs are those who seek to achieve the degree of diversification commensurate to mutual funds, but do not trust money managers enough to make trades on their behalves. In this regard, the UITs are similar to index mutual funds, though UITs do not follow any particular indices.

As UITs are structured as pass-through entities, the tax implication regarding dividends is the same as if the underlying securities are held by the investors directly. In terms of capital gain, since the termination date is pre-specified, investors have better control over when the capital gain is realized. In contrast, mutual funds tend to realize gain at the year-end.

1.3 Data

The main source of data is the collection of Form 487s, the prospectus of UITs. We focus on equity funds from three of the largest sponsors: First Trust, Advisors Asset Management, and Guggenheim. Together these three sponsors account for 55% of all Form 487s available in SEC's EDGAR database. For each UIT, we extract from its prospectus the complete information on its portfolio. For each underlying investment, we obtain the ticker, the name, and the portfolio weight. Also, for each UIT, we observe the inception date, termination date, as well as information on fees, including sales charge, organization cost, and expense ratio.

We then matched the portfolio information with data on returns from CRSP to construct a panel data set of monthly UIT returns. In addition, we extract the time series of market returns from CRSP, factors from Ken French's website, and benchmark Vanguard index funds from Yahoo Finance. Doing so allows us to evaluate the performance of the UIT investments against various benchmark models. From CRSP we also obtain share class (common equity vs. CEF) and the NAICS industry code for each underlying investment. Doing so enables us to examine asset allocation across industries and asset types. After the necessary cleaning and formatting, the final sample consists of 7490 UITs. The earliest UIT in the sample begins in June 2004 and the latest July 2017.

1.3.1 Fees and Durations

Table 1.A reports the summary statistics on fees, duration, and net asset. Panel A shows the aggregate statistics. The average UIT has a life span of 2 years. The annualized organization cost and sales load paid to sponsors are 0.17% and 2.02%, respectively. The average expense ratio paid to the trustee is 0.81%. The average net asset as shown on the prospectus is at $164,339. Net assets tend to cluster around $150,000 because it is the minimum required initial seeding capital for any investment companies. The trust may receive additional flow after the initial issuing period. Hence we do not interpret it as the actual size of the UIT.

Panel B reports the same statistics for each trust type. Among all UITs in the data set, 4396 invest solely in common equity; 1744 invest solely in Closed-End Funds; The remaining 1350 are hybrid UITs investing in a mixture of the common equities and CEFs. We hereafter refer to them as equity UITs, CEF UITs, and hybrid UITs.

All three types of UITs have similar durations and fees. The only significant difference is in expense ratio, where CEF UITs and hybrid UITs on average charge 1.45% and 1.75%. In contrast, the average equity trust charges only 0.25%. This is because the expense ratio reported on prospectus also includes the expense ratio of the underlying funds if there are CEFs in the portfolio.

Panel C further breakdowns the sample by sponsor identity. Here we see First Trust (FT) accounts for over half of the sample. However, there are no significant differences across sponsors in terms of fees and duration, conditional on the trust types.

Note from above that the sales load accounts for the majority of the total compensation to the sponsors. In addition, the dispersion in sales load across all UITs is larger than that of the organization costs: the standard deviation is 0.41% for sales load and only 0.09% for organization costs. The organization costs can thus be interpreted as a constant overhead. As such, we only consider the sales load when exploring the variation in the costs of UITs in the later sections.

Here we need to note the correlation between sales load and duration. While the previous statistics are based on annualized sales load, the prospectus only reports the total sales load. Figure

1.1 illustrates the schedule of sales charge, which plots the total sales load against the duration of the UITs. The majority of UITs fall into one of three clusters: 15-month duration charging 2.95% sales fee; 2-year duration charging 3.95% sales fee; and 5-year duration charging 4.95%. Effectively, the longer the duration of a UIT, the lower the sales load is in annualized terms.

Since duration and sales load are highly correlated, it will be difficult to disentangle which is causing which. Here we take the stance that the dispersion in fees is the driver. Specifically, given the established sales fee schedule, sponsors effectively charge a higher annualized fee by shortening the duration of the UITs. While the data or the institutional backgrounds provide little clarification, there is anecdotal evidence in support of this assumption[2].

1.3.2 Portfolio and Investment Objectives

We mentioned that UITs have buy-and-hold portfolios. So what kind of securities do they hold? How many securities do each hold? and how are assets allocated across those securities? In this section, we describe the portfolios of the UITs in our sample and provide answers to these questions.

From Panel A of Table 1.2, we see the median UIT portfolio holds 30 different securities. 90% of the UIT portfolios hold less than 60 securities. Hybrid UITs on average hold more securities with an average of 40, followed by equity UITs' 38 and CEF UITs' 38. The UIT with the most unique number of holdings is an equity fund holding 187 different securities. Compared with index funds tracking S$P 500 or Russell 2000's, UITs are more selective in terms of portfolio formation.

UITs usually assign portfolio weights of 1/N to all the securities in the portfolio. Let $[w_1, w_2, \ldots, w_N]$ be the portfolio weights (in percentage) of a UIT investing in N distinct companies (such that $\sum_i^N w_i = 100$). We define the *maximum 1/N deviation* as $\max_{i=1,\ldots,N} |w_i - 100/N|$. [3] Panel B of table 1.2 reports that 75% of the UITs have maximum 1/N deviation less than 0.87. In particular,

[2] "Raymond James Agrees to Pay $15 Million for Improperly Charging Retail Investors" Associates at Raymond James were fined for recommending that customers liquidate their positions earlier than necessary, effectively shortening the holding period and incurring higher annualized fee. This suggests sponsors at least are aware that variation in duration can be used to charge differential fees.

[3] For example, for a portfolio with five stocks with portfolio weights 16%, 20%, 20%, 22% and 22%, the maximum 1/N deviation is $|20 - 16| = 4$.

equity UITs have a median maximum 1/N deviation of only 0.03. For comparison, the largest position in S&P 500 is Apple, which account for more than 4% of the total portfolio value. Thus the maximum 1/N deviation for an S&P Index would be closed to 4-0.2 = 3.8.

Note that the 1/N portfolio weights apply only at the inception of the UIT. Since UITs have buy-and-hold portfolios, as the prices of underlying securities change during the trust's life, the portfolio weight would typically deviate from 1/N.

Panel C of Table 1.2 briefly comments on the investment objectives of the trusts. Here we only report the word counts. The average investment objective contains 305 words. The dispersion in word counts is larger among equity UITs than among CEF or hybrid UITs. We defer a more detailed analysis of the language in the next section.

Table 1.3 lists the names and counts of the top common equities and Closed-end funds held by UITs in our sample. Among 4396 equity funds, 1113 have a positive position in Pfizer, 1019 in Chevron, 1018 in AT&T, 999 in Verizon. Each of the above companies appears in more than a quarter of all the equity trusts. All top 20 companies held by equity UITs are blue-chip companies. For CEF UITs, closed-end funds from Invesco, Wells Fargo, and Eaton Vance are among the most held.

While some common equities and CEFs are held frequently by the UITs, most securities in the universe appear sparingly. Figure 1.2 plots the numbers of appearances for all equities and CEFs. For equity UITs, the blue-chip companies from Table 3 only represent a tiny slice of the entire set of firms. Specifically, less than 1% of the companies are held by more than 10% of the equity UITs, whereas more than 75% of the companies are held by less than 1% of the companies. The difference within CEFs is smaller. Around 2% of the CEFs are held by more than 10% of the CEF UITs and only 15% of the CEFs are held by less than 1% of the CEF UITs. In other words, it is much more likely for CEFs to occur in multiple UITs than common equities.

1.3.3 Performance

Based on portfolio information, we construct monthly returns for each UIT. We then construct aggregate UIT portfolios as well as UIT portfolios based on trust types. Specifically, for each month, the return of the aggregate portfolio is the simple average of all active UITs in that month. We then evaluate the time series of returns against the CAPM, Fama-French 3 Factor Model, the Carhart 4-factor model, and the model relative a set of Vanguard Index funds as suggest by Berk and van Binsbergen (2015). There is no evidence of outperformance or underperformance.

1.4 Language as drivers of dispersion in fees

The main goal of this paper is to show how the description of an investment product by the money managers can influence the fees investors are willing to pay. If there is a relationship, what are investors inferring from the descriptions? Are the descriptions informative about the skill of the managers or are the writings more of a marketing tool that misguides investors into paying a higher fee?

Isolating the effect of language is difficult since how the investment product is described is likely correlated with other features of the products. For example, mutual funds focusing on the energy sector will devote a sizable portion of their marketing materials on topics related to the energy sector. Without controlling for what is being invested, one might falsely conclude it is the *description* on the energy sector that allows a fund to charge a higher fee, whereas it is that the higher fees reflect the higher transaction cost of trading companies in the energy sector.

For any investment products, two essential features that need to be controlled for are asset allocation and market timing. In terms of asset allocation, even though certain mutual funds are required to disclose their holding on a quarterly basis, there is a concern of window dressing. So even if two funds disclose similar holdings on the reports, they could have different positions most of the remaining time. Market timing is even more problematic. When an investor invests with a certain active fund, she is trusting the *dynamic* decisions of the money manager. Two money managers can trade different assets at different times, charge different prices, and provide different descriptions of their strategies. Correlation cannot be established if the first two cannot be reasonably controlled.

The unique legal features of UITs allow us to plausibly control for asset allocation and market timing. Specifically, the law prohibiting interim trading eliminates any variation in market timing. Given the buy-and-hold nature of the portfolios, it suffices to compare only UITs at their inceptions. Our goal is to investigate whether two UITs with similar portfolios indeed charge similar fees. If not, how much of the difference in fees is correlated with the difference in how they are described?

1.4.1 Portfolio representation

By focusing on UITs, we decrease the number of controls significantly. However, even for buy-and-holding portfolios, the dimensionality problem still exists. For equity trusts, there are over 4000 potential securities. Even if we only consider the binary choices of whether to invest in each company, there are still more than 2^{4000} possible combinations. As such, rarely are two portfolios identical. It is then impossible to explore the additional impact of language since fees can already be explained by the portfolio alone. In this section, we lay out our procedure of projecting portfolios onto a much more manageable lower-dimensional latent space.

In addition to dimensionality concerns, there is another reason for finding latent representations of the portfolios. Since our goal is to investigate the effects of portfolio description on consumers, we should summarize the portfolio with the level of sophistication comparable to the mental capacity of the UIT investors. It is unlikely that the investors of the UITs scrutinize every single holding when comparing two UIT. The latent representation essentially captures the investors' perception of the portfolio. The assumption is that two UITs with the identical latent representations are the same subject to investors' mental bandwidth.

To obtain the latent representation, we employ collaborative filtering, a class of algorithms for detecting patterns in large sparse datasets. Collaborative filtering is most commonly used in recommender systems by Netflix and Amazon. In Netflix's case, they maintain a sparse dataset connecting a set of users and a set of movies. In order to decide which movies a given user is more likely to enjoy, the algorithm ranks all movies based on ratings of other users with similar preferences. To determine which users have similar preferences, the algorithm compares their existing ratings on movies.

In the context of collaborative filtering, each dimension of the latent representation should capture an aspect of the movie that best explains the observed rating patterns. For Netflix, each dimension might represent a particular genre, say thrillers. For each movie, the value along that dimension in its latent representation captures how closely that movie can be categorized as a thriller. And for the latent representation of a user, the value would represent how much the user prefers thrillers.

In our context, for the underlying companies, each dimension of latent representation could represent its industry, size, or other characteristics. For the UITs, the same latent dimension could correspond to how much of the strategy prioritizes the specific characteristics.

The advantage of collaborative filtering is that it imposes no prior on how we should compare different portfolios. As we saw in the previous section, UITs portfolios do not fall cleanly into sector buckets. While companies differ in numerous ways, the features derived from the latent representations are what UIT managers and investors deem the most important.

The specific algorithm we use is Non-negative Matrix Factorization (NMF). Formally, The collection of all portfolios can be considered as an aggregate holding matrix H with n rows and m columns. Here n is the number of equity UITs and m is the number of companies that appears in at least one UIT. The ij'th entry of H is 1 if the ith UIT has a positive position in security j, and 0 otherwise. Figure 1.3 illustrates the aggregate holding matrix of all equity UITs. The matrix is sparse as the average UIT invests only 30 out of more than 4000 potential companies. Certain columns are relatively dense with dots – these correspond to the popular blue-chip companies listed in Table 1.3.

Let k be the number of latent dimensions. We can express the collection of latent representations of the UIT as $T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix}$ where $T_i = \begin{bmatrix} T_{i1} & T_{i2} & \dots & T_{ik} \end{bmatrix}$ is the latent representation for the ith trust. Similarly let S be representation for the securities. NMF can be formulated as the following minimization problem:

$$\min_{T,S} ||H - TS'||_2^2 = \sum_i^n \sum_i^m \left(H_{ij} - \sum_l^k T_{il} S_{jl} \right)^2 \tag{1.1}$$

where T is $n \times k$, S is $m \times k$. Both T and S have non-negative entries.

There are two reasons why we use a binary entry for H instead of the actual portfolio weight. First, as noted in the previous section, since most trusts tend to hold 1/N portfolios, conditioning on the number of holdings, there is little variation in portfolio weights. Secondly, a binary entry offers an intuitive interpretation. The latent representation of a UIT and each company can be regarded as a vector of unit length. Each entry H_{ij} in the holding matrix approximated by the inner product of T_i and S_j. Note the inner product is at most 1 when $T_i = S_j$ and at least 0 when T_i and S_j are orthogonal. In this case, the entries in H are indeed the similarity measures.

Alternatively, NMF provides an estimator to the following underlying data generating process:

$$H_{ij} = T_i \times S_j + \epsilon_{ij} \tag{1.2}$$

with the added assumption that ϵ_{ij} are iid and satisfies the regularity conditions.

An iterative algorithm is used for estimation. The algorithm also sheds light on the intuition of what we are trying to achieve. Within each step, the algorithm first fixes T and optimizes with respect to S; the algorithm then fixes S and optimizes with respect to T. The former embodies the idea that a UIT should invest in companies consistent with its objective; the latter states the objective of a UIT is informed by the companies it invests in.

Table 1.4 summarizes the estimation results when k is 20.[4] To interpret what each dimension represents, we list the names of the top UITs and securities that load the most along the particular dimension. The order in Table 1.4 has no particular meaning since all dimensions are estimated simultaneously. Certain dimensions represent specific industries. Examples are dimension 2 for Health Care, 3 For Real Estate, 5 for financial firms, 6 for agriculture, 16 for technology, 18 for utility, and 20 for energy.

More interesting are the dimensions representing other characteristics of the underlying companies. From the names of representative UITs for dimension 4,9,14,15,17, all seem to aim for high dividend or equity income. However, the estimation results suggest they are different enough. Indeed, by examining the top 2 representative companies, we see that these 5 dimensions are picking up different companies.

The above highlights the advantage of employing collaborative filtering as a tool for dimension reduction. Since the market for UITs is separated from the market of the typical investment products, the metrics UIT investors use to compare companies are not known to us. We could have used a set of common controls like industries and key financial variables as the latent representation. However, we would capture certain features that UIT investors are indifferent about, and overlook certain features normally not used to describe the firms – like the features that distinguish companies with high dividends.

For robustness checks, we also implement the same algorithms (1) without the non-negativity conditions (2) with Principal Component Analysis instead of NMF, and (3) use actual portfolio weights. The results are robust to different specifications.

In the subsequent sections, we will use these latent representations T for UITs as inputs for further analysis and assume that they capture all significant variations in portfolios.

[4]Here the number 20 is chosen as a heuristics. We run the same exercise for $k = 10, 20, 30$ and obtain similar results.

1.4.2 Text Representation

We adopt a bag-of-words model. Conceptually, we deconstruct a block of text into word counts, ignoring their orders. For example, "When you invest in a Trust you are purchasing a quality portfolio of attractive common stocks in one convenient purchase" will be translated into $[When : 1, You : 2, invest : 1, Trust : 1, purchasing : 1, \ldots]$. In practice, we remove stop words such as the, a, in since they provide little insight. We also combine derivationally related words. In the previous example, both *purchase* and *purchasing* would then be counted toward the same token.

Similar to the portfolio, texts even simplified as bags of words are still high dimensional objects. Hence we rely on Latent Dirichlet Allocation (LDA), a generative language model commonly used in topic modeling. The most prevalent usage of LDA is to classify news articles into topics such as politics, sport, and entertainment. The idea is that certain words appear often and concurrently only in certain contexts. Here we are doing the same thing for the collection of investment objectives. Similar to collaborative filtering, LDA lets the texts speak for themselves what the common topics should be.

We formally describe the underlying model of LDA. Suppose there are k topics. For each document i, let $P_i = \{p_{i1}, p_{i2}, \ldots, p_{ik}\}$ be the distribution over the potential topics. Intuitively, p_{ij} measures the fraction of document i written on topic j. $\sum_j^k p_{ij} = 1$

For each topic j, let $Q_j = \{q_{j1}, q_{j2}, \ldots, q_{jN}\}$ be the discrete distribution over the dictionary of all available words. Here N is the number of distinct words in the corpus, the collection of all documents. Let $W = [w_1, w_2, \ldots, w_N]$ be the dictionary. q_{j1} measures how often word w_1 appears under topic j.

Under LDA, each document i is generated in the following fashion. For each word in document i, a topic j is sampled according to the topic distribution P_i. Then based on the selected topic j, a word is chosen according to Q_j. [5]

[5] In fact, the topic distribution for each document, P_i, is first sampled from a Dirichlet distribution characterized by parameter $\alpha_i, i = 1, 2, \ldots, k$. α and q are the parameters estimated by LDA. P_i is in fact the latent topic distribution that is neither the parameter nor observable.

Here we combine the investment objectives of all UITs when estimating the LDA model. Specifically, we do not distinguish between equity UITs and CEF UITs as we did for the portfolio representation. Even though the portfolios of equity UITs and CEF UITs do not overlap, the language the sponsor use to describe them could share common themes. By estimating over the entire sample, we are capturing these commonalities. We estimate the model with 14 topics.[6]

To best interpret each topic, we find words that are most representative of the topic. Specifically, for each word w, we compute the conditional distributions for all topics. That is

$$r_{jw} = P(\text{topic } j|\text{word } w) = \frac{q_{jw}}{\sum_{i=1}^{k} q_{iw}} \tag{1.3}$$

In figure 1.4, for each topic, we show the word cloud consisted of the top 15 words based on the words' conditional probability. The size of the words reflects the magnitudes of the conditional probability.

Here it is important to note the difference between q_{jw}, $P(\text{word } w|\text{topic } j)$ and r_{jw}, $P(\text{topic } j|\text{word } w)$. Though the former is defined as part of the data generating process, it is not useful in uncovering the meanings of the topics. In fact, words like *stock* and *invest* rank consistently high for multiple topics. And that is precisely why they are not informative about the unique aspects of each topic. In Figure 1.4, the overall frequency of a word is represented by its darkness. We see most words that are reliable signals of certain topics are not frequently observed words, with the exceptions of *closed*, *end*, *fund*, and *tax*.

From the word clouds, we can infer the meaning of each topic. In the Appendix, we also provide the sample investment objectives for each topic. Similar to the latent representation of portfolios, certain topics are industry-specific. For example, there is one topic dedicated to REIT, one to infrastructure, two on different aspects of the tech industry, and one for CEF overall.

The majority of the topics, however, are not industry-specific. For example, the topic which we denote *stock_basic* includes basic accounting and finance terms typically used to analyze a stock – *cap*, *style*, *book*, *analyst*; The topic *stock_Selection* include words associated with the exact

[6]There is no strict guideline regarding what the optimal number of topics should be. Here choose 14 based on perpexity, which measures the performance of language model over the holdout sample.

steps of constructing a portfolio – $method, step, target, inclusion$. Topics $Diversification$ and $Stock\_Weighting$ are self-explanatory and echos the features of UITs discussed in the previous sections.

Perhaps the most interesting is the topic which we call $marketing$. Word $best$ ranks highest within this group. Based on other signal words such as $rank$, $favor$, $insider$, and $prospect$, and the representative documents, we can infer that this topic is less about objective description and more about subjective persuasion, hence the name $marketing$[7]. As we investigate the effects of language, we will pay particular attention to the specific impact of this topic.

Figure 1.5 shows the average topic allocation for both equity UITs and CEF UITs. Topics such as $Stock\_Basic$ and $Stock\_Selection$ are mainly included in equity UITs whereas CEF and Tax are only discussed in CEF UITs' investment objective.

The result of LDA yields 14 topics with our inferred meaning. For each UIT, the estimated topic distribution P_i informs how the entire investment objective is split among discussions on these 14 topics. In the next section, we will use P_i, the distribution over 14 discrete topics as input for text.

While there are other state-of-the-art language models that can account for structure and grammar, the bag-of-word model we use here serves as a first pass. In the next section we will show that even after discarding some information within the text, the remained still captures a significant portion of the variation in fees, conditional on the portfolio. Future extensions could employ more complex language models and see if more insights can be drawn.

1.4.3 Do Portfolio Descriptions Matter for Fee?

One direct way of testing whether the texts in investment objectives can explain the dispersion in fees beyond the portfolios is to regress fee with and without text and compare the adjusted R^2.

[7]Here the appearance of *insider* may seem surprising. A sample usage is "our conviction related to companies which may have had recent insider buying in open market purchases by officers and directors" the UIT managers are not claiming access to insider information. The above usage is more akin to the usage of *best*

If the adjusted R^2 increases significantly, then we can conclude that at least a portion of the fees is correlated with the language that is used.

Specifically, for both samples of equity UITs and CEF UITs, we estimate the following 5 specifications:

$$Fee_i = \beta_0 + Controls_i + \epsilon_i \tag{1.4}$$

$$Fee_i = \beta_0 + Controls_i + \beta_1 Port\_Basics_i + \epsilon_i \tag{1.5}$$

$$Fee_i = \beta_0 + Controls_i + \beta_1 Port\_Basics_i + \beta_2 C\left(T_i\right) + \epsilon_i \tag{1.6}$$

$$Fee_i = \beta_0 + Controls_i + \beta_1 Port\_Basics_i + \beta_2 C\left(T_i\right) + \beta_3 Text\_basics_i + \epsilon_i \tag{1.7}$$

$$Fee_i = \beta_0 + Controls_i + \beta_1 Port\_Basics_i + \beta_2 C\left(T_i\right) + \beta_3 Text\_basics_i + \beta_4 P_i + \epsilon_i \tag{1.8}$$

For each UIT i, $Control_i$ includes sponsor and year fixed-effects. $Port\_Basics$ are basic portfolio characteristics outlined in Table 1.2, which includes how many securities are in the portfolio, and how close are the portfolio weights to 1/N. $Text\_Basics$ include word counts as reported in Panel C of 1.2. P_i is the output of LDA.

Unlike P_i, which are distributions operating in a linear fashion, T_i, the latent representation of portfolio is an non-linear object that cannot be directly feed into the linear specification. As a result, we assign each UIT into one of 20 groups corresponding to the dimension of the latent representation that T_i loads the most on. Specifically, we assign UIT i to group $C\left(T_i\right) = \arg\max_{j \in \{1,2,\ldots,k\}} T_{ij}$. We then use group $C\left(T_i\right)$ to capture the variation in portfolio compositions.

Figure 1.6 shows the level of R^2 for each of the 5 specifications on both Equity trusts and CEF trusts. For Equity trusts. The fixed effects alone explain less than 5% of the total variation. Introducing basic portfolio variables does little. By introducing the $C(T_i)$, which by assumption encompass all there is about the portfolio, we increase the adjusted R^2 by 12 percentage points. We then investigate the effect of including text features. Adding word counts alone has little effects. However, once we introduce topic distribution P_i into the right-hand side, the adjusted R^2 increases by an additional 12 percentage points from 19% to 31%. We conclude that for equity UITs, the language used in the investment objective is indeed correlated with the fee, even after accounting for the actual portfolio.

For CEF UITs, however, we do not observe an increase in adjusted R^2 of the same magnitudes. Sponsor and year fixed effects alone can already account for 30% of the total variations in fees. Introducing text into the specification increases the adjusted R^2 by less than 3 percentage points.

One valid critique on constructing portfolio groups $C(T_i)$ is that certain variation in the latent representation is lost in the clustering process. If so, the increase in adjusted R^2 from texts could be the result of fees and texts both correlating to the omitted portion of the portfolio. To address this concern, we run a robustness test with T_i as the input instead $C(T_i)$. We see similar levels of adjusted R^2 across all 5 specifications.

1.4.4 Which topics have the most impact on fees?

The previous section establishes that language is correlated with fees for equity UITs. In this section, we further investigate which topics are associated with higher fees and which with lower fees. We adopt the full specification in (1.8). We exclude the element in P_i corresponding to $Stock\_Basic$ since it is the most common topic among equity UITs. As such, the coefficient corresponding to other topics is measured relative to $Stock\_Basic$.

Figure 1.7 plots $\hat{\beta}_4$, the estimated vector of coefficients for all 14 topics other than $Stock\_Basic$. We also highlight the significance by drawing the confidence interval. Among all topics, $Marketing$, $Diversification$, $Stock\_Selection$, and $Stock\_Weighting$ have the largest significant estimates. Specifically, a 1% increase in the portion of investment objective devoting to marketing is associated with a 1.42 basis point increase in the annualized sales fee charged. Similarly, an 1% increase in the discussions on stock selection, stock weighting, and diversification is associated with 0.97, 0.84, 0.79 basis points, respectively.

The effects of these four topics are also economically significant. The standard deviations in the fraction spent on these topics are 4.2% for marketing, 3.4% for Diversification, 11% for stock selection, 9% for stock weighting. Hence one standard deviation increase in each of the above topics is associated with 6.00, 3.34, 9.77, 7.80 basis points increase in annual load, all of which are sizeable relative to the sample standard deviation of 42 basis points among fees of equity UITs.

In contrast, discussions on Technology, CEF, and options are associated with lower fees, with corresponding coefficients of -1.13, -0.96, -0.86. The magnitudes here are also economically significant.

1.4.5 Heterogeneous Effect of Language

So far the specification assumes the topic distributions P_i affect the fees of UIT homogeneously. In this section, we allow for heterogeneous effects on different types of portfolios. Specifically, we estimate the following specification:

$$Fee_i = \beta_0 + Controls_i + \beta_1 Port\_Basics_i + \beta_2 (T_i) + \beta_3 Text\_basics_i + \beta_4 (T_i) \times P_i + \epsilon_i \tag{1.9}$$

Here β_4 is a 20 by 13 matrix of coefficients. The ij'th entry of β_4 measures the correlation between topic j and fee for UITs in group i. Figure 1.8 reports the estimated coefficients for the above mentioned topic with positive average correlation with fees. For each topic, we plot its coefficient in each sub sample of portfolios.

The positive effects of marketing language on fees are most pronounced in certain sector UITs such as R&D, Foreign Finance, and Utility, Tech, and Health. In particular, among R&D UITs, a 1% increase in the proportion document devoted to marketing is associated with an almost 6 basis point increase in annual fees.

As for stock selection, stock weighting, and diversification, the positive average correlation with fees is driven largely by the subsample effects over certain strategy UITs. For example, all three topics have larger than the average correlation with fees among the group of income UITs labeled as $Income_2$.

1.5 Discussion

1.5.1 Why are fees and language correlated?

The results in the previous sections imply how the investment objective is written have significant explanatory power in the variation in fees, even after controlling for all relevant features of the underlying investment product. There are several potential reasons. The first is behavioral. Investors are blinded by the sales pitch that they forgot the underlying portfolio is what matters at the end of the day. This is consistent with the average effects of the discussions on marketing on the fees.

Alternatively, the correlation can be rationalized under an environment with search friction. Consider a sponsor issuing several funds with one flagship fund. The prospectus of the flagship fund likely provides more detailed descriptions of how stocks are selected, more thoroughly reviews the benefits of diversification, and makes a more convincing marketing pitch. The flagship fund can also charge a higher fee because it is likely the first option a prospective investor encounters. This effectively reduces its search cost to investors. Hence investors are willing to pay higher for the flagship fund even if the underlying portfolios are identical. Under this setup, the correlation between language and fees are rooted in the common correlation with the status of the flagship.

1.5.2 Why haven't money managers exploits the effect of language?

If how the prospectus is written can influence how much investors are willing to pay for a fund, why aren't all UIT portfolios described optimally? In particular, if a higher fraction of marketing languages can increase fees charged, we should expect its usage to be much more prevalent and take a larger portion of the investment objective. It is unlikely, though possible, that UIT managers simply have not fully understood the effect of the descriptions themselves.

One explanation is that the fees can only be altered by the investment objectives to a certain degree. While the UIT managers can marginally devote a larger portion of the text on marketing and charge a higher fee, there is an upper bound over which the usage of marketing becomes excessive and lose its effectiveness. After all, the investment objective needs to be about the strategy. It is possible the amount of marketing we observe in the description of each UIT is already optimal given this constraint.

Alternatively, money managers may use different styles of descriptions to achieve price discrimination. Suppose only some UIT investors are influenced by the descriptions while the others are unaffected. With the same portfolio, the UIT managers can retain the latter group with a low fee option with plain description, and exploit the former group with a high fee option with flashy languages.

1.6 Related Literature

There are many papers in the finance literature investigating the roles of text. Antweiler and Frank (2004) shows that the texts extracted from online stock trading boards can significantly predict future trading volumes. Loughran and Mcdonald (2011) shows that sentiment analysis used elsewhere cannot be directly applied to financial reports since words take on different sentiments. Loughran and Mcdonald (2016) provides a survey of recent applications of textual analysis to problems in finance and accounting. Papers closest to ours are Hwang and Kim (2017), who show that closed-end funds with easier-to-read disclosure documents tend to have higher values, and Hillert et al. (2014), who show the tone mutual fund managers use in the letters to their investors significantly affect flow. This paper is the first to look at the texts in investment objectives of funds, and examine its relationship with fees.

This paper provides an explanation of the enormous dispersion in mutual fund fees, as first documented by Hortacsu and Syverson (2004). Under the friction-less model of Berk and Green (2004), the dispersion in fee reflects the discrepancies in money managers' skill. Empirical evidence from Bailey et al. (2011) and others suggest that behavioral biases of investors also play a significant role[8].

[8] Choi et al. (2010) and Beshears et al. (2009) conducted experiments and found investors confused or overlook fees. Bergstresser et al. (2009) shows that broker-sold stocks perform worse but charge higher fees. Cronqvist (2006) and Aydogdu and Wellman (2011) show a significant effect of advertising on fund flows despite no relationship between ad expenditure and fund performance. GENNAIOLI et al. (2015) propose a theory where the trust the money managers provide can also justify a portion of the fees. All of the above are consistent with the behavior bias of investors.

Alternatively, the dispersions in fees can be interpreted as the equilibrium outcome featuring search frictions. Hortacsu and Syverson (2004) structurally estimate a search model and quantify the magnitudes of investor search cost. Carlin (2009) develops a similar model featuring endogenous product complexity and differentiation. Armstrong et al. (2009) in their model show that high-quality funds have the greatest incentive to become prominent, which is associated with the lowest search cost. The results of this paper provide additional support for both the behavioral and the search cost stories.

Methodology wise, this paper relies on the work of Blei et al. (2003) on Latent Dirichlet Allocation. Lee and Seung (1999) first outlines the implementation of non-negative matrix factorization as a way to learn parts of faces and semantic features of texts. To measure fund performance, we rely on the 4-factor model put forth in Carhart (1997). We also use a set of Vanguard index funds as alternatives for benchmark assets as suggested by Berk and van Binsbergen (2015).

1.7 Conclusion

We explore the relationship between UIT sales fees and how the investment objectives are stated. By focusing on UITs and leveraging collaborative filtering and LDA, we can isolate and quantify the relationship between texts and the fees. Under our statistical assumptions, how a UIT is described can affect the fee charged. In particular, more marketing languages, discussion on stock selection, diversification are associated with higher fees. The above results are consistent with investors possessing behavior bias. It also sheds light on the role strategy descriptions play in an environment with search frictions.

While the environment of UITs provides a laboratory where the effect of market timing is eliminated, the external validity concern arises. After all, UITs constitute a rather small and obscure segment of the much larger asset management market. How UIT investors behave may not be representative of the average investor. Information on the identities of the UIT investors may be informative regarding the severity of this concern.

Overall, UITs deserve much more attention from researchers. While this paper constructs the first UIT dataset containing portfolio information, future researches including additional information on size, flow, and investor identity will be better equipped to leverage the its unique features.

1.A Figures and Tables

.

Panel A: Aggregate Statistics

duration		*annual_org_cost*		*expense*		*annual_load*		*net_asset*		
mean	std	mean	std	mean	std	mean	std	mean	std	N
24.15	11.61	0.18	0.09	0.82	0.74	2.02	0.41	164339.03	58533.09	7490

Panel B: Breakdown by trust types

	duration		*annual_org_cost*		*expense*		*annual_load*		*net_asset*		
trust_type	mean	std	mean	std	mean	std	mean	std	mean	std	N
CEF	25.70	10.83	0.15	0.07	1.76	0.34	1.98	0.36	156615.07	11003.81	1744
equity	22.60	10.79	0.19	0.10	0.26	0.14	2.07	0.42	169929.08	74414.34	4396
hybrid	27.16	14.05	0.16	0.08	1.44	0.56	1.94	0.42	156114.39	24002.17	1350

Panel C: breakdown by trust types and sponsor

		duration		*annual_org_cost*		*expense*		*annual_load*		*net_asset*		
sponsor	trust_type	mean	std	mean	std	mean	std	mean	std	mean	std	N
AAM	CEF	22.66	5.75	0.25	0.06	1.96	0.42	2.02	0.25	148934.13	556.30	212
	equity	18.77	4.55	0.30	0.07	0.30	0.04	2.17	0.28	181232.01	120509.57	837
	hybrid	21.94	6.46	0.26	0.07	1.81	0.56	2.06	0.25	148786.14	685.30	205
FT	CEF	26.55	11.68	0.16	0.05	1.73	0.31	1.96	0.38	158520.68	11848.36	1345
	equity	24.05	12.11	0.19	0.08	0.23	0.05	2.03	0.46	167967.64	58137.26	3037
	hybrid	29.23	15.53	0.15	0.06	1.37	0.49	1.89	0.46	157599.32	25411.69	985
Guggen	CEF	23.07	7.29	0.04	0.01	1.74	0.29	2.05	0.27	151616.82	1147.95	187
	equity	20.34	7.25	0.05	0.01	0.38	0.35	2.12	0.29	163217.09	58660.58	522
	hybrid	21.14	5.42	0.05	0.01	1.37	0.76	2.14	0.22	156362.08	28423.55	160

Table 1.1. Summary statistics *Duration* is the number of months between the inception date and the termination date; *annual_org_cost* is the annualized organization cost in percentage; *expense* is the annual expense ratio in percentage; *annual_load* is the annualized max sales charge in percentage; *net_asset* is the total portfolio value indicated on the prospectus. N is the number of trusts. Panel A reports the aggregate mean and standard deviation for all above variables. Panel B reports the same statistics for groups based on trust types. CEF UITs are those investing solely in Closed-End funds; Equity UITs solely in common equities; Hybrid UITs have positions in both CEFs and equities. Panel C provides further breakdown based on the sponsor.

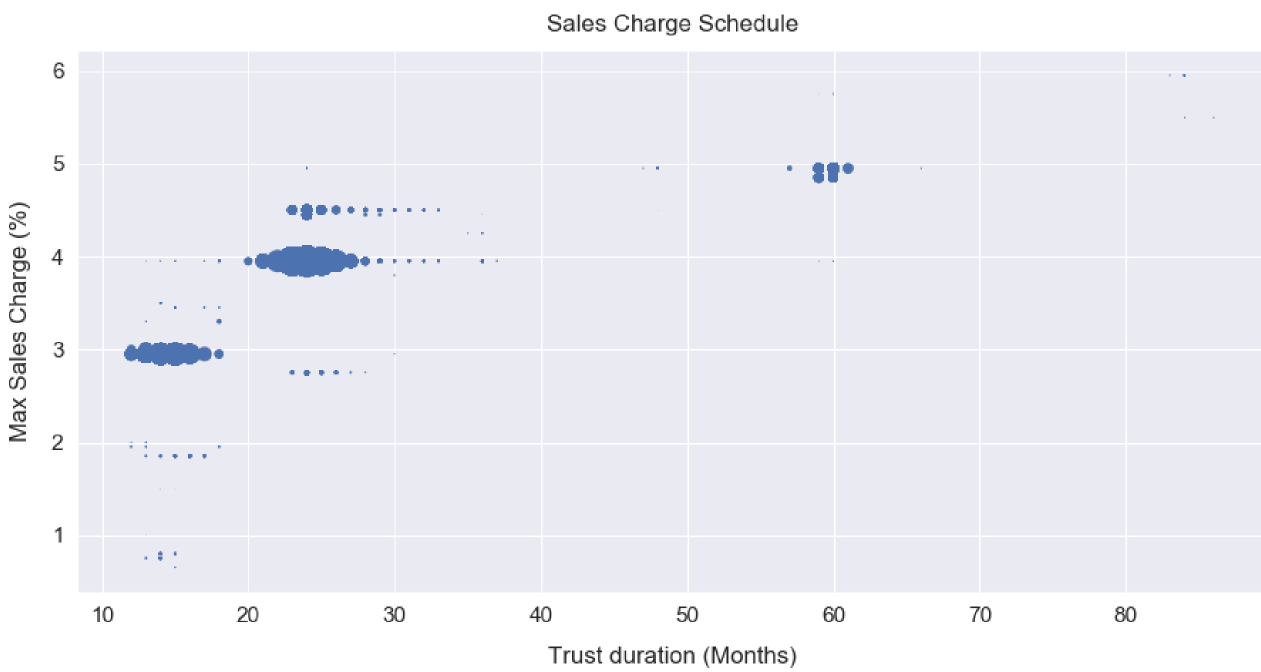


Figure 1.1. Scatter plot of UIT duration and max sales charge Duration is the number of months between the months of inception and termination. Max sales charge is how much the sponsor of the UIT is compensated for the construction of the UIT. The size of the dots indicate the number of UIT with corresponding sales charge and duration. sales fees and duration are highly correlated and form three clusters. UIT with around 15 months duration typically charge total sales fee of 2.95%; UIT with around 24 months duration typically charge total sales fee of 3.95%; UIT with around 60 months duration typically charge total sales fee of 4.95%. Therefore, the shorter the UIT, the higher the effect annualized sales fee.

Panel A: Number of Holdings

	mean	std	min	10%	25%	50%	75%	90%	max
Aggregate	35.25	24.41	5	15	25	30	39	60	187
CEF	23.78	9.43	7	11	16	24	30	36	78
equity	38.30	26.90	5	20	25	30	40	70	187
hybrid	40.15	24.75	8	16	25	32	50	75	140

Panel B: Portfolio Weights distribution

	mean	std	min	10%	25%	50%	75%	90%	max
Aggregate	0.83	1.69	0.00	0.00	0.01	0.06	0.87	2.46	12.61
CEF	1.44	1.88	0.00	0.02	0.18	0.80	1.83	3.74	12.56
equity	0.51	1.49	0.00	0.00	0.00	0.03	0.11	1.41	12.38
hybrid	1.06	1.80	0.00	0.00	0.01	0.33	1.21	3.20	12.61

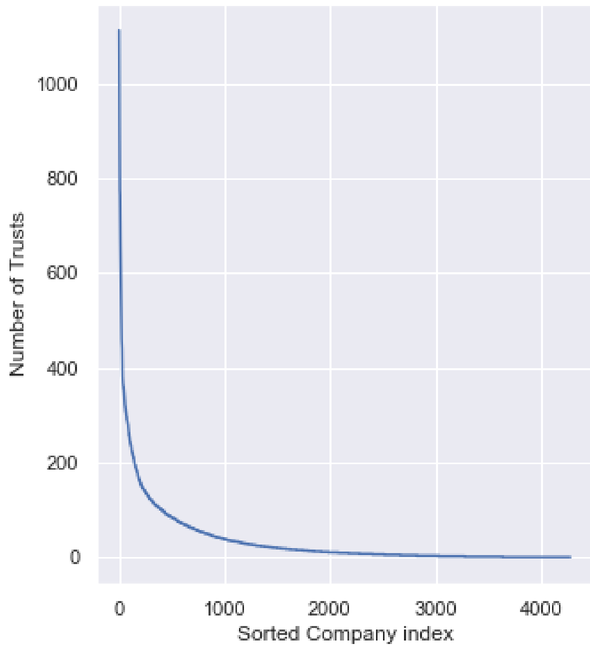
Panel C: Investment Objective Word Count

	mean	std	min	10%	25%	50%	75%	90%	max
Aggregate	451.77	305.23	0	44.50	231	437.00	647.00	773	2559
CEF	398.97	166.42	0	76.00	335	423.00	487.25	601	924
equity	497.62	361.24	0	39.00	157	566.25	719.00	889	2559
hybrid	370.66	198.09	0	76.00	223	386.00	492.50	642	1340

Table 1.2. Portfolio and Text Overview Panel A outlines the number of unique securities each UIT holds. The median UIT holds 30 different securities; Panel B illustrates how portfolio weights are distributed across portfolio companies. The entry measures the max deviation in percentage from $1/N$ allocation. For example, for a portfolio holding 20 companies, the average portfolio weight is 5%. A max deviation of 1% would indicates each securites portfolio weight is between 4% and 6%. We can see that for the median UIT, this is 0.06%, which close to 0. This shows almost all UITs hold close to $1/N$. For reference, the SP500 index fund hold around 500 companies with the biggest position Apple accounting for 4% of the overall portfolio, which yields a 4% max deviation. Panel C describes the number of words used in the investment objective.

Common Equity		Closed-end Funds	
name	count	name	count
Pfizer	1113	Invesco Dynamic Credit Opportunities Fund	382
Chevron	1019	Invesco Senior Income Trust	323
AT&T	1018	Wells Fargo Advantage Income Opportunities Fund	319
Verizon	999	Eaton Vance Limited Duration Income Fund	298
Johnson & Johnson	902	Voya Global Equity Dividend and Premium Opportunity Fund	283
Intel	824	Wells Fargo Multi-Sector Income Fund	271
Microsoft	806	Eaton Vance Tax-Managed Global Diversified Equity Income Fund	267
Exxon Mobil	776	Voya Prime Rate Trust	263
Merck & Company(5)	772	Pioneer Floating Rate Trust	252
Cisco Systems	765	Nuveen Credit Strategies Income Fund	247
Altria Group	722	MFS Intermediate Income Trust	246
Apple	701	Wells Fargo Global Dividend Opportunity Fund	239
JPMorgan Chase & Company	687	Nuveen Floating Rate Income Fund	237
General Electric Company	667	Eaton Vance Municipal Bond Fund	233
ConocoPhillips	647	BlackRock Global Opportunities Equity Trust	232
IBM	607	Eaton Vance Tax-Managed Diversified Equity Income Fund	229
PepsiCo	590	Eaton Vance Senior Floating-Rate Trust	225
McDonald's Corporation	582	Eaton Vance Tax-Managed Global Buy-Write Opportunities Fund	221
Qualcomm	552	BlackRock Corporate High Yield Fund, Inc.	219
Total S.A.	545	Liberty All Star Equity Fund	210

Table 1.3. Companies and CEFs most held by UITs We sort all companies held by any UITs by the number of appearances in UIT. The first two columns list the names and the number of UITs each company appears in. The next two columns are the same for closed-end funds.



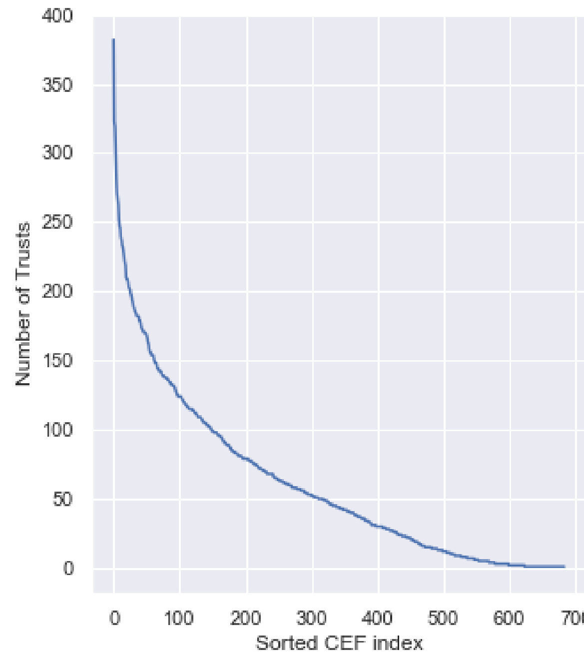


Figure 1.2. Sorted number of appearance in UITs The graph on the left illustrates equities; the graph on the right illustrates the CEF trusts. On the x-axis is the indices corresponding to portfolio companies, sorted by number of appearances in UIT portfolios. For each company, the value on the graph is the number of UITs with positive position in the company. Specifically, 32 companies ($< 1\%$) that are held by more than 400 UITs (10% of equity UIT) . 3287 companies ($> 75\%$) are held by less than 40 UITs. (1% of equity UIT); On the CEF side, 39 (2%) CEFs are held by more than 180 CEF UITs (10% of CEF UIT), 220 (15%) are held by less than 18 CEFs (1 % of total CEF UITs)

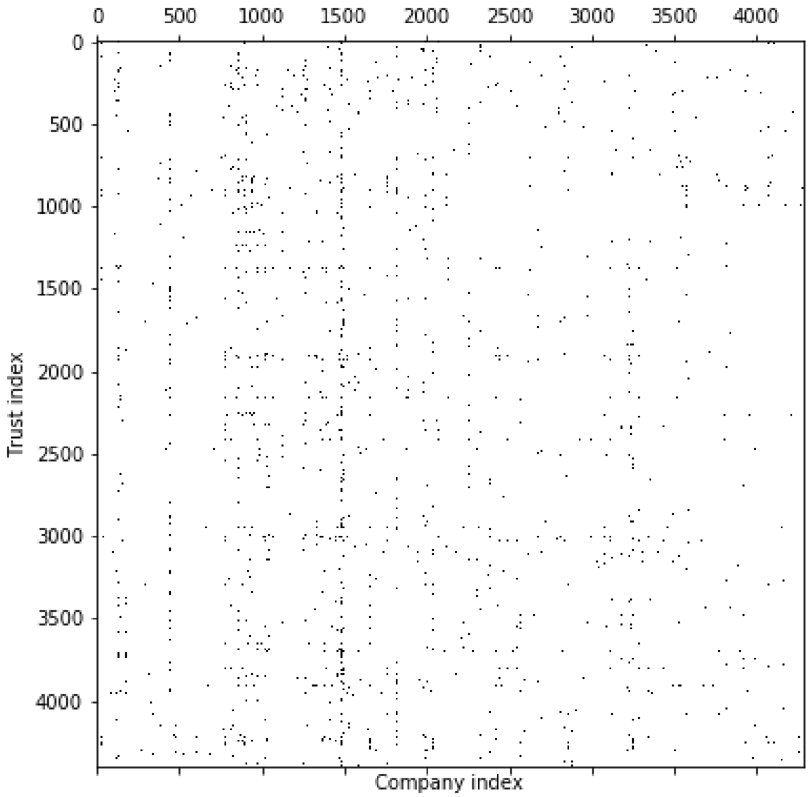


Figure 1.3. Visualization of portfolios of all equity UITs The above graph illustrates the aggregate holding matrix H for all equity UITs. Rows correspond to UITs and the columns represents portfolio companies. Each row is the portfolio for the corresponding UIT. A dot on row i and column j indicates UIT i has a positive position in security j. The graph is sparse since the average UIT holds less than 40 copmanies(out of more than 4000 possible securities). Certain columns are dense relative to others. They correspond to the blue chip companies shown in Table 1.3

	Representative Trust	Representative Company 1	Representative Company 2
1	Lone Star Core Portfolio	Barclays Plc (3	Nippon Telegraph & Telephone
2	Health Care Opportunities Portfolio	Amgen	Gilead Sciences
3	Guggenheim Reit Portfolio	Simon Property Group	Epr Propertie
4	Strategic High 50(R)/Q-50 Dividend And Growth Portfolio	At&T	Philip Morris International
5	Financials Select Portfolio	Jpmorgan Chase & Compan	Wells Fargo & Compan
6	Agriculture Opportunity Portfolio	Cf Industries Holdings	The Mosaic Compan
7	Market Strength Allocation Plus Portfolio	Waddell & Reed Financial	Aflac
8	Advisors Core Equity Strategy Portfolio	Apple Inc. (5)(6	Facebook
9	Equity Income Portfolio	Verizon Communications	Pfizer
10	Target Focus Four Portfolio	Energy East Corporatio	Bce
11	Bny Mellon Brazil, Russia, India And China (Bric) Portfolio	Cnooc Limited (Adr)	China Mobile Limited (Adr)
12	Market Strength Allocation Plus Portfolio	Lyondellbasell Industries Nv (3	The Boeing Compan
13	Global Water Portfolio	Aqua America	Danaher Corporatio
14	Strategic High 50/Q-50 Dividend And Growth Portfolio	Eli Lilly & Compan	Reynolds American
15	Global 100 Dividend Strategy Portfolio	British American Tobacco Pl	Total S.A
16	Technology Select Portfolio	Cisco Systems Inc	Microsoft Corporatio
17	Income Dividend Equity Allocation (Idea) Portfolio	Pepsico	Johnson & Johnso
18	Utilities Portfolio	Public Service Enterprise Group	American Electric Power Company
19	Tarantula Portfolio	Nvidia Corporatio	Moody'S Corporatio
20	Energy Portfolio	Chevron Corporatio	Exxon Mobil Corporatio

Table 1.4. Interpreting trained latent dimensions - equity UITs The above helps to understand the results of Matrix factorization. We project each trust and each company into a 20 dimensional vector. For each dimension, we sort all UITs and companies based on the loadings of their latent representations on that dimension. Each row lists the names of the top UIT and the top 2 companies.

	Representative Trust	Representative CEF
1	U.S. Equity Closed-End Portfolio	Liberty All Star Equity Fun
2	Diversified Income Wave Portfolio (15-Month)	Allianzgi Nfj Dividend Interest & Premium Strategy Fun
3	Global Bond Income Closed-End Portfolio	Western Asset Emerging Markets Debt Fund
4	Senior Loan & Limited Duration Plus Closed-End Portfolio	Barings Global Short Duration High Yield Fun
5	High-Yield Income Closed-End Portfolio	Wells Fargo Advantage Income Opportunities Fun
6	Municipal Income Opportunities Closed-End Portfolio	Putnam Managed Municipal Income Trus
7	Strategic Income Advantage Select, Closed-End Portfolio	Allianzgi Equity & Convertible Income Fun
8	California Municipal Income Plus Closed-End Portfolio	Blackrock Muniyield California Quality Fund
9	Senior Loan Select Closed-End Portfolio	Pioneer Floating Rate Trus
10	Municipal Income Select, Closed-End Portfolio	Putnam Managed Municipal Income Trus
11	Senior Loan & Limited Duration Closed-End Portfolio	Eaton Vance Limited Duration Income Fun
12	Tax-Advantaged Municipal Closed-End Portfolio	Blackrock Municipal Income Investment Quality Trus
13	New York Municipal Income Closed-End Portfolio	Nuveen New York Dividend Advantage Municipal Fun
14	Insured Municipal Income Select, Closed-End Portfolio	Blackrock Muniyield Quality Fund Ii
15	Covered Call Select Closed-End Portfolio	Nuveen S&P 500 Buy-Write Income Fun
16	New Jersey Municipal Income Select Closed-End Portfolio Serie	Blackrock Muniholdings New Jersey Quality Fund
17	Municipal Income Plus Closed-End Portfolio	Nuveen Enhanced Municipal Credit Opportunities Fun
18	Strategic Income Portfolio (15-Month)	John Hancock Tax-Advantaged Dividend Income Fun
19	Global Equity Income Closed-End Portfolio	Eaton Vance Tax-Managed Global Diversified Equity Income Fun
20	Investment Grade Select Closed-End Portfolio	Alliancebernstein Income Fund

Table 1.5. Interpreting trained latent dimensions - CEF UITs The above table is constructed in similar fashion to 1.4, done with CEF UITs. Due to space constraints, we only report the top representative CEF.

Figure 1.4. Word Clouds of LDA Topics The above are the visualizations of the results from LDA estimation. Each word cloud selects the top 15 associated with the given topic. The size of the word reflects the the conditional probability of the word being drawn from the particular topic. For example, the word *best* is a strong indicator of the topic named *marketing*. We use both the word cloud and the representative documents in the Apppendix to infer the meaning of each topic.

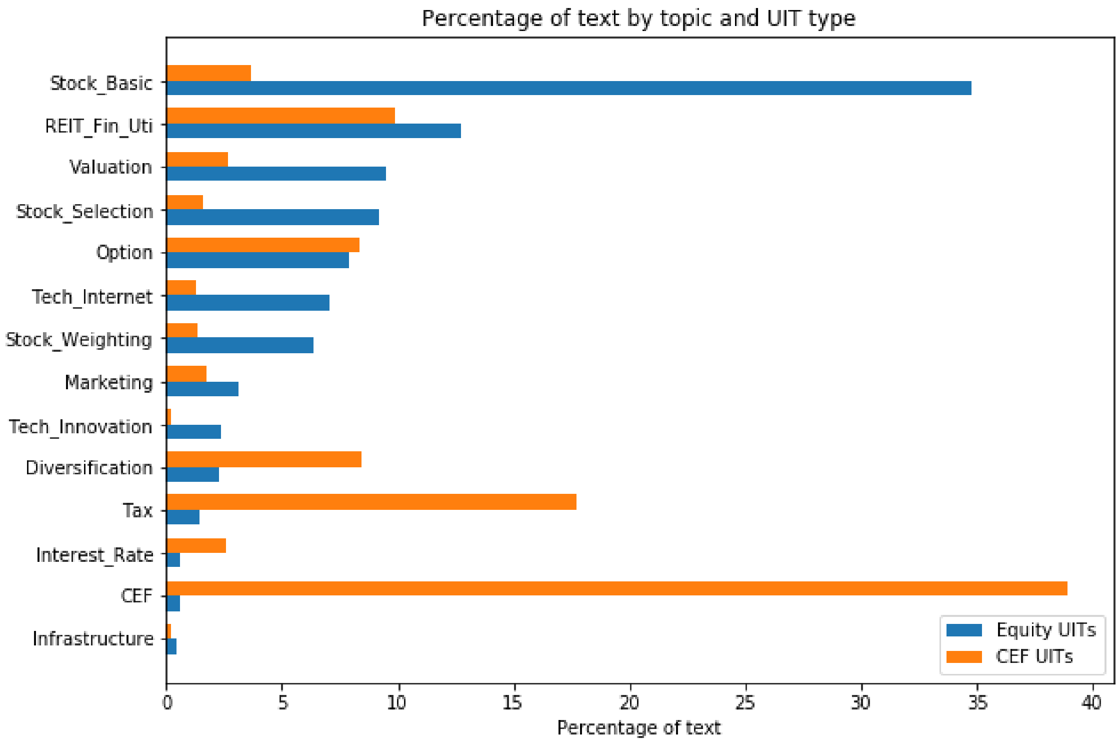


Figure 1.5. Average Topic Distributions for equity trusts and CEF Trusts The above bar graph plots the average topic distribution for both equity trusts and CEF trusts. The topics correspond to those in Figure 1.4

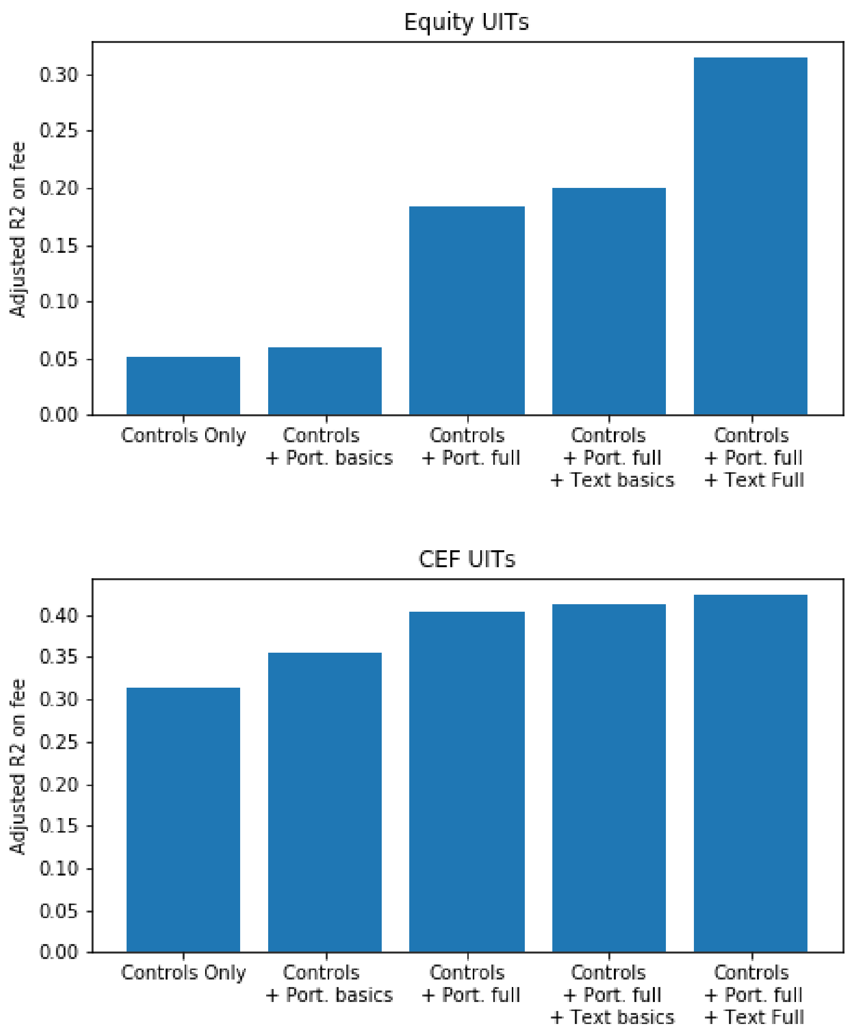


Figure 1.6. Adjusted R^2 when regressing fee on portfolio and text characteristics The topic graph plots the adjusted R^2 when regressing UIT fees over equity UITs. The label indicates the independent variables. Controls include sponsor and year fixed effects. Port. basic inlcude number of holdings and the portfolio weights; Port. Full is the clusters constructed from the output of Matrix Factorization on portfolio; Text. Basic is the word count of investment objective; Text Full is the topic distribution learned from LDA. The bottom graph do the same for CET UITs

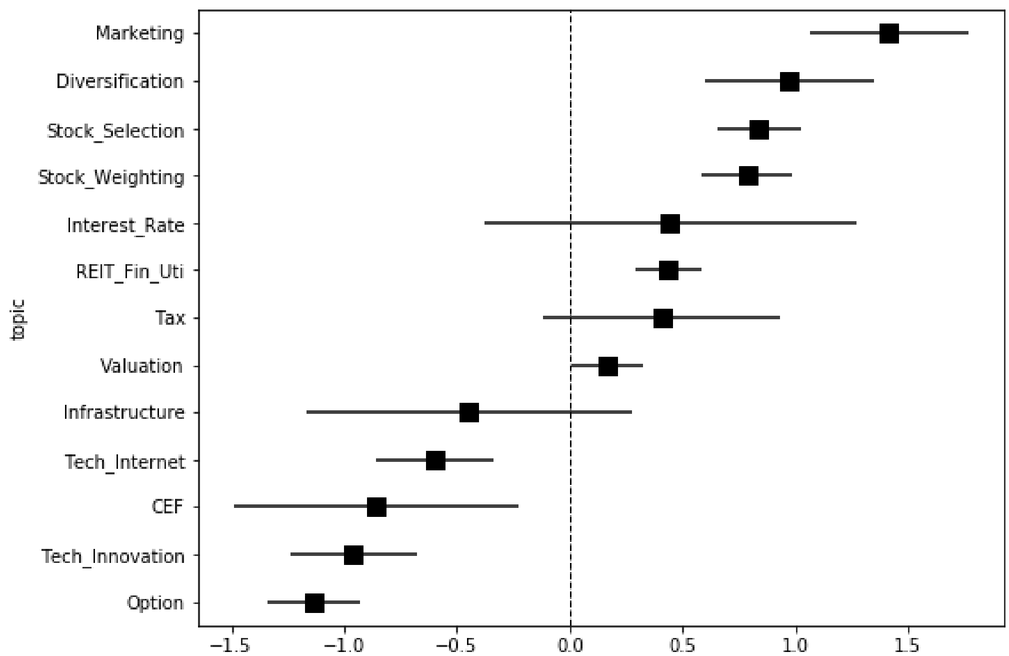


Figure 1.7. Effects on fee for different topic probabilities Each line plots the point estimate along with the confidence interval for the coefficient corresponding to the label. For example, the first row means that an 1% increase in the amount spent discussing the marketing is associated with a 1.41 basis increase in fee.

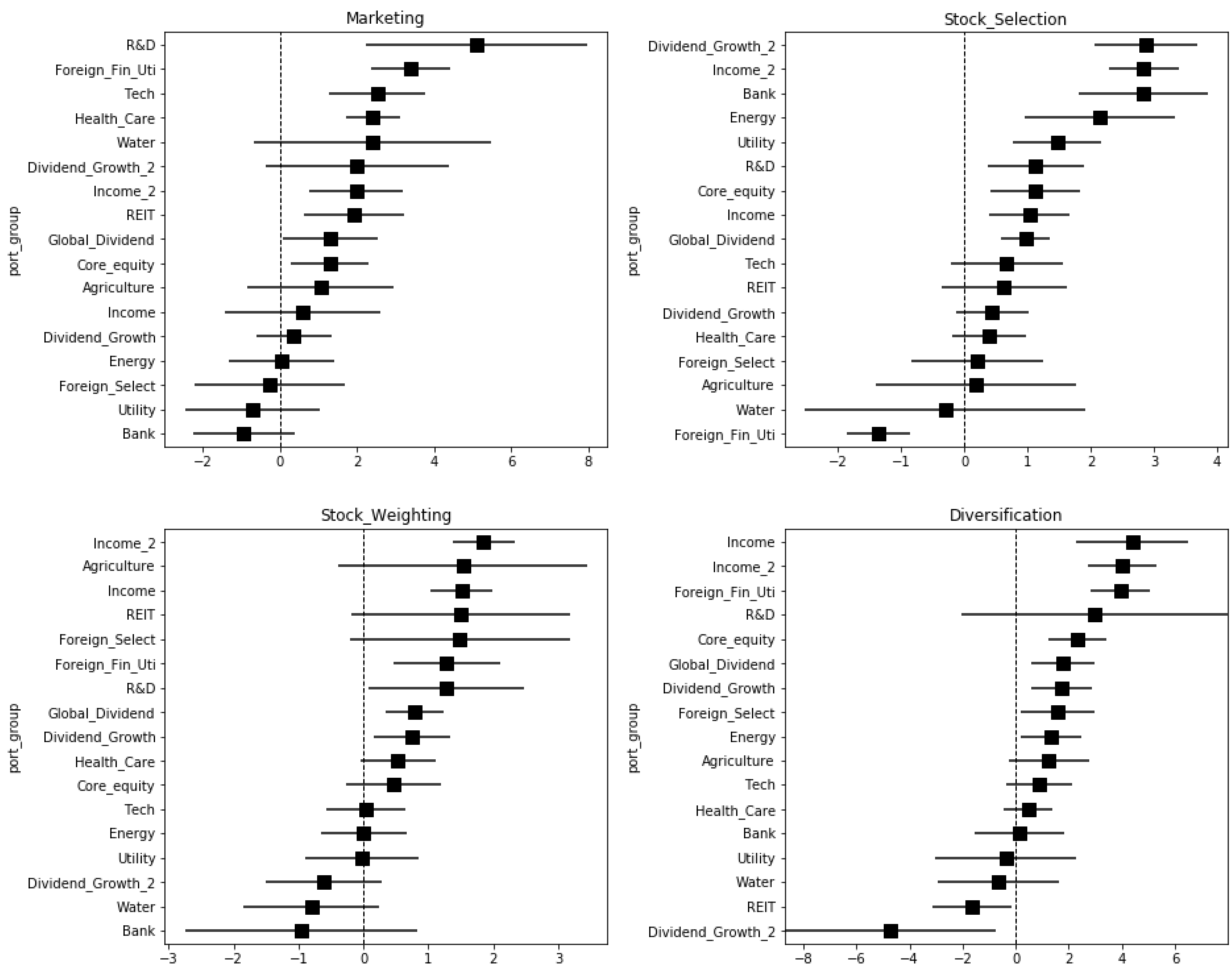


Figure 1.8. Effects on fee for different topic probabilities for various groups of portfolios The graphs should be read similarly as Fig 1.7. The title of each subplots represents the language topic; the y labels correspond to different portfolio groups.

1.B Sample investment objective

Topic 1: Market Cap and Growth vs. Value

The objective of the Trust is to provide investors the potential for above-average capital appreciation by investing in equity securities.

Based on the composition of the portfolio on the Initial Date of Deposit, the Trust is considered a **Large-Cap** Blend Trust.

From time to time in the prospectus or in marketing materials we may identify a portfolio's style and capitalization characteristics to describe a trust. These characteristics are designed to help you better understand how the Trust fits into your overall investment plan. These characteristics are determined by the Sponsor as of the Initial Date of Deposit and, due to changes in the value of the Securities, may vary thereafter. In addition, from time to time, analysts and research professionals may apply different criteria to determine a Security's style and **capitalization** characteristics, which may result in designations which differ from those arrived at by the Sponsor. In general, **growth stocks** are those with high relative price-to-book ratios while value stocks are those with low relative **price-to-book** ratios. At least 65% of the stocks in a trust on the trust's initial date of deposit must fall into either the **growth or value** category for a trust itself to receive the designation. Trusts that do not meet this criteria are designated as blend trusts. Both the weighted average **market capitalization** of a trust and at least half of the Securities in a trust must fall into the following ranges to determine its **market capitalization** designation: **Small-Cap-less than $2.5 billion; Mid-Cap- $2.5 billion to $10 billion; Large-Cap-over $10 billion** . Trusts, however, may contain individual stocks that do not fall into its stated style or market capitalization designation. Of course, as with any similar investments, there can be no assurance that the objective of the Trust will be achieved. See "Risk Factors" for a discussion of the risks of investing in the Trust.

Topic 2: Real Estate Investment Trusts (REIT), finance, and utility

The sponsor, with the assistance of Guggenheim Partners Investment Management, LLC ("GPIM"), an affiliate of the sponsor and Guggenheim Partners, LLC, has selected the securities to be included in the trust's portfolio The U.S.-listed common stocks held by the trust may include the common stocks of U.S. and non-U.S. companies. The trust may include securities of **real estate investment trusts ("REITs")** . As a result of the strategy, the trust is concentrated in the **financial sector** and invests significantly in the **utilities sector**.

Topic 3: Stock Selection Criterion

When you invest in a Trust you are purchasing a quality portfolio of attractive common stocks in one convenient purchase. Each Trust seeks above-average total return. To achieve this objective, each Trust will invest in the common stocks of companies which are **selected by applying a unique specialized strategy**. While the Trusts seek above-average total return, each follows a different investment strategy. We cannot guarantee that a Trust will achieve its objective or that a Trust will make money once expenses are deducted.

The Dow(R) Target 5 Portfolio invests in stocks with high dividend yields. **By selecting stocks** with the highest dividend yields, The Dow(R) Target 5 Strategy seeks to uncover stocks that may be out of favor or undervalued. Investing in stocks with high dividend yields may be effective in achieving the investment objective of the Trust, because regular dividends are common for established companies, and dividends have historically accounted for a large portion of the total return on stocks. The Dow(R) Target 5 Strategy seeks to amplify this dividend yield strategy by selecting the five lowest priced stocks of the 10 highest dividend-yielding stocks in the Dow Jones Industrial Average ("DJIA(R)").

The Dow(R) Target 5 Strategy stocks are determined as follows:

- **Step 1: We rank all 30 stocks contained in the DJIA(R) by their current indicated dividend yield as of the business day prior to the date of this prospectus.**
- **Step 2: We then select the 10 highest dividend-yielding stocks from this group.**
- **Step 3: From the 10 stocks selected in Step 2, we select an equally-weighted portfolio of the five stocks with the lowest per share stock price for The Dow(R) Target 5 Strategy.**

Topic 4: Diversification

The objective of the Trust is to provide investors with the potential for a high rate of current monthly income, with capital appreciation as a secondary objective. In addition, based on current publicly available information, none of the closed-end funds selected for the portfolio are reporting the use of leverage. Closed-end funds which employ leverage are more volatile than those that do not. However, certain or all of these Closed-End Funds may have utilized leverage in the past and may elect to utilize leverage in the future.

The Multi-Sector Approach.

The Trust is designed to enable investors who are seeking a high rate of current monthly income to reduce some of the volatility typically associated with high-income investments. To accomplish this, the portfolio is **diversified across a broad range** of closed-end funds including convertible securities funds, general bond and equity funds, high-yield bond funds, investment grade bond funds, specialized equity funds, and world equity and income funds. Because different sectors follow different cycles and react differently to changes in global economies and interest rates, spreading assets across this spectrum of closed-end funds has the potential to reduce the overall risk of the portfolio.

Closed-End Advantages.

- - Portfolio Control. Unlike open-end mutual funds, closed-end funds maintain a relatively fixed pool of investment capital. This allows portfolio managers to better adhere to their investment philosophies through greater flexibility and control. In addition, closed-end funds don't have to manage fund liquidity to meet potentially large redemptions.

- - Diversification. The portfolio offers investors diversification by investing in a broad range of closed-end funds that are further diversified across hundreds of individual securities. Diversification does not guarantee a profit or protect against loss.

- - Income Distributions. Because they are not subjected to cash inflows and outflows, which can dilute income distributions over time, closed- end funds can generally provide a more stable income stream than other managed fixed-income investment products. However, stable income cannot be assured.

Topic 5: Equal weighting

The trust seeks to provide total return, through dividend income and capital appreciation, by investing in a portfolio consisting of the 10 stocks with the highest dividend yield chosen from among the 30 stocks in the Dow Jones Industrial Average(SM) ("DJIA(SM)") as of February

16, 2011. **The stocks are held in approximately equal dollar amounts as of the trust's inception.**

Please note that the strategy was applied to select the portfolio at a particular time. If we* create additional units of the trust after the trust's inception date, the trust will purchase the securities originally selected by applying the strategy. This is true even if a later application of the strategy would have resulted in the selection of different securities. We disregard any stock in the DJIA(SM) where the U.S. government or any of its agencies can exercise any veto power over any current or future dividends, whether common or preferred, to be distributed in connection with such stock. In such cases, we will automatically select the next highest dividend yielding stock in the DJIA(SM). In addition, companies which, based on publicly available information as of two business days prior to the date of this prospectus, are the target of an announced business acquisition which are expected to happen within six months of the date of this prospectus have been excluded from the universe of securities from which the trust's securities are selected.

Topic 6: Balance sheet and valuation

Under normal circumstances, the trust will invest at least 80% of the value of its assets in common stocks of European companies. The trust invests in a diversified portfolio of 30 European companies that the sponsor believes have had **strong valuations**, returns on capital and **balance sheets**. To determine whether a company has an **attractive valuation**, the sponsor compares valuation metrics against the selected company's peer group. Strong returns on capital are evidenced by the company's return on capital compared to the selected company's peer group. Companies with strong balance sheets are typically those entities that are **less levered than their peers.** The trust's investment process is designed to favor strong cash flow generating companies that trade at reasonable multiples of their excess profits. However, their can be no assurance that the trust's investment strategy will identify companies that will perform well in the future. See "Investment Policies" in Part B of the prospectus for more information.

Topic 7: Options

The objective of **Covered Call Closed-End Portfolio**, Series 3 is to provide investors with the potential for income and, to a lesser extent, capital appreciation.

Covered Call Closed-End Portfolio, Series 3 invests in Closed-End Funds whose investments are diversified among common stocks and, on an ongoing and consistent basis, will sell **covered call options** to seek a more controlled risk/reward outcome.

What is a **Covered Call Option**? A call option is a contractual obligation which gives the buyer of the option the right to purchase a certain number of shares of common stock from the writer (seller) of the option at a predetermined price. If the predetermined price is reached, the buyer has the right, depending on the type of option, to exercise the option at the option's expiration date or at any time up until the option's expiration.

An option is considered "covered" when a Closed-End Fund owns the equity securities against which the options are sold. You should be aware that a product which includes Closed-End Funds utilizing covered call strategies may not be suitable for all investors. It may not be appropriate for investors seeking preservation of capital or high current income. Before investing, you should make sure you understand the risks of this type of product, and whether it suits your current financial objectives.

Benefits of Using Options.

Though call options can be used for many investment purposes, they are typically used as a tool to potentially enhance returns, offer a current yield to investors, and provide limited downside protection. A comparison of the Standard and Poor's ("S&P") 500 Index to the CBOE S&P 500 BuyWrite Index ("BXM Index") can illustrate how an options strategy could work.

The S&P 500 Index is an unmanaged index of 500 stocks used to measure large-cap U.S. stock market performance. The BXM Index is a passive total return index maintained by the Chicago Board Options Exchange which measures the total return of a hypothetical covered call strategy applied to the S&P 500 Index. The BXM Index consists of a hypothetical "long" position in the S&P 500 Index on which are deemed written a succession of **one-month, at-the-money call options.**

Topic 8: Closed-end fund

The Trust seeks a high rate of current monthly income, with capital appreciation as a secondary objective. One thing that has not changed over the years is the need for some investors to earn high current income. Investors willing to assume certain credit and market risks have the potential to earn a high level of current monthly income by investing in high-yield bonds. The **High-Yield Income Closed-End Portfolio** is comprised of a well-diversified pool of **closed-end funds** that invest in U.S. and foreign high-yield bonds.

Although subject to greater risks, high-yield bond investors have historically received greater returns from their high-yield investments than investment grade bond investors. Of course, there can be no assurance that high-yield securities will outperform investment grade bonds in the future or that the default rate on high-yield securities will not rise.

Closed-End Fund Selection. The Closed-End Funds were selected by our research department based on a number of factors including, but not limited to, the size and liquidity of the Closed-End Fund, the current dividend yield of the Closed-End Fund, the quality and character of the securities held by the Closed-End Fund, and the expense ratio of the Closed-End Fund, while attempting to limit the overlap of the securities held by the Closed-End Funds.

Closed-End Features.

- Income Distributions. Closed-end funds are structured to generally provide a more stable income stream than other managed fixed-income investment products because they are not subjected to cash inflows and outflows, which can dilute dividends over time. However, as a result of bond calls, redemptions and advanced refundings, which can dilute a fund's income, the Trust cannot guarantee consistent income. Although one of the Trust's objectives is to seek a high rate of current monthly income, there is no assurance the objective will be met.

- Portfolio Control. Since closed-end funds maintain a relatively fixed pool of investment capital, portfolio managers are better able to adhere to their investment philosophies through greater flexibility and control. In addition, closed-end funds don't have to manage fund liquidity to meet potentially large redemptions.

Topic 9: Bond Market, inflation, and interest rates

The Trust seeks above-average total return. Because of an improving economy and the possibility of **rising inflation**, we believe there is the potential for rising **interest rates** in the coming years. Bonds tend to lose value in a rising interest rate environment. Equities, on the other hand, have often historically provided positive returns after the **Federal Reserve's initial rate hike**. Like stock returns, economic growth, and inflation, interest rates are one of those variables that you can't control. But, as an investor, you can control how your investment dollars are allocated.

The **Interest Rate Hedge Portfolio** is a professionally-selected unit investment trust which invests in common stocks of companies that have a history of dividend growth, as well as closed-end funds ("CEFs") which invest in convertible securities, **Treasury Inflation Protected**

Securities ("TIPS"), master limited partnerships ("MLPs"), limited duration bonds and real estate investment trusts ("REITs"). Of course, there can be no assurance that the Trust will achieve a positive return if interest rates rise.

Topic 10: Technology - I

The Trust seeks above-average capital appreciation. **Social media** refers to interaction among individuals in which they create, share and exchange information and ideas in virtual communities and networks. The **Internet** is now used by hundreds of millions of people worldwide and has become something upon which we rely. Social media sites have changed the way entities circulate information and **communicate on a global scale.**

The Trust invests in **Internet retail** and **Internet software** and services companies that we believe will benefit as social media becomes more universal.

Topic 11: Technology II

Under normal circumstances, the trust will invest at least 80% of the value of its assets in companies that the sponsor believes are **technologically innovative**. Technologically innovative companies are innovative companies that create new technology or use current technology in a new way to create new growth opportunities. These technological innovations are being **engineered and applied** in multiple sectors of the economy to improve corporate profit, marketing and operations as well as the efficiency, productivity and enjoyment in the daily lives of individuals.

The trust consists of companies that the sponsor feels have demonstrated a capacity to be innovative or are involved in an industry that has been involved in **technological advances in various fields, including, but not limited to, health sciences, global security, information access and manufacturing improvements.** The U.S.-listed common stocks held by the trust may include the common stocks of U.S. and non-U.S. companies.

Topic 12: Tax

The Trust seeks monthly income that is **exempt from federal income taxes** by investing in a well-diversified portfolio of common stocks of Closed-End Funds that invest in **municipal bonds.** However, certain distributions paid by certain funds may be **subject to federal income taxes**. In addition, a portion of the income may be subject to the **alternative minimum tax.**

Americans deal with a number of **different taxes** in their everyday lives. Perhaps none are more noticeable than **individual income taxes**. In fact, individual income taxes comprise the largest component of Americans' tax bill. On average, Americans had to work a full 36 days in the year 2011 just to earn enough money to pay for them.

Tax Freedom Day, the day on which Americans have earned enough money to pay all federal, state and local taxes for the year, was three days later in 2011 than in 2010. Americans paid more in taxes in 2011 than they spent on food, clothing, and shelter combined. These examples are based on an overall average tax rate for the nation which is done by dividing the nation's total tax payments by the nation's income as projected by the Tax Foundation for 2011.

Topic 13: Expertise, Optimism, persuasion

The Trust seeks above-average capital appreciation.

Sabrient Systems, LLC ("Sabrient") is an independent equity research firm that builds**powerful investment strategies** by using a fundamentals-based, quantitative approach. The strategies are used to create **rankings and ratings** on more than 7,000 stocks, indices, sectors, and ETFs. Their models are designed to identify those companies that are anticipated to outperform or underperform the market.

Historically, certain stocks have been able to outperform the market during periods in which the market itself is declining. The Trust is a unit investment trust that seeks to find companies that Sabrient **believes are positioned to perform well** in environments of falling stock prices but also those companies that have the potential to provide solid performance in rising markets. The stocks in the portfolio are selected by **applying a seven-step investment strategy process developed by Sabrient.**

Topic 14: Infrastructure

The Trust seeks above-average capital appreciation by investing in companies that provide products or services that aid the **development and maintenance of our country's infrastructure system** as well as companies that tend to benefit from an increasing economic growth rate.

A sound infrastructure is fundamental to the quality of life and the economic growth and development of a community. **Infrastructure assets** help provide products and services that have become essential to everyday life such as **water and wastewater systems, roads, railways, airports, utilities, hospitals, schools and communications systems.** Developing countries realize that infrastructure will play a key role in economic development and in their ability to attain projected growth rates. However, the need to improve infrastructure extends beyond rapidly developing countries. The United States, while considered to be a world leader in many aspects, lags other developed regions, especially Europe, when it comes to the condition of its aging infrastructure.

The Aging of America.

Infrastructure funding gaps in the United States are immense in comparison to the rest of the world. Total investment needs for America's infrastructure are estimated to be approximately $3.6 trillion by 2020. Our aging infrastructure comes with additional costs in terms of economic growth.

CHAPTER 2. Team Production in Money Management

2.1 Introduction

Institutional investors play a vital role in the money management industry: collectively, institutions held 68 percent of the US equity in 2007 [1]. In addition to managing the majority of household wealth, institutions also commonly delegate investment decisions among themselves. Entities such as endowments, pension funds, and foundations, on average delegate \$36 trillion in assets to external asset management firms between 2000 and 2012, which represents 29% of worldwide investable capital. [2]. This paper seeks to explore the dynamics among institutions where they serve both as money managers and their clients.

To avoid confusion, in this paper, the word *institution* is used only when referring to the likes of pension funds that delegate asset management to other institutions. We will use *asset managers, money managers, managers* interchangeably when referring to asset management firms managing assets of institutions.

Delegation to asset managers from institutions typically takes place via institutional funds. Unlike retail mutual funds, institutional funds are specialized investment vehicles constructed for only one or a few clients. Another salient feature of institutional funds is that clients have investment discretion over the specific holdings of the funds. Such feature suggests these institutional clients have an edge in investment.[3] The following quote taken from the investment objective of Princo (Princeton University Investment Company) explicitly points out the value-added of institutions as the clients of institutional funds:

> "Princo partners with best-in-class investment management firms across the globe and in diverse asset categories . . . our strong partnership with managers has led to additional value creation."

[1] Computed by Lewellen (2011) using 13F filings.
[2] Gerakos et al. (2016)
[3] Sometimes institutions also impose constraints on portfolios to comply with certain regulations.

However, recent empirical findings seemingly reject the notion that institutions have an edge in investment: while institutional funds can outperform corresponding strategy benchmarks (Gerakos et al. (2016)), Lewellen (2011) shows that for the capital managed in-house by institutions themselves, there is no evidence of such outperformance. If institutions do have an edge in investment, why does it not show up when they invest without the help of asset managers?

We consider a model that is able to answer the above question. In the model, some institutions have private information that can lead to outperformance. In order to generate the observed outperformance regarding delegated institutional capital through institutional funds and the lack of outperformance regarding capital managed in-house by institutions themselves, two ingredients are needed: (a) complementarity between institution's information and asset manager's investment skill, and (b) costly search of skilled managers.

The skill of money managers is better utilized by institutions with private information than by institutions without. As such, informed institutions are willing to pay a higher search cost for skilled management. Search frictions then act as a separating mechanism in equilibrium: only informed institutions choose to invest through the institutional funds, whereas institutions without private information find the search cost too high and invest on their own instead. As a result, all clients of institutional funds have private information, hence their investment discretion is justified. The outperformance of institutional funds is the team production of both institutions and asset managers. On the other hand, institutions that invest in-house are precisely the ones with no investment expertise, hence they do not beat the benchmark. In other words, if one concludes from the lack of in-house outperformance that institutions on average are not skilled, he is ignoring the underlying selection by institutions on whether to invest via institutional funds[4].

We then use the model as a novel lens to examine other questions in money management using comparative statics. Specifically, (1) What determines the management fee? (2) What determines

[4]Instead of informed institutions vs. uninformed institution, another interpretation is the specific areas the institutions are experts in. A common investment approach of institutions is to first determine strategy level allocation. Then for each strategy, the institutions decide whether to outsource or to manage in house. The model result under this setting is that institutions delegate portions of their portfolio which that have the most private information. They do so because the asset managers can make the best use of that information

the equilibrium number of skilled money managers? Important to our model, relative to previous models on money management, is that investors (in this case, the institutions) play an active role. Therefore in answering the above questions, we focus on implications of change in information allocation across institutions, both in quantity (number of institutions with private information), and in quality (how precise the private information is).

First consider the case when each institution becomes more informed, holding fixed the number of informed institutions. As a result, what institutions can earn on their own increases. Therefore the money managers need to lower the management fee in order to retain their clients. it follows that there are fewer skill money managers due to lower fee revenue.

When there are more informed institutions, management fees also drop. Since each informed institution invests through skilled asset managers, there is more trading on managers' information. Therefore, price becomes more reflective of managers' signals. Consequently, managers are forced to charge a lower fee. The number of skilled managers, however, can either increase or decrease. While the fee rate decreases, there are also more institutions paying the fee. So the net effect on total fee revenue is ambiguous.

Lastly, we consider efficiency implication of institutions' information allocation. Specifically, is price more informative when there are many moderately informed institutions or a few highly informed institutions? We show that under team production between institutions and managers, the quantity has an amplified effect on price informativeness. Therefore, price is more informative when institutions' private information is more diffused.

There is an abundance of literature, both empirical and theoretical, that explore whether skill in investment exists and what skill is[5]. Berk and Green (2004) use rational arguments featuring learning and competition among investors to reconcile lack of performance persistence and fund flows that follow performance.

[5] Examples of evidences against the existences of investment skill are Carhart (1997), French (2008), FAMA and FRENCH (2010). Examples in favor of the existences of investment skill include Berk and Green (2004) and Berk and van Binsbergen (2015). Papers investigating the details of investment skills include BERK et al. (2017), Kacperczyk et al. (2014), Kacperczyk et al. (2016), Pastor et al. (2017), and Manconi and Spalt (2012).

Regarding model mechanics, this paper is closely related to Pedersen and Garleanu (2015), who first considers embedding an asset market equilibrium similar to Grossman and Stiglitz (1980) in a matching equilibrium of the market of asset managers. Our model's main distinction is the investor's private information. Doing so allows us to explain the puzzle laid out in the beginning. It also allows for analysis of changes in investors' information later on.

While in this model only focuses on equilibriums where complementarity exists between investors and managers, Admati and Pfleiderer (1987) provides the first systematic treatment on when complementarity and substitutability occurs. Another similar paper is Goldstein and Yang (2015), who also consider the model with one risky asset and two signals. Their focus is on how the realization of the signal affects one's trading of the asset. In their setup, complementarity is imposed via a functional form. For a more comprehensive survey of similar models, see Veldkamp (2011).

Lastly, this paper provides insights into the chain of money management. Concretely, consider a government employee, who invests his saving via a pension fund, who then invest in an asset manager, who eventually invest in stocks, the model speaks to the role pension funds play as the middle layer – both in providing additional investment insights and connecting parties with better skilled. Other papers featuring chains of intermediation includes Glode (to include), who consider a sequence of bilateral trades, each of which reduces asymmetric information by a little.

The rest of the paper is organized as follows. Section 2.2 describes the model setup and defines the equilibrium. Section 2.3 focuses on how complementarity and search friction generate the observed pattern on outperformance. Section 2.4 conducts comparative statics. Section 2.5 deals with price informativeness. Section 2.6 concludes.

2.2 Model

2.2.1 Setup

We builds upon the model of Grossman and Stiglitz (1980) featuring noisy rational expectation equilibrium. The model is static. There is one risky asset with random payoff given by

$$\tilde{v} \sim N\left(0, \gamma_0^{-1}\right) \tag{2.1}$$

where γ_0 captures the prior precision shared among all agents in the economy.

There is a continuum of money managers (henceforth manager for short) with mass $\overline{M}$. All managers are ex-ante identical with risk neutral preference. Each manager can pay cost κ to gain access to a signal on asset payoff. The signal is given by

$$\tilde{s}_m = \tilde{v} + \tilde{\epsilon}_m, \qquad \tilde{\epsilon}_m \sim N\left(0, \gamma_m^{-1}\right) \tag{2.2}$$

Managers who pay for the signal is said to have acquired skill. We refer them as skilled managers. Here the signal s_m is common across all skilled managers. We can interpret manager skill as the expertise in processing publicly available data and generating deeper insights. We denote the mass of skilled manager as M, which is determined in equilibrium.

There is a continuum of institutional investors with identical initial wealth $\overline{W}$ and CARA utility[6]:

$$U(W) = -\frac{1}{\alpha}\log(E(e^{-\alpha W})). \tag{2.3}$$

Investors are not endowed with the same information set. Some investors have access to the signal

$$\tilde{s}_i = \tilde{v} + \tilde{\epsilon}_i, \qquad \tilde{\epsilon}_i \sim N\left(0, \gamma_i^{-1}\right) \tag{2.4}$$

s_i has the same normal structure as s_m, with potentially different precision γ_i [7]. Note that s_i is the same for all investors[8]. Furthermore, we assume $\tilde{\epsilon}_i$ and $\tilde{\epsilon}_m$ are independent. We refer

[6] With the normalization by α and the log transformation, $U(\tilde{W})$ is equal to the certainty equivalent of $\tilde{W}$

[7] As we will see later, The difference in precision allows for study of how changes in quality of investor's information can impact managers response.

[8] In this paper, the subscript i is short for investor, or institutional investor. This is to be distinguished from its common usage as an index.

to investors endowed with $\tilde{s}_i$ as informed investors, and rest of the investors as uninformed. Let λ_i and λ_u denote the mass of informed investors and uninformed, respectively. For simplicity, we assume λ_i and λ_u are exogeneous[9].

Both informed and uninformed investors have the option to invest with a skilled manager. However, screening of managers and the related due diligence on asset managers are costly. This is modeled by the search cost that all investors face:

$$c(M) = \bar{c}M^{-\psi} \tag{2.5}$$

where $\psi > 0$. Following the literature on searching, the search cost is increasing in M. Naturally, as there are more skilled managers, it becomes less costly to search for them. Let $I_i \in [0, \lambda_i], I_u \in [0, \lambda_u]$ denote the mass of informed and uninformed investors that search for skilled management. I_i, I_m are determined in equilibrium, as is M.

Paying the search cost, $c(M)$, allows the investor to find a skilled manager. When investing with an skilled manager, the investor first disclose her own information s_i (if she is informed). The manager then combines any information shared from the investor to his own signal s_m, and fulfills his fiduciary duty by investing each clients asset exactly according to their' CARA utility. As such, from the investor's perspective, searching and investing through a skilled manager is equivalent to acquiring s_m indirectly. Investor who invest with the skilled manager also needs to pay a fee of f. The fee is determined via Nash Bargaining, as we will see.

On the other hand, if an investor does not search for a skilled manager, his only option is to invest directly in the asset market [10].

[9] Though this assumption is mainly for tractability, we provide one rational.

For managers, skill can be acquired because the public information is available to all to process. However, uninformed institutions may not be

This is due to the innate heterogeneity among institutions. Whereas it is more likely that all money managers can exert effort and dissecting public information in the same depth, (or compete for the same pool of talented investment professionals), institutions are more different – the very source of information for some institutions, which rely on its place in the network, could be drastically different from that of the other institution. For example, the endowment of some small university simply cannot obtain the information available to Yale endowment, regardless of the cost of acquiring it.

[10] Alternatively, an investor may choose to invest with an arbitrary manager without searching, thus avoiding the search cost c. Doing so results in a positive probability where the manager happens to be skilled. This additional choice is discussed in Pedersen and Garleanu (2015). TODO: consider this extension in this model

To prevent a fully revealing equilibrium, we assume the supply of asset is noisy:

$$\tilde{q} \sim N\left(1, \gamma_q^{-1}\right) \tag{2.6}$$

As the mean of $\tilde{q}$ suggests, we normalize the supply of asset and suppose there is one share of asset on average.

Throughout the paper, we will use tilde to denote random variables. symbols without tilde (i.e. v, s_i) are used to denote specific realizations of the random variables. This distinction helps with clarification when we consider investor's problems in the next section.

Throughout the paper, we only consider equilibrium featuring linear price specified as follows:

$$p = a_0 + a_i s_i + a_m s_m + a_q q \tag{2.7}$$

Events unfold in four stages. At $t = 1$, managers decide whether to acquire s_m, and investors decide whether to search for skilled managers, both taking place simultaneously. At $t = 2$, investors who has found a skilled manager negotiate fees with their money manager. At $t = 3$, signals arrive; Investors (either directly or indirectly through money managers) trade for the asset based on their available information. At $t = 4$, asset value realizes and all agents consume.

2.2.2 Investor's Problem

Each institutional investor faces two decisions. First, each investor needs to decide whether to search and pay for skilled management; then, based on the investor's knowledge, which could include s_m if she chooses to invest with a manager, the investor formulates the demand for the asset. We analyze investor's problem backwards and first consider the asset demand.

For any given investor, let $\mathcal{F}$ be her the information set at $t = 3$. Recall from the time line, trading takes place after signal has realized, so inside $\mathcal{F}$ are the **realizations** of the signals, s_i and s_m, as well as the realization of price, p. Investor's CARA utility, coupled with normal random variables, in addition to the zero interest rate and no borrowing constraint, yields the following demand function:

$$D = \frac{E[\tilde{v}|\mathcal{F}] - p}{\alpha \operatorname{var}[\tilde{v}|\mathcal{F}]} \tag{2.8}$$

which is increasing in the conditional expectation, decreasing in risk aversion and conditional uncertainty, and independent of initial wealth. The derivation of (2.8) is deferred to the Appendix. Investors' trading decision at $t = 3$ is then given by the demand function (2.8). Note that implicitly we have assumed zero interest rate and no borrowing constraint. The interior solution for demand arises not because of budget constraint, but because of the trade-off between perceived risk and return.

We now consider investors' decision at $t = 1$ of whether to search and invest with a skilled manager. With demand function known, we first calculate investor's value function for each realization of the signals. We then integrate over all realizations and calculate the investor's **ex-ante** utility based on what signals are known. Explicitly, the information sets of interest are $\{\{\tilde{p}\}, \{\tilde{s}_i, \tilde{p}\}, \{\tilde{s}_m, \tilde{p}\}, \{\tilde{s}_i, \tilde{s}_m\}\}$. There are 4 (2 by 2) possibilities since investors can either be informed or uninformed and choose either invest with managers or not. Note that unless the investor is herself informed and search for skilled manager, price $\tilde{p}$ will enter her asset demand. Also note that each of the above information set contains price and signals as random variables (as indicated by $\tilde{}$), not their realizations. This is because searching takes place before signals are realized, so investors need to formulate ex-ante utility.

For each $\tilde{\mathcal{F}} \in \{\{\tilde{p}\}, \{\tilde{s}_i, \tilde{p}\}, \{\tilde{s}_m, \tilde{p}\}, \{\tilde{s}_i, \tilde{s}_m\}\}$, the value function at $t = 1$ is a function of $\tilde{\mathcal{F}}$ and is given by :

$$U(\tilde{\mathcal{F}}) = \frac{1}{2\alpha} \log\left(\frac{var(\tilde{v} - \tilde{p})}{var(\tilde{v}|\tilde{\mathcal{F}})}\right) + \text{constant} \tag{2.9}$$

The only dependency of ex-ante utility on $\tilde{\mathcal{F}}$ is through $var(\tilde{v}|\tilde{\mathcal{F}})$. Hence when comparing value functions corresponding to two information sets $\tilde{\mathcal{F}}_1$ and $\tilde{\mathcal{F}}_2$, the difference has the following convenient expression

$$U(\tilde{\mathcal{F}}_1) - U(\tilde{\mathcal{F}}_2) = \frac{1}{2\alpha} \log\left(\frac{var(\tilde{v}|\tilde{\mathcal{F}}_2)}{var(\tilde{v}|\tilde{\mathcal{F}}_1)}\right) \tag{2.10}$$

When we plug $\tilde{\mathcal{F}}_1 = \{\tilde{s}_i, \tilde{p}\}$, $\tilde{\mathcal{F}}_2 = \{\tilde{s}_i, \tilde{s}_m\}$ into the above expression, the outcome represents the gain from skilled management to an informed investor. Whether or not an informed investor would want to search depends on its magnitude relative to the search cost and the fee rate. That is, an informed investor's decision to search for skilled manager at $t = 1$ depends on whether the following inequality holds:

$$\frac{1}{2\alpha} \log \left(\frac{var(\tilde{v}|\tilde{s}_i, \tilde{p})}{var(\tilde{v}|\tilde{s}_i, \tilde{s}_m)} \right) > f_i + c \tag{2.11}$$

The left-hand-side is the certainty-equivalent from knowing $\tilde{s}_m$, whereas the right-hand-side represents the sum of search cost and the fee. Thus, an informed investor search if and only if (2.11) holds.

Similarly, uninformed investors search iff

$$\frac{1}{2\alpha} \log \left(\frac{var(\tilde{v}|\tilde{p})}{var(\tilde{v}|\tilde{s}_m, \tilde{p})} \right) > f_u + c \tag{2.12}$$

Note the distinction between fee paid by informed investor f_i and uninformed investor f_u. We will examine the process in which fees are determined and how they may be different for different investors.

2.2.3 Management Fee

The management fee paid by the investors to their money managers is determined from Nash Bargaining. Fee negotiation takes place after the managers have already acquired the information and investors have already conducted the due diligence. Therefore both manager's cost of skill, κ, investors cost of search c are sunk and do not affect the Bargaining outcome. For simplicity, we assume equal Bargaining power.[11] This sets fee as:

$$f = \arg\max_f \left(\text{investor's gain from } \tilde{s}_m - f\right) \cdot (f - 0) = \frac{1}{2} \cdot \text{investor's gain from } \tilde{s}_m$$

The objective of the maximization is the product of investor's net gain (gain less fee), and manager's gain (fee collected) from delegation. Conceptually, given the outside option of the investor, which

[11] Whereas managers are skilled by default given the matching process, we also assume whether the investor is informed or not is known to the manager. This simplifies the bargaining analysis. As we will see, informed and uninformed investors could pay different fee. If investor's type is not known to the managers, there may be incentive for investors to mask their true type if possible. Managers would then need to incorporate the potential untruthful reporting into their Bargaining decision.

is to invest directly in the asset market without the knowledge of $\tilde{s}_m$, and the outside option of the manager, which is to collect zero fee from the said investor, the total gain from combining their capital and expertise is investor's gain in certainty equivalent from knowing $\tilde{s}_m$. And from the assumption of equal bargaining power and manager's risk-neutral preference, the resulting fee distributes the total gain equally between the investor and the manager, which is precisely half the gain from $\tilde{s}_m$.

The bargaining solution also sheds light on why informed investors and uninformed investors might pay different fees. This is because they value skilled management differently – value of knowing $\tilde{s}_m$ depends on whether $\tilde{s}_i$ is also known. From expressions on the gain from s_m, the fee paid by informed investor is given by:

$$f_i = \frac{1}{4\alpha} \log \left(\frac{var(\tilde{v}|\tilde{s}_i, \tilde{p})}{var(\tilde{v}|\tilde{s}_i, \tilde{s}_m)} \right) \tag{2.13}$$

Similarly for uninformed investor, the fee is given by:

$$f_u = \frac{1}{4\alpha} \log \left(\frac{var(\tilde{v}|\tilde{p})}{var(\tilde{v}|\tilde{p}, \tilde{s}_m)} \right) \tag{2.14}$$

2.2.4 Manager Equalizing Condition

While skilled managers make a portfolio decision, the portfolio decision is mechanical from the assumption of fiduciary duty and is covered in the analysis of investor's asset demand. The remaining choice for any given manager is whether to pay the skill cost κ and acquire $\tilde{s}_m$ at $t = 1$. The marginal manager will become skilled if and only if

$$\underbrace{\frac{f_u I_u + f_i I_i}{M}}_{\text{Expected fee revenue}} > \underbrace{\kappa}_{\text{Cost of becoming skilled}}$$

When deciding whether to acquire skill or not, each manager takes as given other managers and investors decision at $t = 1$ regarding matching. In addition, all managers foresee the ex-ante value function based on the size I_i, I_m and can compute the fees that will arise at $t = 2$.

On the left hand side is the benefit from becoming skilled, which is the total fee revenue collected from investors who search for skilled management, normalized by the size of skilled management.

If the benefit is greater than the cost, more manager will become skilled. As we will see in the equilibrium, this in turn reduces the net gain of becoming skilled. Managers will keep adjusting until the indifference condition is satisfied:

$$\frac{f_u I_u + f_i I_i}{M} = \kappa \tag{2.15}$$

2.2.5 Equilibrium

An equilibrium is defined as the mapping from the set of parameters $\{\alpha, \gamma_0, \gamma_i, \gamma_m, \gamma_q, \lambda_i, \lambda_u, \bar{c}, \psi, \kappa\}$ to equilibrium objects $\{I_i, I_u, M, f_i, f_u, p(\cdot)\}$ such that

1. Each investor's decision on whether to search for skilled management maximizes the ex-ante utility specified in (2.9), taking $I_i, I_u, M, f_i, f_u, p(\cdot)$ as given; after both signals $\tilde{s}_m$ and $\tilde{s}_i$ realize, each investor's demand for asset follows the demand function specified in (2.8).

2. The marginal manager is indifferent on whether to acquire skill. That is, the indifference condition (2.15) holds.

3. For any investor that invest through a skilled manager, the management fee is determined via Nash Bargaining between the two parties and is as specified in (2.13) and (2.14).

4. Market for asset clears for all realizations of the signals:

$$q = D_{\{s_i, s_m\}} I_i + D_{\{s_i, p\}}(\lambda_i - I_i) + D_{\{s_m, p\}} I_u + D_{\{p\}}(\lambda_u - I_u) \tag{2.16}$$

For simplicity, we restrict our following analysis to equilibrium that are interior and symmetric. Here we refer interior equilibrium as those where there are positive mass of investors who search, and positive mass of managers who are skilled (The economy where no managers are skilled and no investors search for skilled managers naturally sustains an equilibrium. The marginal manager is content with not acquiring skill since no investor is searching; the marginal investor is content with not searching because the search cost is infinity with no skilled managers)

2.3 Equilibrium Featuring Complementarity

In our following analysis, we focus on equilibrium where investor's information is a complement to manager's skill. Here complementarity is defined similarly as in Admati (1985). That is

$$\underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|s_i,p)}{var(v|s_m,s_i)}\right)}_{\text{Gain from } s_m \text{ knowing } s_i} > \underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|p)}{var(v|s_m,p)}\right)}_{\text{Gain from } s_m \text{ not knowing } s_i} \tag{2.17}$$

The primary reason for focusing on this type of equilibrium is that it better represents the investments settings between institutions and their asset managers. As we will see, the results on investment performance in these equilibriums are able to replicate some empirical findings that are hard to reconcile otherwise.

2.3.1 Existence

While the model laid out in the previous section does not inherently assume complementarity, the following result shows that for certain general ranges of parameter, the complementarity do arise in equilibrium.

PROPOSITION 2.1. *Given parameter values $\Theta = \{\alpha, \gamma_0, \gamma_i, \gamma_m, \lambda_i, \lambda_u, \bar{c}, \psi\}$, there are open sets of Γ_q such that for each $\gamma_q \in \Gamma_q$, in the resulting equilibrium with Θ and γ_q, investor information and manager skill are complements as in (). In addition, for each γ_q, there exists K such that for any $\kappa \in K$, there exists symmetric equilibrium with Θ, γ_q, κ, where all informed investors invest through skilled managers, whereas all uninformed investors invest directly on their own.*

The proof is given in the Appendix, part of which borrows from Proposition 6.2 of Admati and Pfleiderer (1987). The idea is that when there is little noise in asset supply, signals are complements. This is because the value of a certain signal, from expression of the demand function and conditional expectations, is proportional to how much weight the agent put on it when forming their demand. When there is little noise in the asset supply, price become a near-perfect summary statistics of s_i and s_m. Therefore unless both s_i and s_m are observed, the agent will put significant weight on the information reflect in price. As a result, uninformed investor, who do not observe s_i, assign little value to s_m. On the other hand, informed investor values s_m greatly since once s_m is acquired, p will no longer be used in formulating demand. As a result, significant weight is put on s_m and thus their valuation of skilled management. This discrepancy is the main driver of the complementarity.

The second part of the Proposition shows that whenever complementarity holds, there are regions of manager's skill cost κ where the resulting equilibrium is symmetric and separating. The proof is constructive. We first compute the value of skilled management s_m, taking investors searching decision as given. In particular, we assume $I_i = \lambda_i$ and $I_0 = 0$. From complementarity, informed investor necessarily receive a higher value from s_m than uninformed investors. The search cost that can sustain this separating searching pattern then must lie between the values. From the cost function (2.5), the range in searching cost translates to an upper and lower bound for the size of skilled management M, which then can be translated to a range in terms of κ when fed into the manager indifference condition (2.15).

2.3.2 Outperformance relative to benchmark

In this section we analyze the investment performance of investors and managers in an equilibrium featuring complementarity. First we define benchmark performance as the ex-ante expected earning when only source of information is price. That is, all value function will be measured relative to $U(\{\tilde{p}\})$ where $U(\cdot)$ is given in (2.9). As such, uninformed investors, who also choose not to invest with managers in equilibrium, invest observing only p and earn benchmark by definition. On the other hand, informed investors, who search for skilled managers, do beat the benchmark. Particularly, the magnitude of gross outperformance (i.e. outperformance before fee) can be dissected in the following way:

$$\underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|p)}{var(v|s_m,s_i)}\right)}_{\text{Gross outperformance}} = \begin{cases} \underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|p)}{var(v|s_i,p)}\right)}_{\text{Value of } s_i} + \underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|p)}{var(v|s_m,p)}\right)}_{\text{Value of } s_m} \\ + \underbrace{\frac{1}{2\alpha}\log\left(\frac{var(v|s_i,p)var(v|s_m,p)}{var(v|p)var(v|s_i,s_m)}\right)}_{\text{Synergy}} \end{cases} \tag{2.18}$$

The first two terms on the right are values of s_i and s_m on their own, conditioning on the price. The third term is the synergy term and is positive due to complementarity.

The results above on outperformance relative to benchmark replicate the empirical findings. From Gerakos et al. (2016), institutional funds generate above benchmark returns. In our model, managers managing (institutional) investors capital always beat the market. This is because in equilibrium, only informed investors invest with skilled managers. Hence the outperformance is a combination of investors' private information, managers' skill, and the synergy.

On the other hand, the result where uninformed investors receiving benchmark is consistent with findings on performance of institutions overall (i.e. Lewellen (2011) from 13F). In particular, one puzzling fact is the lack of outperformance on the capital managed in-house by the institutions themselves, despite the implications from interactions between institutions and their asset managers that institution do possess advantage in investing. Supposedly if institutions have any advantage in investment, it should be reflected through outperformance. According to the model, this is because the institutions who invest directly are precisely the ones without private information. Alternatively, for institutions who is experienced and able to provide value-added to only some segments of the market, they will choose to delegate that part of the asset under management to external managers, who they deem would make better use of their information. The assets they end up investing themselves will earn benchmark return.

Put differently, had the investors of the institutional funds invested themselves, they would still beat the benchmark because of their private information. They choose to invest through the managers because of the additional synergy. Similarly, had institutions allocate their in-house asset to external managers, they would also earn an above benchmark return. However, they chose not to do so due to the high search cost.

2.4 Comparative Statics

In this section, we study how certain equilibrium objects vary with respect to changes in the underlying parameters. Throughout this section, we will emphasize the effect of change in information allocation among investors. Note that all comparative statics discussed here are local, in that all equilibrium must still lie in the regime where complementary holds.

2.4.1 Gross outperformance

We first examines the gross outperformance received by skilled managers with private information of the investors. For simplicity, we will refer it as the gross outperform and denote it as π_g. π_g is defined as the left-hand side of (2.18). The following proposition states how π_g changes with respect to the underlying parameters.

PROPOSITION 2.2. *Within regions of parameters where resulting equilibriums featuring complementarity, equilibrum gross outperformance π_g is increasing in investors' risk aversion α, and decreasing in mass informed investors λ_i, and precision in asset supply γ_q. π_g is increasing in γ_i and γ_m only when $\gamma_i + \gamma_m$ is sufficiently small. Specifically,*

$$\frac{\partial \pi_g}{\partial \gamma_i} = \frac{\partial \pi_g}{\partial \gamma_m} > 0 \quad \textit{iff} \quad \underbrace{\gamma_0}_{\textit{Precision of prior}} > \underbrace{\frac{(\gamma_i + \gamma_m)^2 \gamma_q \lambda_i^2}{\alpha^2}}_{\textit{Precision of price}} \tag{2.19}$$

The results on α, λ_i and γ_g are intuitive. Since gross outperformance originates from the unique access to s_i and s_m, it is less when price is more reflective of these underlying signals, which occurs if there is more informed trading, either because there are more informed investor λ_i, or that each informed investor is trading a greater amount due to lower risk aversion α. Price is also more reflective when there is less noise in the supply, that is, when the prior precision on supply γ_q is greater.

The interesting result is how π_g varies when the same mass of informed investors and skilled managers each hold better signals. On the one hand, having better signals leads informed investors and skilled managers to formulate better asset demands, which would exert an upward pressure to the outperformance. On the other hand, the equilibrium effects of better signal translate to less noise in price, which allow uninformed investors to better infer the underlying signals. This in turn pushes up benchmark performance and in turn reduces outperformance. The net effect is thus ambiguous and would depend on the relative magnitudes of the two mechanism. As the expression suggests, the condition in which the directly effect dominates the equilibrium effect is rather simple – we only needs to compare γ_0, which is the prior precision on asset payoff, to the precision of price as the noisy signal of the summary statistics.

Intuitively, prior precision γ_0 proxies the direct impact of a unit increase in signal precision (either γ_i or γ_m) has on informed investor's absolute performance; Similarly, the precision of price proxies the indirect impact the increase in signal precision has on the performance of the uninformed. Concretely, when $\gamma_m + \gamma_i$ is small, the condition in (2.19) will hold. This suggests that when informed investors and skilled managers do not have much of an edge in investments, an increase in the quality of either's signal would always generate a net improvement in terms of gross outperformance. On the contrary, suppose investors and managers collectively have precise signals, then $\gamma_m + \gamma_i$ is large and (2.19) would not hold. As such, agents have incentive to lower the quality of signal by (creditably) not fully utilize their signals.

2.4.2 Management Fee

Whereas the previous section focuses on the total benefit of edge in investment shared among skilled managers and informed managers, this section sheds light on how the benefit is split between the two parties. Specifically, we examine how the fee (managers' share of reward for their skill) varies when the underlying parameter varies. Note that since only the informed investors invest through money managers, by fee we refer to the fee paid by informed investors, which is given in (2.13).

We first state the proposition:

PROPOSITION 2.3. *Within regions of parameters where resulting equilibrium featuring complementarity, equilibrium fee f is increasing in investors' risk aversion α, and decreasing in mass informed investors λ_i, and precision in asset supply γ_q. f is decreasing in γ_i, and is increasing in γ_m only when γ_m is sufficiently small. Specifically,*

$$\frac{\partial f}{\partial \gamma_m} = \frac{\partial \pi_g}{\partial \gamma_m} > 0 \quad \textit{iff} \quad \underbrace{\gamma_0 + \gamma_i}_{\textit{Precision of prior}} > \underbrace{\frac{\gamma_m^2 \gamma_q \lambda_i^2}{\alpha^2}}_{\textit{Precision of price}} \tag{2.20}$$

To better understand the results in Proposition 3, we interpret it along side Proposition 2. For many underlying parameters, change in them result in the same qualitative change in f as in π_g. This is driven by the Bargaining outcome: fee represents manager's share from the bargaining outcome, which is proportional to the total gain π_g. As such, the same comparative statics arises for many parameters for π_g and f.

However, for γ_i and γ_m, things are different because they affect how the total gain is splitted between investors and managers. Intuitively, as γ_i (γ_m) increases, investors (managers) are contributing more to the team production of outperformance, resulting in a smaller (larger) portion of the outperformance going to the managers in the forms of fee. Mechanically, Proposition 2.3 is the c

Indeed, proposition states that f is decreasing in γ_i, this is in spite of the fact that greater γ_i can drive up the total outperformance π_g – though the total size of the pie becomes larger, managers end up receiving less fee. This is because investor's outside option of investing directly also increases as the result of the improvements in the quality of investors' signal.

Similarly, there are regions for γ_i, γ_m such that condition in (2.20) holds whereas (2.19) does not. That is, an increase in γ_m leads to a smaller total outperformance, yet manager receives a higher level of fee.

2.4.3 Mass of skilled managers

What portion of the money managers are actually skilled? In particular, how would the money managers as a whole respond in terms of skill acquisition with regard to changes in investors' information structure. In this section, we focus on two aspects on the investors side: how many of investors are informed— size of λ_i, and the quality of those investor's information — level of γ_i. As before, we focus on equilibrium where all informed investors invest with skilled managers.

From manager's indifference (2.15), since manager's skilled cost κ is fixed, M is proportional to the total fee revenue $\lambda_i f$. Therefore it is equivalent to consider comparative statics of $\lambda_i f$.

First consider a change in γ_i, the quality of each informed investor's information. Here the number of informed investors λ_i is fixed. From Proposition 3, each investor will pay a lower fee rate due to their improved outside option. As such, total fee revenue for all skilled managers decrease. To restore manager indifference, there must then be a smaller number of skilled managers. In this sense, the increase in investor ability crowds out manager skill.

On the other hand, when there are more informed investors, holding the quality of their information fixed, how the number of skilled managers change is ambiguous. There are two opposing forces at work. First, more informed investors mechanically increase fee revenue, since informed investors are the ones paying the fee in equilibrium; secondly, as proposition 3 stated, as there are more informed investors that invest with managers, manager signal s_m is reflected more in the price, driving down both π_g and f. Therefore the net effect of increase in λ_i on total fee revenue depends on which of the two forces are dominant.

The above analysis can be summarized in the following Proposition.

PROPOSITION 2.4. *Within regions of parameters where resulting equilibrium featuring complementarity, equilibrium mass of skilled manager M is decreasing in the quality of investors' private information γ_i. M can be increasing or decreasing with respect to λ_i*

2.5 Price informativeness and Information concentration

The previous section shed light on how change in investor's information endowment can affect managers skill acquisition in equilibrium. In this section, we take a normative perspective, and examine how a policy maker, whose goal is to maximize the informativeness of price, may want to change the initial information structure of the investors. Specifically, would the policy maker prefer more and moderately informed investors, or less but significantly informed investors?

First we define price informativeness as the inverse of the conditional variance of asset payoff, conditioning only on the price. We denote price informativeness as η. Explicitly

$$\eta = \frac{1}{var(v|p)} = \gamma_0 + \gamma_i + \gamma_m - \frac{(\gamma_i + \gamma_m)A}{\gamma_i + \gamma_m + A} \tag{2.21}$$

where $A = \frac{\alpha^2}{\lambda_i^2 \gamma_q}$. This definition of price informativeness is natural and is similar to the price efficiency in Grossman and Stiglitz (1980). As η increases, there is less uncertainty regarding asset payoff when conditioning only on price.

It is straightforward to check that η is increasing in both γ_i and λ_i. That is, price will become more informative if more investors are endowed with information, or if quality of information endowed to each investor is better. Both are expected. However, how do the quantity and quality trade off is less obvious.

In order to quantify and compare the effects of change in γ_i and λ_i, we introduce of the notion of the **amount of investor information** for an economy with given set of parameters. Specifically, consider the following two economies 1 and 2. Let $(\lambda_{i,1}, \gamma_{i,1})$, $(\lambda_{i,2}, \gamma_{i,2})$ be the parameters related to investor information structures for economy 1 and 2, respectively. Furthermore, suppose all other parameters are identical. Economy 1 and 2 are said to have the **amount of investor information** if $\eta_1 = \eta_2$ absent any money managers. That is, if

$$\gamma_0 + \gamma_{i,1} - \frac{\gamma_{i,1} A_1}{\gamma_{i,1} + A_1} = \gamma_0 + \gamma_{i,2} - \frac{\gamma_{i,2} A_2}{\gamma_{i,2} + A_2} \tag{2.22}$$

Now with the amount of investor information fixed as the unit of analysis, we can discuss the optimal concentration of quality and quantity of investor information. Result can be summarized in the following proposition

PROPOSITION 2.5. *Holding fixed the amount of investor information, price informativeness is greater when investor information is diffused.*

The proof consider marginal rates of substitution between γ_i and λ_i, first holding fixed the amount of investor information, then holding fixed the price informativeness. We can show that one always dominates the other. Intuitively, consider an increase in λ_i and decrease in γ_i that keep the amount of investor information unchanged. As we introduce skilled managers into the analysis, and focus on equilibrium where informed investors all search for skilled managers, the effect of quantity λ_i has on price informativeness is amplified since each informed investors will end up trade on manager's signal s_m as well. As such, the resulting price informativeness increases due to the spillover effects of increase in λ_i.

From a policy maker's perspective, if the goal is to maximize price informativeness, he should prefer quantity of investor information over quality of investor information. More concretely, suppose there are a few small circles of elite institutions, whose board members are closely connected and often shares private information, the policy maker should have incentive to, to the extent possible, break apart the coalitions. As a result, each institution's private information will be less valuable, but there will be more institutions with the private information.

2.6 Conclusion

In this paper, we consider a model where asset managers collaborate in generating returns. In doing so, we provide an explanation of the lack of in-house outperformance by the skilled institutional investors. Specifically, institutions act on their information by passing it onto external asset managers. Hence evaluating institutions' investment expertise based solely on their in-house performance likely leads to understatements.

2.A Proofs

2.A.1 Preliminaries

FACT 1 (Conditional Moments of Multivariate Normal Distribution)**.** *Consider the multivariate normal vector* $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$*. Consider partitioning* $\boldsymbol{\mu}$ *and* $\boldsymbol{Y}$ *into* $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$ $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}$ *with a similar partition of* σ *into* $\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$*. Then,* $(\boldsymbol{y}_1 | \boldsymbol{y}_2 = \boldsymbol{a})$*, the conditional distribution of the first partition given the second, is* $\mathcal{N}(\overline{\boldsymbol{\mu}}, \overline{\Sigma})$*, with mean* $\overline{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \Sigma_{12}{\Sigma_{22}}^{-1}(\boldsymbol{a} - \boldsymbol{\mu}_2)$ *and covariance matrix* $\overline{\Sigma} = \Sigma_{11} - \Sigma_{12}{\Sigma_{22}}^{-1}\Sigma_{21}$*.*

FACT 2. *If* $y \sim \mathcal{N}(\mu_y, \sigma_y)$*, then* $E[e^{-ay}] = \exp\left(-a\mu_y + \frac{a^2\sigma_y^2}{2}\right)$

Proof.

$$
\begin{aligned}
& E[e^{-ay}] \\
= \; & \int e^{-at} \tfrac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(t-\mu_y)^2}{2\sigma_y^2}} \\
= \; & \int \tfrac{1}{\sqrt{2\pi}\sigma_y} \exp(-\tfrac{(t-\mu_y)^2}{2\sigma_y^2} - at) \\
= \; & \int \tfrac{1}{\sqrt{2\pi}\sigma_y} \exp(-\tfrac{t^2 - 2\mu_y t + \mu_y^2}{2\sigma_y^2} - at) \\
= \; & \int \tfrac{1}{\sqrt{2\pi}\sigma_y} \exp(-\tfrac{t^2 - 2(\mu_y - a\sigma_y^2)t + \mu_y^2}{2\sigma_y^2}) \\
= \; & \int \tfrac{1}{\sqrt{2\pi}\sigma_y} \exp(-\tfrac{t^2 - 2(\mu_y - a\sigma_y^2)t + (\mu_y - a\sigma_y^2)^2 - a^2\sigma_y^4 + 2\mu_y a\sigma_y^2}{2\sigma_y^2}) \\
= \; & \int \tfrac{1}{\sqrt{2\pi}\sigma_y} \exp(-\tfrac{\left[t - (\mu_y + a\sigma_y^2)\right]^2}{2\sigma_y^2} + \tfrac{a^2\sigma_y^2}{2} - a\mu_y) \\
= \; & \exp(-a\mu_y + \tfrac{a^2\sigma_y^2}{2})
\end{aligned}
$$

$\square$

FACT 3. *If* $z \sim \mathcal{N}(\mu_z, \sigma_z)$*, then* $E[e^{-\frac{1}{2}z^2}] = \frac{1}{\sqrt{1+\sigma_z^2}} \cdot \exp\left(-\frac{\mu_z^2}{2(1+\sigma_z^2)}\right)$

Proof.

$$
\begin{aligned}
E[e^{-\frac{1}{2}z^2}] &= \int \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(t-\mu_z)^2}{2\sigma_z^2}\right) \exp\left(-\frac{1}{2}t^2\right) dt \\
&= \int \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{t^2 - 2\mu_z t + \mu_z^2 + t^2\sigma_z^2}{2\sigma_z^2}\right) dt \\
&= \int \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{\left(1+\sigma_z^2\right)t^2 - 2\mu_z t + \mu_z^2}{2\sigma_z^2}\right) dt \\
&= \int \frac{1}{\sqrt{1+\sigma_z^2}} \frac{1}{\sqrt{2\pi}\frac{\sigma_z}{\sqrt{1+\sigma_z^2}}} \exp\left(-\frac{t^2 - 2\frac{\mu_z}{(1+\sigma_z^2)}t + \frac{\mu_z^2}{(1+\sigma_z^2)}}{2\left(\frac{\sigma_z}{\sqrt{1+\sigma_z^2}}\right)^2}\right) dt \\
&= \int \frac{1}{\sqrt{1+\sigma_z^2}} \frac{1}{\sqrt{2\pi}\frac{\sigma_z}{\sqrt{1+\sigma_z^2}}} \exp\left(-\frac{t^2 - 2\frac{\mu_z}{(1+\sigma_z^2)}t + \frac{\mu_z^2}{(1+\sigma_z^2)^2} - \frac{\mu_z^2}{(1+\sigma_z^2)^2} + \frac{\mu_z^2}{(1+\sigma_z^2)}}{2\left(\frac{\sigma_z}{\sqrt{1+\sigma_z^2}}\right)^2}\right) dt \\
&= \frac{1}{\sqrt{1+\sigma_z^2}} \cdot \exp\left(-\frac{-\frac{\mu_z^2}{(1+\sigma_z^2)} + \mu_z^2}{2\sigma_z^2}\right) \\
&= \frac{1}{\sqrt{1+\sigma_z^2}} \cdot \exp\left(-\frac{\mu_z^2}{2\left(1+\sigma_z^2\right)}\right)
\end{aligned}
$$

□

LEMMA 1. *The maximizer of the CARA objective in (??) is given by*

$$x_D = \frac{E(v|D) - p}{\gamma var(v|D)} \tag{2.23}$$

if D contains only realizations of normal variables

Proof. Since D only contain information about other normal variables. Fact 1 implies $v|D$ is also normal. Hence we can apply Fact 2 to the objective and obtain:

$$
\begin{aligned}
& -\tfrac{1}{\gamma}\log E(e^{-\gamma x(v-p)}|D) \\
= & -\tfrac{1}{\gamma}\log E(e^{-\gamma xv} \cdot e^{\gamma xp}|D) \\
= & -xp - \tfrac{1}{\gamma}\log E(e^{-\gamma xv}|D) \\
= & -xp - \tfrac{1}{\gamma}\left[\tfrac{\gamma^2 x^2 var(v|D)}{2} - \gamma x E(v|D)\right] \\
= & \; x\left[E(v|D) - p\right] - \tfrac{\gamma x^2 var(v|D)}{2}
\end{aligned}
$$

Demand x_D can then be obtained from the first order condition. □

LEMMA 2. *The ex-ante value of CARA utility in* **??** *with access of information D is given by*

$$u_D = \frac{1}{\gamma}\log\left(\frac{\sigma_{v-p}}{\sigma_{v|p}}\right) + \frac{1}{2\gamma}\frac{(\bar{v}-\theta_0)^2}{\sigma_{v-p}^2}$$

Proof. Using demand x_D in 2.23 we have

$$\begin{aligned}
E\left[e^{-\gamma x_D(v-p)}|D\right] &= E\left[\exp\left(-\gamma x_D(v-p)\right)|D\right] \\
&= E\left[\exp\left(-\frac{E\left(v-p|D\right)}{var(v|D)}(v-p)\right)|D\right] \\
&= \exp\left[-\frac{E\left(v-p|D\right)}{var(v|D)}E(v-p|D) + \frac{1}{2}\frac{\left[E\left(v-p|D\right)\right]^2}{var(v|D)}\right] \\
&= \exp\left[-\frac{1}{2}\frac{\left[E\left(v-p|D\right)\right]^2}{var(v|D)}\right]
\end{aligned}$$

where the equality between second and third line uses Fact 2. Furthermore, since $E\left(v|D\right)$ and p are normal, and $var(v|D)$ only depend on parameters, $\frac{\left[E(v-p|D)\right]^2}{var(v|D)}$ has chi-squared distribution. Specifically $\frac{E(v-p|D)}{\sigma_{v|D}} \sim \mathcal{N}\left(\overline{\mu}, \overline{\sigma}_z^2\right)$ where

$$\overline{\mu} = \frac{E\left[E\left(v-p|D\right)\right]}{\sigma_{v|D}} = \frac{E\left[v-p\right]}{\sigma_{v|D}} = \frac{\overline{v}-\theta_0}{\sigma_{v|D}}$$

$$\overline{\sigma}_z^2 = \frac{var\left[E\left(v-p|D\right)\right]}{\sigma_{v|D}^2} = \frac{var\left(v-p\right) - E\left[var\left(v-p|D\right)\right]}{\sigma_{v|D}^2}$$

Applying Fact 3, we have

$$\begin{aligned}
\exp\left[-\frac{1}{2}\frac{\left[E\left(v-p|D\right)\right]^2}{var(v|D)}\right] &= \frac{1}{\sqrt{1+\overline{\sigma}_z^2}}\cdot\exp\left(-\frac{\overline{\mu}_z^2}{2\left(1+\overline{\sigma}_z^2\right)}\right) \\
&= \frac{\sigma_{v|D}}{\sigma_{v-p}}\cdot\exp\left(-\frac{(\overline{v}-\theta_0)^2}{2\sigma_{v-p}^2}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
u_D &= -\frac{1}{\gamma}\log E_D\left[E(e^{-\gamma x(v-p)}|D)\right] \\
&= -\frac{1}{\gamma}\log E_D\left[\exp\left[-\frac{1}{2}\frac{\left[E\left(v|D\right)-p\right]^2}{var(v|D)}\right]\right] \\
&= -\frac{1}{\gamma}\log E_D\left[\frac{\sigma_{v|D}}{\sigma_{v-p}}\cdot\exp\left(-\frac{(\overline{v}-\theta_0)^2}{2\sigma_{v-p}^2}\right)\right] \\
&= \frac{1}{\gamma}\log\left(\frac{\sigma_{v-p}}{\sigma_{v|p}}\right) + \frac{1}{2\gamma}\frac{(\bar{v}-\theta_0)^2}{\sigma_{v-p}^2}
\end{aligned}$$

□

2.A.2 Proofs of Propositions

2.A.2.1 Proposition 1

Proof. We begin by positing a symmetric equilibrium where only informed investors search for skilled managers at $t = 1$. Then we will show that at $t = 3$, the linear price exists that would clear the market. Lastly, we use the derived equilibrium price to show that there exists ranges of γ_q and k such that the presumed searching pattern are indeed optimal at $t = 1$.

1. Existence of equilibrium price Suppose $I_i = \lambda_i$, $I_u = 0$. Then whomever have access to $\tilde{s}_i$ also has access to $\tilde{s}_m$ and vice versa. The resulting information structure is equivalent to the setting with only one signal on the asset payoff, namely the summary statistics of $\tilde{s}_i$ and $\tilde{s}_m$:

$$\tilde{s} = \frac{\gamma_i \tilde{s}_i + \gamma_m \tilde{s}_m}{\gamma_i + \gamma_m} \tag{2.24}$$

with the property

$$\tilde{s} = \tilde{v} + \tilde{\epsilon} \quad \text{where} \quad \tilde{\epsilon} \sim N\left(0, \gamma_s^{-1}\right) \tag{2.25}$$

where $\gamma_s = \gamma_i + \gamma_m$ By assumption, there are mass λ_i of investors with access to signal $\tilde{s}$ and λ_u without the signal. As such, we can rewrite linear price (2.7), which is in terms of both s_i and s_m, to the following price function, which only depends on s:

$$p = b_0 + b_s(s + b_q q) \tag{2.26}$$

By GS (1980), there exists unique solution (b_0, b_s, b_q) such that the resulting linear price (2.7) always lead to market clearing as defined in (2.16). We briefly outline the arguments here.

First derive the conditional expectations and variances of asset payoff for investors with the signal s. This is straightforward since p is not used:

$$E(v|s) = -\frac{\gamma_s s}{\gamma_s + \gamma_0} \tag{2.27}$$

$$var(v|s) = (\gamma_s + \gamma_0)^{-1} \tag{2.28}$$

On the other hand, the uninformed investors need to infer s from price. It is helpful to consider the linear transformation of p:

$$p' = \frac{p - b_0}{b_s} = s + b_q q \tag{2.29}$$

which can be interpreted as a signal of v with precision from which it is easy to compute that

$$\gamma_p = \left[(\gamma_i + \gamma_m)^{-1} + b_q^2 \gamma_q^{-1}\right]^{-1} \tag{2.30}$$

As such, we have:

$$E(v|p) = -\frac{\gamma_p (p - b_0)/b_s}{\gamma_p + \gamma_0} \tag{2.31}$$

$$var(v|p) = (\gamma_p + \gamma_0)^{-1} \tag{2.32}$$

We can then plug the conditional expectations and variances into the right-hand side of the market clearing condition (2.16). As such, both sides will be linear in s and q. In order to market clearing to always hold, we match coefficients for s, q, and the constant, and solve the following system of equations in term of b_0, b_s and b_q:

$$\begin{cases} 0 = \lambda_i \dfrac{\frac{\gamma_s}{\gamma_s + \gamma_0} - b_s}{\alpha var(v|s)} + \lambda_u \dfrac{\frac{\gamma_p}{\gamma_p + \gamma_0} - b_s}{\alpha var(v|p)} \\ 1 = \lambda_i \dfrac{-b_s b_q}{\alpha var(v|s)} + \lambda_u \dfrac{\frac{\gamma_p}{\gamma_p + \gamma_0} b_q - b_s b_q}{\alpha var(v|p)} \\ 0 = \lambda_i \dfrac{-b_0}{\alpha var(v|s)} + \lambda_u \dfrac{-b_0}{\alpha var(v|p)} \end{cases}$$

Multiply the first equation with b_q then subtract it from the second equation. Rearranging the terms yields:

$$b_q = -\frac{\alpha}{\lambda_i \gamma_s} \tag{2.33}$$

Given the above expression for b_q, both $var(v|p)$ and γ_p can be computed directly. Rearranging the second yields solution for b_s, and rearranging the third equation yields solution for b_0.

To map the solutions to the original price (2.7), apply the following transformation

$$a_0 = b_0 \tag{2.34}$$

$$a_i = b_s \frac{\gamma_i}{\gamma_i + \gamma_m} \tag{2.35}$$

$$a_m = b_s \frac{\gamma_m}{\gamma_i + \gamma_m} \tag{2.36}$$

$$a_q = b_s b_q \tag{2.37}$$

Hence the equilibrium price exists by construction.

2. Demonstrate complementarity Our goal is to show that given any parameters other than γ_q, there exists $\bar{\gamma}_q$ such that (2.17) holds as long as $\gamma_q \in (0, \bar{\gamma}_q)$.

From equilibrium price derived above, we can derive conditional variances used in the complementarity condition (2.17) as functions of parameters. Below we express the precisions (reciprocal of conditional variances) for convenience and more insights:

$$var(v|s_i, s_m)^{-1} = \gamma_0 + \gamma_i + \gamma_m \tag{2.38}$$

$$var(v|s_i, p)^{-1} = \gamma_0 + \gamma_i + \gamma_m - \frac{\gamma_m A}{\gamma_m + A} \tag{2.39}$$

$$var(v|s_m, p)^{-1} = \gamma_0 + \gamma_i + \gamma_m - \frac{\gamma_i A}{\gamma_i + A} \tag{2.40}$$

$$var(v|p)^{-1} = \gamma_0 + \gamma_i + \gamma_m - \frac{(\gamma_i + \gamma_m) A}{\gamma_i + \gamma_m + A} \tag{2.41}$$

where

$$A = \frac{\alpha^2}{\lambda_i^2 \gamma_q} \tag{2.42}$$

can be interpreted as the overall noise in equilibrium price.

The above expression of precisions provide direct interpretation. For instance, $var(v|s_i, p)^{-1}$ can be decomposed as the precision when all three signals are known, less the noise in infering s_m from price. The term representing the magnitude of the reduction in precision due to noise, $\frac{xA}{x+A}$ is increasing in both x and A. That is, the lost in precision is greater if either the unknown signal carries more information itself, or if the price is more noisy in general.

Going back to the complementarity condition (2.17), it is equivalent to show

$$var(v|p)^{-1}var(v|s_i, s_m)^{-1} - var(v|s_i, p)^{-1}var(v|p, s_m)^{-1} > 0$$

plugging in the expression for the precisions to the left hand side we have

$$\begin{aligned}
&var(v|p)^{-1}var(v|s_i, s_m)^{-1} - var(v|s_i, p)^{-1}var(v|p, s_m)^{-1} \\
=&\left(\gamma_0 + \gamma_i + \gamma_m - \frac{(\gamma_i + \gamma_m)A}{\gamma_i + \gamma_m + A}\right)(\gamma_0 + \gamma_i + \gamma_m) - \left(\gamma_0 + \gamma_i + \gamma_m - \frac{\gamma_i A}{\gamma_i + A}\right)\left(\gamma_0 + \gamma_i + \gamma_m - \frac{\gamma_m A}{\gamma_m + A}\right) \\
=&A\left[(\gamma_0 + \gamma_i + \gamma_m)\left(\frac{\gamma_m}{\gamma_m + A} + \frac{\gamma_i}{\gamma_i + A} - \frac{(\gamma_i + \gamma_m)}{\gamma_i + \gamma_m + A}\right) - \frac{\gamma_i \gamma_m A}{(\gamma_i + A)(\gamma_m + A)}\right]
\end{aligned}$$

which evaluates to 0 when $A = 0$. Furthermore, its derivative with respect to A at $A = 0$ is given by $\gamma_0 + \gamma_i + \gamma_m > 0$. Since the entire expression is continuous in A, for any $\gamma_0, \gamma_i, \gamma_m$, there exists $\bar{A}$ small enough such that for all $A < \bar{A}$, complementarity condition (2.17) holds.

3. Verifying searching pattern Now suppose with $\gamma_0, \gamma_i, \gamma_m$ given, and $A < \bar{A}$ so that complementarity holds. We show that there exists region of k such that the presupposed searching pattern — only informed investors invest through skilled managers — is indeed incentive compatible. Using expressions on precisions (2.38) - (2.41), we can compute the gain from s_m both for informed and uninformed investors. By complementarity, the gain for informed investors is greater than the gain for uninformed investor. (Note that since fee will be precisely half of the gain. this ordering is sustained even accounting for the fee). As such, we can denote $\bar{c}$ as the gain from s_m for the informed, and $\underline{c}$ for the uninformed. From cost function (2.5), we can find $\underline{M}$ and $\bar{M}$ that correspond to $\underline{c}$ and $\bar{c}$. Lastly, from manager's indifference condition (2.15), since total revenue is fixed, we can find $\underline{k}$ and $\bar{k}$ such that

$$k \in [\underline{k}, \bar{k}] \Rightarrow M \in [\underline{M}, \bar{M}] \Rightarrow c(M) \in [\underline{c}, \bar{c}]$$

which implies the assumed searching pattern is indeed incentive compatible.

Detail derivation on $var(v|s_i, p)$, ($var(v|s_m, p)$ is almost identical)

$$\begin{aligned}
var(v|s_i,p)^{-1} &= \gamma_0 + \gamma_i + \left(\gamma_m^{-1} + \left(\frac{a_q}{a_m}\right)^2 \gamma_q^{-1}\right)^{-1} \\
&= \gamma_0 + \gamma_i + \left(\gamma_m^{-1} + \left(b_q \frac{(\gamma_i+\gamma_m)}{\gamma_m}\right)^2 \gamma_q^{-1}\right)^{-1} \\
&= \gamma_0 + \gamma_i + \left(\gamma_m^{-1} + \left(\frac{\alpha}{\lambda_i \gamma_m}\right)^2 \gamma_q^{-1}\right)^{-1} \\
&= \gamma_0 + \gamma_i + \left(\gamma_m^{-1} + A\gamma_m^{-2}\right)^{-1} \qquad \text{where} \qquad A = \frac{\alpha^2}{\lambda_i^2 \gamma_q} \\
&= \gamma_0 + \gamma_i + \frac{\gamma_m^2}{\gamma_m + A} \\
&= \gamma_0 + \gamma_i + \gamma_m - \frac{\gamma_m A}{\gamma_m + A}
\end{aligned}$$

□

2.A.2.2 Proposition 2

Proof. We can write π_g explicitly in terms of parameters by plugging (2.38) and (2.41) into left-hand side of (2.18):

$$\pi_g = \frac{1}{2\alpha}\left[log\left(\gamma_i + \gamma_m + \gamma_0\right) - log\left(\gamma_i + \gamma_m + \gamma_0 - \frac{(\gamma_i+\gamma_m)A}{\gamma_i+\gamma_m+A}\right)\right] \tag{2.43}$$

First, we have

$$\frac{\partial \pi_g}{\partial A} = \frac{1}{2\alpha var(v|p)} \cdot \frac{(\gamma_i+\gamma_m)^2}{(\gamma_i+\gamma_m+A)^2} > 0 \tag{2.44}$$

It then follows from (2.42) and the chain rule that $\frac{\partial \pi_g}{\partial \alpha} > 0$, $\frac{\partial \pi_g}{\partial \lambda_i}, \frac{\partial \pi_g}{\partial \gamma_q} < 0$

We also have $\frac{\partial \pi_g}{\partial \gamma_0} < 0$ from direct computation:

$$\frac{\partial \pi_g}{\partial \gamma_0} = \frac{1}{2\alpha}\left[\frac{1}{\gamma_i+\gamma_m+\gamma_0} - \frac{1}{\gamma_i+\gamma_m+\gamma_0 - \frac{(\gamma_i+\gamma_m)A}{\gamma_i+\gamma_m+A}}\right] < 0.$$

We now investigate $\frac{\partial \pi_g}{\partial \gamma_i}$ and $\frac{\partial \pi_g}{\partial \gamma_m}$. For $\frac{\partial \pi_g}{\partial \gamma_i}$, we have

$$\begin{aligned}
\frac{\partial \pi_g}{\partial \gamma_i} &= \frac{1}{2\alpha}\left[\frac{1}{\gamma_i+\gamma_m+\gamma_0} - \frac{1}{\gamma_i+\gamma_m+\gamma_0-\frac{(\gamma_i+\gamma_m)A}{\gamma_i+\gamma_m+A}}\cdot\left(1-\frac{A^2}{(\gamma_i+\gamma_m+A)^2}\right)\right] \\
&\propto \left(\gamma_i+\gamma_m+\gamma_0-\frac{(\gamma_i+\gamma_m)A}{\gamma_i+\gamma_m+A}\right) - (\gamma_i+\gamma_m+\gamma_0)\left(1-\frac{A^2}{(\gamma_i+\gamma_m+A)^2}\right) \\
&= (\gamma_i+\gamma_m+\gamma_0)\frac{A^2}{(\gamma_i+\gamma_m+A)^2} - \frac{(\gamma_i+\gamma_m)A}{\gamma_i+\gamma_m+A} \\
&\propto (\gamma_i+\gamma_m+\gamma_0)\,A - (\gamma_i+\gamma_m)(\gamma_i+\gamma_m+A) \\
&\propto \gamma_0 - \frac{(\gamma_i+\gamma_m)^2}{A}
\end{aligned}$$

Note that in the above derivations each $\propto$ corresponds to scaling by a positive number. Hence

$$\frac{\partial \pi_g}{\partial \gamma_i} > 0 \Leftrightarrow \gamma_0 > \frac{(\gamma_i+\gamma_m)^2}{A}$$

We can also verify that $\frac{\partial \pi_g}{\partial \gamma_i} = \frac{\partial \pi_g}{\partial \gamma_i}$. □

2.A.2.3 Proposition 3

Proof. We first write f explicitly in terms of parameters by plugging (2.38) and (**??**) into left-hand side of (2.13):

$$f = \frac{1}{4\alpha}\left[log\,(\gamma_i+\gamma_m+\gamma_0) - log\left(\gamma_i+\gamma_m+\gamma_0-\frac{\gamma_m A}{\gamma_m+A}\right)\right] \tag{2.45}$$

Note the connection the above expression has with equation 2.43. Explicitly, f can be written as $\frac{1}{2}\pi_g\left(\gamma_0', \gamma_i', \gamma_m'\right)$ where $\gamma_i' + \gamma_m' = \gamma_m$ and $\gamma_0' = \gamma_0 + \gamma_i$. Therefore

$$\frac{\partial f}{\partial A} = \frac{1}{2}\frac{\partial \pi}{\partial A} > 0,$$

$$\frac{\partial f}{\partial \gamma_0} = \frac{\partial f}{\partial \gamma_i} = \frac{1}{2}\frac{\partial \pi_g}{\partial \gamma_0} < 0,$$

and

$$\frac{\partial f}{\partial \gamma_m} > 0 \quad \Leftrightarrow \quad \gamma_i + \gamma_0 > \frac{\gamma_m^2}{A}$$

□

2.A.2.4 Proposition 4

Proof. From manager indifference condition (2.15),

$$M = \frac{\lambda_i f}{k}$$

Hence

$$\frac{\partial M}{\partial \gamma_i} = \frac{\lambda_i}{k}\frac{\partial f}{\partial \gamma_i} < 0$$

where the last inequality rely on results from Proposition 2.3

And

$$\frac{\partial M}{\partial \lambda_i} = \frac{1}{k}\left[\lambda_i \cdot \frac{\partial f}{\partial \lambda_i} + f\right]$$

The sign of the term inside the bracket is ambiguous since $f > 0$ and $\frac{\partial f}{\partial \lambda_i} < 0$. □

2.A.2.5 Proposition 5

Proof. We first derive the price informativeness absent money manager

$$\eta' = \gamma_0 + \gamma_i - \frac{\gamma_i A}{\gamma_i + A}$$

Conceptually, it is equivalent to the setting where manager signals are white noise, that is, $\gamma_m = 0$. Note that the dependence on λ_i is through A. As such, the problem of maximizing price informativeness fixing amount of investor information can be formulated as

$$\max_{\gamma_i, \lambda_i} \quad \gamma_0 + \gamma_i + \gamma_m - \frac{(\gamma_i + \gamma_m)A}{\gamma_i + \gamma_m + A}$$

subject to

$$\gamma_0 + \gamma_i - \frac{\gamma_i A}{\gamma_i + A} = \bar{\eta}'$$

The constraint can be transformed into a function $A(\gamma_i)$, we can then plug $A(\cdot)$ into the objective and consider its derivative with respect to γ_i. We have

$$\frac{\partial \eta}{\partial \gamma_i} = \frac{\partial \eta}{\partial \gamma_i} + \frac{\partial \eta}{\partial A} \cdot \frac{\partial A}{\partial \gamma_i}$$

$$= \qquad = \left(1 - \frac{A^2}{(\gamma_i + \gamma_m + A)^2}\right) + \frac{\gamma_i^2}{(\gamma_i + A)^2}$$

The marginal rate of substitution between γ_i and λ_i, fixing the amount of investor information, is given by

$$MRS_{\gamma_i,\lambda_i,\eta'} = -\frac{\partial\eta'/\partial\gamma_i}{\partial\eta'/\partial\lambda_i} = -\frac{1-\frac{A^2}{(\gamma_i+A)^2}}{\frac{\gamma_i^2}{(\gamma_i+A)^2}\cdot\frac{\partial A}{\partial\lambda_i}} = \left(1+\frac{2A}{\gamma_i}\right)\cdot\left(-\frac{1}{\partial A/\partial\lambda_i}\right)$$

Similar derivation yields the marginal rate of substitution between γ_i and λ_i, fixing the price informativeness:

$$MRS_{\gamma_i,\lambda_i,\eta} = \left(1+\frac{2A}{\gamma_i+\gamma_m}\right)\cdot\left(-\frac{1}{\partial A/\partial\lambda_i}\right)$$

□

CHAPTER 3. Venture Capital Skill, Size, and Capital Deployment

3.1 Introduction

The venture capital (VC) industry serves a vital role in the growth of the economy. Though the relative size of the VC industry is small compared to other asset classes[1], startups receiving VC funding go on to become some of the largest companies in the world. Given the potential impact VCs have on the economy through growth and innovation, understanding how VCs create value and identifying superior VCs is important.

However, the uniqueness of VC operation and limitation on return data make the evaluation of VCs difficult. Unlike other forms of investment companies such as mutual funds, payoffs of VC portfolio companies are sporadic (at the single point of exit) and skewed (a small fraction of start-ups account for the majority of the gains). Therefore it is difficult to evaluate VC based on return data.

Moreover, the “skill” of VCs depends on many things. Some common traits of a successful VC include the ability to screen promising start-ups from others and providing value-added expertise and support. In this paper, we highlight another important component of VC skill: access/deal flows. While both value-added and screening are important for VC, both assume an existing pool of start-ups for VC to select from and improve upon. In reality, different VCs face different options. The most prestigious VC firms tend to get early access to investments. Not only do their connections allow them to discover promising early-stage companies, but their reputations also lead more founders to voluntarily come to their ways.

The co-existence of better deal flow and better screening ability makes the evaluation of VC more difficult. For example, consider Sequoia, one of the most respected firms in the VC industry with a proven track record. Precisely due to its established reputation, it is hard to pin down

[1] In 2018, global VC investments totaled $254B according to “Venture Pulse Q4 2018”, KPMG

the root of its success. On the one hand, the partners at Sequoia likely have years of experience and industry expertise, which will play huge roles in identifying promising start-ups and providing invaluable support. On the other hand, Sequoia receives more than average deal flows because of its reputation. It is not obvious how much one can attribute Sequoia's success to each of the above components.

How do we distinguish better deal flows from better screening abilities? In which aspects do funds differ more? Can we evaluate VCs without relying on return data? In this paper, we answer the above questions by developing and structurally estimating a dynamic model of VC fundraising and capital deployments. Deal flows/access is modeled as the arrival probability of investment opportunities (capital commitment) during the investment (fundraising) stage. Screening ability is modeled as the precision of the signal on the NPV of a prospective start-up. Fund size and deployment frequency are endogenously determined based on these parameters.

Both network and evaluations affect the size and frequency of fund deployment. When funds have a better network, they meet with prospective startups more frequently, thus shortens the average wait time between deals. The network effect also affects fund initial size through its impact on the fundraising speed – funds that raise funds quicker set a higher initial fundraising target.

On the other hand, better screening leads to longer waiting times between deals. For VCs that are aware of their abilities to spot a better investment opportunity in the future, there is an incentive to be more selective regarding start-ups that they encounter. In contrast, to a VC with limited screening capability, all start-ups (are perceived to) have similar prospects. Hence there is little gain from reserving capital for future investments. This leads to shorter waiting times. On the other hand, because screening capability has little impact on the fundraising stage, its influence on fund size is orders of magnitude smaller than that of the network effect.

We then use actual VC deal data from VentureXpert to estimate the model parameters. For each fund, based on the actual dates of the deal, we can obtain the time intervals between consecutive deals. We then obtain estimates for deal flow and screening ability as the outcomes of maximum likelihood estimators.

The results suggest the difference in deal flows across funds is much more significant than the difference in screening abilities. Variation in deal flows alone accounts for almost the entire dispersion in estimated fund values. This is consistent with the notion that connection in VC is key.

We also conduct the counterfactual experiment where we assume the deal flows and the screening technology of any given VC is separable. By reallocating greater deal flows to funds with better screening technology, we increase the overall fund value by 2.5%. The improvement in fund value corresponds to the unrealized synergy between better screening technology and greater deal flows.

Our paper contributes to the VC skill literature. similar to Sorensen, we apply a structural model to pass the lack of return data. We can separately identify two attributes of VC: deal flow and screening.

The rest of the paper is organized as follows. Section 3.2 sets up and solves the problems for a VC fund. In Section 3.3, we describe the data used in the exercise. Section 3.4 outlines our estimation procedures. Section 3.5 discusses the estimation results. Section 3.6 concludes.

3.2 Model

There are two distinct stages in a VC's lifecyle– fund raising and capital deployment. We model the dynamic decisions of an individual VC fund in both stages. Time is discrete. The state variable is capital level K. For simplicity, we assume K only takes on integer values.

During the fund raising stage, in each period, κ units of capital arrive with probability λ. When the fund target size, K_0, is reached, the fund is closed to investors and shifts from fund raising to deploying capital. Here the subscript 0 corresponds to time 0 in the investment stage. This transition is irreversible. (We leave aside the ability of the fund manager to raise follow-on fund e.g., "X Fund II").

The deployment of K_0 units of capital is as follows. At each date, t, with probability λ, a single investment opportunity arrives. With probability $1 - \lambda$ no opportunity arrives. Note that λ is the arrival rate for both the fundraising and investment stage. We interpret λ as the network and reputation of the VC fund partners. A well-connected VC partner has better access to both LPs when it comes to raise money, and to greater deal flows when it comes to investing.

If an investment opportunity arrives, the VC receives a signal s_t on the NPV of the project y_t. The pair (s_t, y_t) are i.i.d across periods with the following joint normal distribution:

$$\begin{bmatrix} y_t \\ s_t \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_y \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & \rho\sigma_y \\ \rho\sigma_y & 1 \end{bmatrix}\right)$$

The marginal distribution of y_t is $y_t \sim N(\mu_y, \sigma_y^2)$. This reflects the quality of the overall pool of the potential investment opportunities. We can also compute the distribution of $y_t|s_t$, VC's posterior distribution on NPV after observing the signal:

$$y_t|s_t \sim N\left(\mu_y + \rho\sigma_y s_t, \quad (1-\rho^2)\sigma_y^2\right) \tag{3.1}$$

Here $\rho \in [0, 1]$ is the correlation coefficient between the signal s_t and the actual NPV y_t. ρ is the quality of the signal and represents the VC's ability to evaluate a prospective start-up . When $\rho = 0$, s_t provides no additional information on y_t. This can also be seen by plugging $\rho = 0$ into the conditional distribution of $y_t|s_t$ and verify that it is identical to the marginal distribution. On the other hand, if $\rho = 1$, the value of y_t can be precisely inferred by the VC. This can be seen by verifying that $var(y_t|s_t) = 0$ when $\rho = 1$.

Note that the assumptions that the marginal distribution for s_t is $N(0, 1)$ and that the correlation coefficient $\rho \geq 0$ are without loss of generality. This is because linear transformation of normal signals convey the same information.

Based on the signal, s_t, the VC chooses whether to invest. For simplicity, we assume the investment size of each opportunity is fixed at one unit of capital. If the opportunity is pursued, $K_{t+1} = K_t - 1$. Therefore, the capital level K will only take on integer values. Investment continues until $K_T = 0$, at which point the VC portfolio includes exactly K_0 different investments. (The uninvested capital does not earn the any return. In fact, typically the capital is only committed by the LP and provided when there is an investment; a "cash call.")

The VC is risk neutral. It's goal is to maximize the total discounted value of expected NPVs.

We can characterize the optimal solution working backwards through the two key decisions. First we solve for the optimal investment policy of the VC with K units of capital staring at an investment opportunity. Then, knowing the policy functions and values function from the investment stage, we can solve for the fund size as an optimal stopping point problem. We first lay out the solution for the investment stage.

3.2.1 Investment

The investment decision is characterized by how a VC with K units of capital reacts to an investment signal of s_t. For simplicity, let x_t denote the conditional mean of the NPV. From (3.1), we have:

$$x_t = \mu_y + \rho\sigma_y s_t \tag{3.2}$$

Therefore the ex ante distribution on conditional mean NPV is given by

$$x_t \sim N(\mu_x, \sigma_x^2) \tag{3.3}$$

where

$$\mu_x = \mu_y + \rho\sigma_y E(s_t) = \mu_y \tag{3.4}$$

$$\sigma_x = \rho\sigma_y\sigma_s = \rho\sigma_y \tag{3.5}$$

Note that σ_x is increasing in ρ. Consider a VC with superior screening capability, i.e. $\rho \approx 1$. As a result, ex ante, the conditional mean of NPV largely depends on the realization of the signal s_t. Therefore the variance is large ($\sigma_x^2 \approx \sigma_y^2$). On the contrary, if the signal reveals little information, the conditional mean of NPV will be close to the unconditional mean μ_y, regardless of the realization of s_t.

Because the VC is risk neutral and pairs (y_t, s_t) are i.i.d, x_t is the sufficient summary statistics for VC to make the investment decision at time t. Note the VC's problem is equivalent to the following setup without signals: at each period, an opportunity may arrive with known NPV x_t drawn from $N(\mu_x, \sigma_x^2)$. The VC then chooses to either receive x_t and deploy one unit of capital, or wait for better draws in the subsequent periods.

Let $V(K,x)$ denote the fund's value function with capital level K and inferred mean NPV x if an opportunity arrives; let $V(K,\emptyset)$ denote the value function where no opportunity arrives. Let $\bar{V}(K)$ denote the ex ante value function at capital K when the arrival of the investment opportunity is unknown. We have:

$$\begin{aligned}
\bar{V}(K) &= \lambda E(V(K,x)) + (1-\lambda)V(K,\emptyset) && (3.6)\\
V(K,x) &= \max_{w\in\{0,1\}} \begin{cases} x+\beta\bar{V}(K-1) & \text{if } w=1 \\ \\ V(K,\emptyset) & \text{if } w=0 \end{cases} && (3.7)\\
V(K,\emptyset) &= \beta\bar{V}(K) && (3.8)\\
\bar{V}(0) &= V(0,\emptyset) = V(0,x) = 0 \quad \text{for all } x && (3.9)
\end{aligned}$$

Equation (3.6) follows from the iterated law of expectation. w in equation (3.7) denotes whether the investment is made. If the investment is made, the value of fund becomes the sum of the conditional expected NPV x_t and the discounted continuation value with $K-1$ units of capital. $\beta = \frac{1}{1+r}$ is the discount factor. Alternatively, if the VC chooses to wait, fund value is the same as if the opportunity did not arrive, which is the ex ante value discounted by one period, as given in equation (3.8). Equation (3.9) specifies the boundary condition. When capital level is depleted, fund value becomes 0.

Given the monotonicity in this set up, the optimal policy will be characterized by a cut off, $x(K)$, such that it is optimal to make the investment if and only if $x \geq x(K)$. By smooth pasting, at $x(K)$, VC is indifferent between investing and waiting. From (3.7) and (3.8) we have

$$x(K) = \beta\left(\bar{V}(K) - \bar{V}(K-1)\right) \tag{3.10}$$

We can also re-express $V(K,x)$ in terms of threshold $x(K)$:

$$V(K,x) = \begin{cases} x+\beta\bar{V}(K-1) & \text{if } x \geq x(K) \\ \\ \beta\bar{V}(K) & \text{if } x < x(K) \end{cases} \tag{3.11}$$

Taking expectations with respect to x on both sides, plugging in (3.10), and simplifying

$$\begin{aligned} E[V(K,x)] &= Prob(x \geq x(K)) \cdot \left[E[x|x \geq x(k)] + \beta\bar{V}(K-1)\right] + Prob(x < x(K)) \cdot \beta\bar{V}(K) \\ &= \int_{x(K)}^{\infty} (x - x(K))dF(x) + \beta\bar{V}(K) \end{aligned}$$

The second term denotes the baseline continuation for low realizations of x. The first term is the expected incremental gains when $x \geq x(K)$.

We can then plug the above in (3.6) and obtain a recursive expression of $\bar{V}$ in terms of itself:

$$\begin{aligned} \bar{V}(K) &= \lambda E(V(K,x)) + (1-\lambda)V(K,\emptyset) \\ &= \lambda \int_{x(K)}^{\infty} (x - x(K))dF(x) + \lambda\beta\bar{V}(K) + (1-\lambda)\beta\bar{V}(K) \\ &= \lambda \int_{x(K)}^{\infty} (x - x(K))dF(x) + \beta\bar{V}(K) \end{aligned}$$

which yields

$$\bar{V}(K) = \frac{\lambda}{1-\beta} \int_{x(K)}^{\infty} (x - x(K))dF(x) \tag{3.12}$$

Equation (3.10) and (3.12) is the law of motion for the pair $(X(K), \bar{V}(K))$. We can solve $X(K), \bar{V}(K)$ for any K using backward induction, starting from the boundary condition $\bar{V}(K) = 0$. Formally, we have the following result.

PROPOSITION 3.1. *Given a set of parameters $\{\mu_y, \sigma_y, \rho, \lambda, \beta\}$, the law of motions in (3.10),(3.12), and the boundary condition (3.9) $\bar{V}(0) = 0$, there exists a unique sequence of $x(K)$, $\bar{V}(K)$, $K = 1, 2, 3, \ldots$. Furthermore, $x(K) > x(K+1) > 0$ for all K. That is, fund becomes more selective as capital is depleted.*

Proof. By induction. The initial case is given by the boundary condition $\bar{V}(0) = 0$. We can also use (3.12) to show that $x(0) = \infty > 0$. ($x(0)$ never appears in the VC's problem, but we state it here for consistency)

We now show the inductive step. Suppose we have computed the unique values of $\bar{V}(K)$ and $x(K) > 0$ for some K. Our goal is to show that there exists unique $\bar{V}(K+1)$ and $x(K+1) \in (0, x(K))$ satisfying the law of motions. We first plug $\bar{V}(K)$ into (3.10) and yields

$$x(K+1) - \frac{\beta\lambda}{1-\beta} \int_{x(K+1)}^{\infty} (x - x(K+1))dF(x) + \beta\bar{V}(K) = 0 \tag{3.13}$$

The goal is to show that there exists unique solution z for the following equation:

$$h(z) = z - \frac{\beta\lambda}{1-\beta}\int_z^\infty (x-z)dF(x) + \beta\bar{V}(K) = 0 \tag{3.14}$$

First, note that $\int_z^\infty (x-z)dF(x)$ is decreasing in z: let $z_1 < z_2$,

$$\int_{z_1}^\infty (x-z_1)dF(x) > \int_{z_1}^\infty (x-z_2)dF(x) > \int_{z_2}^\infty (x-z_2)dF(x)$$

Therefore, since $x(K) > 0$ by the inductive hypothesis

$$\begin{aligned} h(0) &= 0 - \frac{\beta\lambda}{1-\beta}\int_0^\infty (x-0)dF(x) + \beta\bar{V}(K) \\ &< \frac{\beta\lambda}{1-\beta}\int_{x(K)}^\infty (x-x(K))dF(x) + \beta\bar{V}(K) \\ &= 0 \end{aligned}$$

The last equality follows from (3.12).

Similarly, we can show $h(x(K)) = x(K)$. Since h is continuous, by the Intermediate Value Theorem, there exists $x(K+1) \in (0, x(K))$ such that $h(x(K+1)) = 0$. Moreover, h is strictly increasing. Thus the solution must be unique. □

3.2.2 Fundraising

Given the projected sequence of $\bar{V}(K)$, the VC fund determine the optimal amount of capital, K_0, to begin with. In the context of VC, the fund raising stage is only gathering commitments; capital does not flow into the fund until needed for deployment (i.e. cash call). However, the distinction between capital inflow and capital commitment is insignificant here. As we will see, the key trade-off in the fund raising stage is not the opportunity

Let $W(K)$ be the value function during the fundraising stage at the start of period t, where K is the units of capital raised so far. We have

$$W(K) = \max_{w\in\{0,1\}} \begin{cases} \beta\left(\lambda W(K+\kappa) + (1-\lambda)W(K)\right) & \text{if } w = 1 \\ \bar{V}(K) & \text{if } w = 0 \end{cases} \tag{3.15}$$

where w represents the binary decision of whether to devote another period to search for a commitment. If $w = 1$, the fund remain in the fundraising stage. With probability λ a LP arrives with κ units of capital. Otherwise, if $w = 0$, the fund closes to investors and shift immediately to investing stage.

Here the trade-off is between the chance of additional units of capital and the cost of delaying the payoffs. This is rooted in the feature that fund must first close before making investments. Contrast this with the investment decision in (3.7), where the trade-off also arise from the finite amount of total resource.

Similar to the investment stage, we conjecture the optimal fundraising policy is characterized by a threshold size K_0. Specifically, the fund will keep raising capital if and only if the existing level of capital $K < K_0$. As such,

$$W(K) = \begin{cases} \beta\left(\lambda W(K+\kappa) + (1-\lambda)W(K)\right) & \text{if } K < K_0 \\ \bar{V}(K) & \text{if } K \geq K_0 \end{cases} \tag{3.16}$$

The optimal fund size K_0 can be solved with the following iterative procedure over $K \in \{0, \kappa, 2\kappa, \ldots\}$. We begin with $K = 0$. In each iteration, we compare the policy that stops at K and the policy that stops at $K + \kappa$. Under the first policy, $W(K) = \bar{V}(K)$. Under the second policy:

$$\begin{aligned} W(K) &= \beta\left(\lambda W(K+\kappa) + (1-\lambda)W(K)\right) \\ &= \beta\left(\lambda \bar{V}(K+\kappa) + (1-\lambda)W(K)\right) \\ &= \frac{\beta\lambda}{1-\beta+\beta\lambda}\bar{V}(K+\kappa) \end{aligned}$$

Therefore the policy corresponding to K dominates the policy corresponding to $K + \kappa$ if and only

$$\bar{V}(K) \geq \frac{\beta\lambda}{1-\beta+\beta\lambda}\bar{V}(K+\kappa) \tag{3.17}$$

. We first show that there exists K such that the above is true. Rearranging terms in (3.17) we arrive at the following equivalent condition

$$(1-\beta)\bar{V}(K) \geq \beta\lambda\left(\bar{V}(K+\kappa) - \bar{V}(K)\right) \tag{3.18}$$

Plugging in (3.12) and (3.10) into the LHS and RHS, respectively, yields:

$$\int_{x(K)}^{\infty}(x - x(K))dF(x) \geq \sum_{i=0}^{\kappa-1} x(K+i) \tag{3.19}$$

which is true since $\int_{x(K)}^{\infty}(x - x(K))dF(x) \to E[x|x \geq 0]$ and $x(K) \to 0$ as $K \to \infty$[2] . Let K_0 be the smallest multiple of κ such that the above holds. By default, the stopping rule corresponding to K_0 dominates that for all $K < K_0$. Furthermore, since LHS is increasing in K and the RHS is decreasing in K, 3.17 also holds for $K > K_0$. Therefore the stopping rule characterized by K_0 is the global optimal.

3.2.3 Numerical Examples: Deal flows and Screening

Our main goal is to separately identify the deal flow and the screening skill of each VC fund. Therefore, we explore how the fund's initial value and decisions are related to λ, the deal arrival probability, which captures deal flow, and ρ, the signal precision, which captures screening ability. We do so by examining various comparative statics via numerical examples.

First, we examine the initial fund value under different parameterizations of ρ and λ. Specifically, we focus on $\bar{V}(K_0)$, the fund value at the inception of the investment stage, with optimal capital K_0. We summarize our observation.

OBSERVATION 1. *The value of a VC fund at the beginning, $\bar{V}(K_0)$ is increasing in the signal precision ρ and deal flow λ. In addition, ρ and λ have complimentary effect on value, i.e. $\frac{\partial^2 \bar{V}(K_0)}{\partial\rho\partial\lambda} > 0$.*

Figure 3.1 shows the value of $\bar{V}(K_0)$ for different combinations of ρ and λ. Not surprisingly, both greater deal flow and higher precision cause funds to have greater values. Moreover, we see that the same increase in ρ has a larger impact on value when λ is higher. This is because higher deal flows magnify the benefit of better screening ability.

[2]To see that $x(K) \to 0$ as $K \to \infty$, note that $x(K)$ is decreasing and bounded below by 0, hence must have a limit. By (3.10), the limit of $x(K)$ must equal to the limit of the RHS, which is 0

While both λ and ρ are positively related to fund value, they have different qualitative effects on investment policy, which then lead to different implications on time between successful deals. We summarize the result in the following observation.

OBSERVATION 2. *All else equal, the probability that an investment is made at any level K is decreasing in ρ and increasing in λ. The optimal fund size K_0 is increasing in λ and decreasing in ρ. Though it seems* $\left|\frac{\partial K_0}{\partial \lambda}\right| >> \left|\frac{\partial K_0}{\partial \rho}\right|$

Figure 3.2 illustrates the optimal investment policy under different parameterizations over ρ and λ. The top left panel plots the NPV threshold that determines investment. Each curve corresponds to one parameterization. Consistent with Proposition 3.1, We see that as capital K decreases to 0, the threshold increases. Funds become more selective as they have lower capital

Comparing the curves representing $\rho = 0.3$ to curves representing $\rho = 0.7$, wee see that NPV threshold increases when ρ increases. For VC that can identify better start-ups, they require higher perceived NPVs to give up 1 unit of capital that could be invested in better startups in the future. On the flip side, if the fund has limited screening ability, all the start-ups have similar perceived NPV, hence the threshold NPV is lower.

On the other hand, comparing curves representing $\lambda = 0.3$ and $\lambda = 0.7$, we see that the increase in λ increases threshold, though with smaller magnitude. This is due to the indirect effect that future projects may arrive sooner, so it makes waiting more appealing, then raises the minimum NPV required.

The upper right panel translates the NPV threshold into the probability of investment. That is, the lines represents $P(x > x(K))$. Since both increases in λ and ρ results in a higher threshold, they lead to a lower probability of investment. Everything for λ carries through. For ρ, however, in addition to the effect on x(K), it also directly impacts the distribution of x. Higher ρ increases the spread of the distribution of x, thus increasing $P(x > x(K))$ for any $x(K)$. Panel B indicates the net effects of ρ on P is still negative.

While the previous panel plots the probability of invest conditional on the arrival of opportunity. The ex-ante probability of whether a deal is made would also need to account for the arrival rate.

The bottom panel plots $\lambda \cdot P(x > x(K))$, which is the ex-ante probability of investing in the period before arrival is realized. This is ultimately what determines the time between deals. Here, the effect of ρ is the same as the conditional probability. However, for λ, the effect on ex-ante probability through higher arrival rate leads to a net positive effect. Therefore we see that fund with higher λ have higher chance to invest at any given point, whereas fund with higher ρ has lower.

The distinctive qualitative effects of λ and ρ on fund investment probability provide one unique ingredient for our identification. However, investment alone is not sufficient. Given frequent successful deals, it is hard to distinguish between a fund with large (low) deal flow and a fund with noisy (precise) signals. To do so we also consider the impact on fund size.

In figure 3.3, each curve plots the optimal fund size for various value of ρ, fixing λ. The negative slope indicates funds with higher ρ choose lower fundraising targets. This is because the increased precision leads to higher expected payoffs for future projects. Therefore the cost of waiting and raising an additional unit of capital becomes greater.

On the contrary, as indicated by the distance between lines, λ is positively correlated with fund size. Though higher λ, like higher ρ, raises the benefit of future fund value thus the cost of waiting, it also expedites the fundraising process since LP commitments arrive more frequently. The latter effect dominates and the net effect of better access overall is a larger fund.

More importantly than the different signs of $\frac{\partial K_0}{\partial i\lambda}$ and $\frac{\partial K_0}{\partial i\lambda}$ is the disparity in their magnitudes. The slope of each line, especially for high ρ and low λ, is relatively flat, indicating the limited effect ρ has on optimal fund size. This is because ρ only affects optimal size indirectly through future value. λ, on the other hand, directly affects the rate at which capital is raised, thus have a much larger effect.

Conceptually, investment frequency and fund size provide a system of two equations based on ρ and λ, which we can then solve for. In the estimation section, we will do exactly this.

3.3 Data

The data we use in this paper is VentureXpert (also known as ThomsonOne), which is one of the main academic sources of data on investments by venture capital funds. The unit of observation is startup-round-investor. For example, if a particular round for a startup has five investors, then there will be five observations. Each observation records the identity of the fund and startup involved in the deal, the date of the deal, the round number, and the total amount of investments. In addition, we observe the state and city for the startup and fund. For each fund, we also have access to the VC firm, its initial size, the founding year of the fund, and the VC firm. The sample period ranges from 1980 to 2015.

3.3.1 Sample Selection

While VentureXpert is one of the most comprehensive datasets on VC deals, it has its drawback as Kaplan et. al. (2002) suggest. Here we outline the necessary steps to clean and preprocess the data. First, we drop all observations with unidentifiable funds. In the dataset, they are either labeled "unspecified fund" or "undisclosed fund". Secondly, we drop observations with missing values. Most missing values appear in the amount invested column.

For our purpose, we want to estimate each fund's skill parameter using its waiting time between deals. Doing so requires us to have complete or near-complete records of all the deals the fund has been involved in. To see this, consider a fund with a portion of its deals missing from the dataset. Running the estimation directly on the incomplete data could lead to the misinterpretation that the fund tends to wait longer between deals.

To determine which funds might have unreported deals, we construct a proxy for the percentage deals reported. For each fund, we collect all deals in which the fund has been a part of. For each deal, we compute the expected investment amount by dividing the total amount invested in the deals by the number of investors. This amount is not exact because co-investors for a deal normally do not invest the same amount. We then aggregate the expected investment amounts for all deals. The proxy for percentage reported is then defined as the ratio between the sum invested and the reported initial fund size. We only retain funds with $> 75\%$ deals matched according to the proxy. Here we use 75% instead of higher because the proxy might be different from the actual ratio.

We also exclude funds whose last observed deal is within 2 years of the end of our sample. This is to exclude funds that likely have not used up their capital, even if the previous proxy rule has not already eliminated them. The 2 years here is a conservative threshold and will likely guarantee all funds in our analysis are terminated funds.

While the remaining funds likely have all deals reported in the dataset, we take additional steps to drop extreme outliers so that the eventual sample is reasonably homogeneous. The main reason, as we will see in the next section, is that the model cannot identify certain parameters such as κ, the average capital committed per LP, or β, the discount factor. As such, we can only interpret the estimation results in relative terms. Doing so requires all funds estimated are similar in most aspects.

First, we drop all funds outside of the US. Secondly, by examining the empirical distribution of the number of deals, we note that more than 75% of funds participate in between 3 and 30 deals. There are masses of funds with 1 or 2 deals. We exclude these because the intervals between deals are either empty or a singleton and thus not enough to analyze. We also include those on the right tail of the distribution in terms of deal invested. There are funds with more than 100 deals. They likely operate under a different investment strategy and style from the rest of the sample. So we avoid pooling them with our core sample. We conduct similar sample selection with the average investment size – we only include funds whose average investment size is between \$0.5M and \$3M (similar to the criterion used for the number of deals, \$0.5M and \$3M correspond to the 25th and 75th percentile among our remaining sample.)

3.3.2 Summary Statistics

Table 3.1 provides the summary statistics at both the fund level and the deal level. From Panel A, The average fund in our sample has \$19.64M capital committed, is founded in 1994 within a VC firm that started 5 years prior in 1989, participates in almost 14 deals in its lifecycle with an average investment of \$1.52M in each company. On average the deal is for the third-round funding. We compare the statistics among our sample to the original sample before any selection. Our sample is not significantly different from the main sample in terms of vintage. However, our sample tends to consist of smaller funds both in terms of the overall asset and asset per investment.

Panel B reports the deal level statistics. Within our sample, the average round injects almost \$8M capital to the underlying company, which is in its 3rd round of funding. Typical deals have a little more than 5 co-investors. The average year of the deal is 1999. Most of these statistics are comparable to those of the entire sample. This suggests our sample is representative at least in regard to the types of deals.

3.4 Estimation Procedure

We conduct the estimation on each fund separately. Below we describe how we bring the model to data, by deriving the likelihood function and formatting data so that it can be used as the input for the likelihood. We also discuss the assumptions we make to achieve identification.

3.4.1 Likelihood function

We first derive the likelihood function. From the model, we see how λ and ρ affects the probability of invest at each time. The probabilities then give rise to a distribution of the waiting time between deals. Let $p_k = \lambda Prob(x \geq x(k))$ be the unconditional probability of investment in a given period when the current level of capital is k. Whether a deal is made in any period is a Bernoulli random variable with probability p_k. As such, the expected waiting time d_k for a deal to arrive and occur has a binomial distribution:

$$Prob(d_k) = (1 - p_k)^{d_k - 1} \cdot p_k \tag{3.20}$$

That is, for the deal at k to take exactly d_k periods, we need $d_k - 1$ periods of non-investment, each with probability $(1 - p_k)$, and the last period to have a successful investment, with probability p_k. We multiply them because of the independence of the draw across periods.

Let $d_1, d_2, \ldots\ldots d_K$ be the periods between deals at different capital levels. By independence, the joint probability of observing the entire sequence $\{d_i\}$ is then the product of (3.20) applied to different values of k. Taking the log yields the following log likelihood function:

$$l_{invest}(\rho, \lambda | d_1, d_2, \ldots, d_K) = \sum_{k=1}^{K} (d_k - 1)\ln(1 - p_k(\rho, \lambda)) + p_k(\rho, \lambda) \tag{3.21}$$

Here we highlight the fact that ρ, λ affect the likelihood through p_k.

However, (3.21) does not provide the complete picture. If $d_1, d_2, \ldots\ldots d_K$ represents all the wait times, then K is the initial size of the fund, which we also need to account for in the likelihood function. From the model, the optimal size of K_0 is the deterministic solution to the stopping time problem in the fundraising stage. That is, $Prob(K_0 = K)$ is either 0 or 1 depending on the parameters. The discontinuity is an undesirable feature for estimation and departs from reality.

We address the discontinuity issue by modifying the model in the following way: suppose that as soon as the fund reached the optimal level of capital K_0 and closes to investors, there is a one time shock that changes the capital level to K_0 to K_0'. Here the discrepancy could arise from certain LP deciding to commit more capital than agreed upon (or pulling back from an agreement). The difference could also capture unexpected windfall or cash withdrawal more generally.

In a reduced form way, we assume K_0' takes on a geometric distribution with K_0 as the expectation. In robustness tests, we can experiment with alternative functional forms. The requirement here is that the support of the distribution is the set of integers so to preserve the discreteness of capital level.

Taken together, the log-likelihood is given by

$$l(\rho, \lambda | d_1, d_2, \ldots, d_K) = (K-1)\ln\left(1 - \frac{1}{K_0}\right) + \ln\left(\frac{1}{K_0}\right) + \sum_{k=1}^{K} (d_k - 1)\ln(1 - p_k(\rho, \lambda)) + \ln(p_k(\rho, \lambda)) \tag{3.22}$$

which can be numerically computed by first solving the model as outlined in the earlier section.

3.4.2 Data Formatting

Recall the unit of observation in VenturExpert is round-investor-company. Our goal is to compute the sequence of wait times $\{d_i\}$, which we can then use to define the likelihood function and then maximize.

For each fund in our dataset, let $t_1, t_2, \ldots, t_K$ be the dates of all investments sorted from the latest to earliest. We can compute the difference in weeks between deals in days and normalize it into weeks. We use weeks to represents periods so that the most wait times lie in the desired range of model outcomes. Partners in VC funds also hold weekly meetings to discuss potential portfolio companies. Hence adopting weeks is also consistent with the decision cycles of funds.

Specifically, we define

$$d_k = \min\left(1, \lfloor \frac{t_{k+1} - t_k}{7} \rfloor\right)$$

Here the unit of $t_{k+1} - t_k$ is in days and are guaranteed to be integers since the deal dates are given. Since the waiting time in our model must also be integers, we use the floor function when transitioning from days to weeks. If consecutive deals occur in the same week, we treat the time difference as 1. Future expansion of the model could also consider the possibility of multiple projects in one period.

Lastly, while we observe K deals for every single fund, we can only compute the $K-1$ wait times. This is because we do not observe the date in which fund closes to the investor and transition to the deployment stage. Therefore we do not observe d_K. Hence in calculating the likelihood function (3.22), we omit the term associated with d_K.

3.4.3 Identification and Parameter Restrictions

So far, we have defined the likelihood function in terms of only ρ and λ. However, the entire set of parameters is $\{\mu_y, \sigma_y, \kappa, \beta, \rho, \lambda\}$. Indeed, we artificially impose restrictions on the degree of freedom. Due to the limitations of our data, this is necessary to ensure identification. Below we describe which parameters we need to fix and provide justifications.

When solving the model, scaling μ_y and σ_y simultaneously result in no change in policy functions. Conceptually, we are only changing the unit of payoff. Since we only match deal frequencies and the total number of deals, we are not able to pin down the absolute scales of μ_y and σ_y. In the actual estimation, we fix μ_y to be one – that is, the unit of the payoff is the expected NPV of all available projects. (Here the implicit assumption is that the mean NPV is nonzero and positive, which is not inconsistent with reality.)

Similarly, we can not separately identify ρ and σ_y. Recall ρ and σ_y affect VC decisions only through $\sigma_x = \rho\sigma_y$, the ex-ante variance of the conditional NPV. Conceptually, we cannot distinguish between imprecise signals from a noisy prior. Here we restrict σ_y to be 10 and let ρ varies. The restriction that $\sigma_y = 10$ imposes an upper bound on all fund's variance. However, from the results in later sections, we see that estimates of ρ do not lie near 1, suggesting the restriction is not binding.

While we cannot use existing data to separately identify μ_y and σ_y, future extensions could incorporate data on the NPVs of the deals, the distribution of which could pin down the μ and σ.

For k and β, though their effects on the model cannot be perfectly replicated by ρ and λ, the effects are close enough that estimating all four parameters simultaneously would cause identification issue. Throughout our estimation, we fix $k = 1$ and $\beta = 0.99$. For future extensions, data on the number of LPs for funds and opportunity cost of the LP can help pin down the estimates of k and β.

3.5 Estimation Results and Analysis

Following the procedures outlined in the previous section, we estimate $\hat{\rho}$ and $\hat{\lambda}$ for each fund. Figure 3.4 is a scatter plot of $\left(\hat{\rho}, \hat{\lambda}\right)$ for all fund in our sample.

3.5.1 Dispersion in Deal Flows vs. Dispersion in Screening

We compare the dispersion in $\hat{\lambda}$ and the dispersion in $\hat{\rho}$. This addresses the question that we raised at the beginning: do VC funds differ more in their ability to screen prospective startup or in their ability to generate larger deal flows? Below we present several pieces of evidence suggesting the dispersion in deal flow is greater and have a much larger impact on fund value.

Figure 3.5 plots both the 2D and 1D densities for $\hat{\lambda}$ and $\hat{\rho}$ over all funds. By inspecting the shape of the marginal densities on each estimated parameters, we see that $\hat{\rho}$ has a unique mode near 0.25. In contrast, for $\hat{\lambda}$, all values in $[0.1, 0.3]$ share the maximum density.

Furthermore, the mean of $\hat{\lambda}$ and $\hat{\rho}$ are 0.231 and 0.263, respectively. The standard deviations are 0.123 and 0.017. Judging by the standard deviation and the mode, funds appear to be more different in terms of deal flows. However, one could argue the magnitude here since λ and ρ are different objects – one is the probability, the other correlation. In particular, recall we could have assumed a smaller prior uncertainty σ_y, which would mechanically scale all the $\hat{\rho}$ and effectively increase the standard deviation.

To address the above concern, we use $\bar{V}(K_0)$, fund value at the beginning of the investment stage, as the common metric. First, we compute $\bar{V}(K_0)$ at 4 sets of (λ, ρ): $\left(E\hat{\lambda}, E\hat{\rho} \pm \sigma_{\hat{\rho}}\right)$ and $\left(E\hat{\lambda} \pm \sigma_{\hat{\lambda}}, E\hat{\rho}\right)$ and label the result on Figure 3.5. We see that 2 sample standard deviation in λ is able to generate a difference of 61.4-19.1 = 42.3 in fund value; whereas 2 sample standard deviation in ρ can only generate a difference of 41.7-37.4 = 4.3 in fund value. Therefore, the dispersion in λ is greater that in ρ in terms of the spread in corresponding fund values.

Alternatively, we can directly plot the density of the estimated fund values. Figure 3.6 and 3.7 compare the actual estimated density with the counterfactual densities where the variation in λ or ρ is muted. Specifically, by setting λ (ρ) to $E\hat{\lambda}$ ($E\hat{\rho}$), we examine the variation in fund value that screening (deal flow) alone can generate. Consistent with the previous results, variation in deal flow accounts for almost all the variation in fund value: $var(\bar{V})$ actually increases from 472.7 to 275.9; on the other hand, $var(\bar{V})$ when λ is fixed is only 4.7.

This result is consistent with the notion that connection is the most valuable resource in the VC industry. This is a clear distinction between VC and other types of investments. For comparison, consider mutual funds specializing in public equity, fixed-income, or commodity. Each firm more or less has access to the same pool of assets (though the connection with traders would often result in better deals). In the above scenarios, screening plays a much more important role.

3.5.2 Optimal Sorting of Deal Flow and Screening

Here we consider the following thought experiment: if each fund's screening capacity and deal flow are separable and transportable, what is the optimal reallocation of skills across funds so that the aggregate fund value is optimized?

Recall from Observation 1 that deal flows and screening ability are complements in that $\frac{\partial^2 V}{\partial \lambda \partial \rho} > 0$. As such, the optimal reallocation would match the best deal flow with the best screening technology. That is, we would to sort both $\hat{\lambda}$ and $\hat{\rho}$ and assign the highest of both technologies to the same VC firm.

Optimal reallocation results in an increase in the average fund value of 1.03 from 40.03 to 41.06, a 2.57% increase. While this magnitude might be small, recall there is little variation in ρ. This suggests the potential loss in synergy in other samples could be much larger.

That being said, we should keep in mind the above result is predicated on the assumption that deal flows and screening are separable. While this is the case in our model since both λ and ρ are exogenous parameters, they can be related in intricate ways. For example, it is reasonable to assume both generating deal flows and screening projects are time- and labor-intensive. To establish connections and improve access, VC partners need to dedicate a significant amount of time attending/organizing networking events or acting as judges on demos or competitions. On the other hand, the necessary due diligence required to accurately assess a start-up's potential is also time-consuming. In this regard, we cannot simply exchange one VC's deal flow with others without impacting their screening abilities.

3.6 Conclusion

We model the dynamic decisions of a VC fund. By mapping the model outcome with actual data on deals, we separately identify the screening ability and deal flows of each fund. Most variation in fund values stems from the difference in deal flows. Counterfactual analysis suggests a potential room for value creation by matching better screening technology with greater deal flows.

We recognize the various potential drawbacks of the current setup. For one, by estimating all funds jointly as a panel, we could potentially identify more model parameters. Another direction for future research is to enrich the model by incorporating heterogeneous effects based on various fund characteristics. For example, the NPV distribution could be related to the state and the vintage of the VC fund.

Despite all the limitations, the paper contributes to the literature on VC evaluation by examining VC decisions through a unique lens. The unique feature of our work is that we do not rely on any return data. As such, we provide a fresh perspective on VC skills in relationship to existing measurements.

3.A Figures and Tables

Panel A: Fund Level Statistics

	Selected Sample					Full Sample				
	mean	std	25%	50%	75%	mean	std	25%	50%	75%
fund size ($M)	19.34	21.64	4.60	11.00	25.60	262.68	935.65	10.70	50.50	200.00
fund year	1994.60	10.12	1986	1997	2001	1997.07	10.41	1989	1999	2005
firm founded year	1989.01	14.20	1983	1992	1999	1988.29	16.44	1982	1990	1999
number of deals	13.91	7.90	7	13	20	20.93	34.52	2	8	25
average investment size	1.52	0.70	0.91	1.43	2.08	10.32	65.52	0.88	2.10	4.26
average deal round	3.25	1.28	2.33	3.08	4.00	3.10	1.71	2.00	3.00	4.00
average # of co-investors	5.27	2.57	3.53	4.78	6.37	4.68	2.78	2.83	4.28	6.00
N	1296					8210				

Panel B: Deal Level Statistics

	Selected Sample					Full Sample				
	mean	std	25%	50%	75%	mean	std	25%	50%	75%
RoundAmountDisclosedThou	7944.07	11375.66	1500	4100	10000	12301.30	54917.96	1500	4800	11875
RoundNumber	3.34	2.48	1	3	4	3.62	2.68	2	3	5
RoundNumberofInvestors	5.15	3.97	2	4	7	5.02	4.00	2	4	7
year	1999.36	8.94	1994	2000	2006	1998.20	9.60	1990	2000	2006
N	18022					171803				

Table 3.1. Descriptive Statistics Panel A summarizes the statistics on fund levels. Fund size, fund year, firm founded year are taken directly from the VenturExpert dataset. Number of deals are computed based on identifiable deals within the dataset. Average deal round, average deal round, and average number of investors are calculated based on characteristics of observed deals for each fund. Panel B provides the summary statistic at the fund level. All features are taken from the dataset directly. Here RoundAmountDisclosed is in thousands and represents the total investment among across all co-investors. For both Panel A and Panel B we compare the full sample and our selected sample.

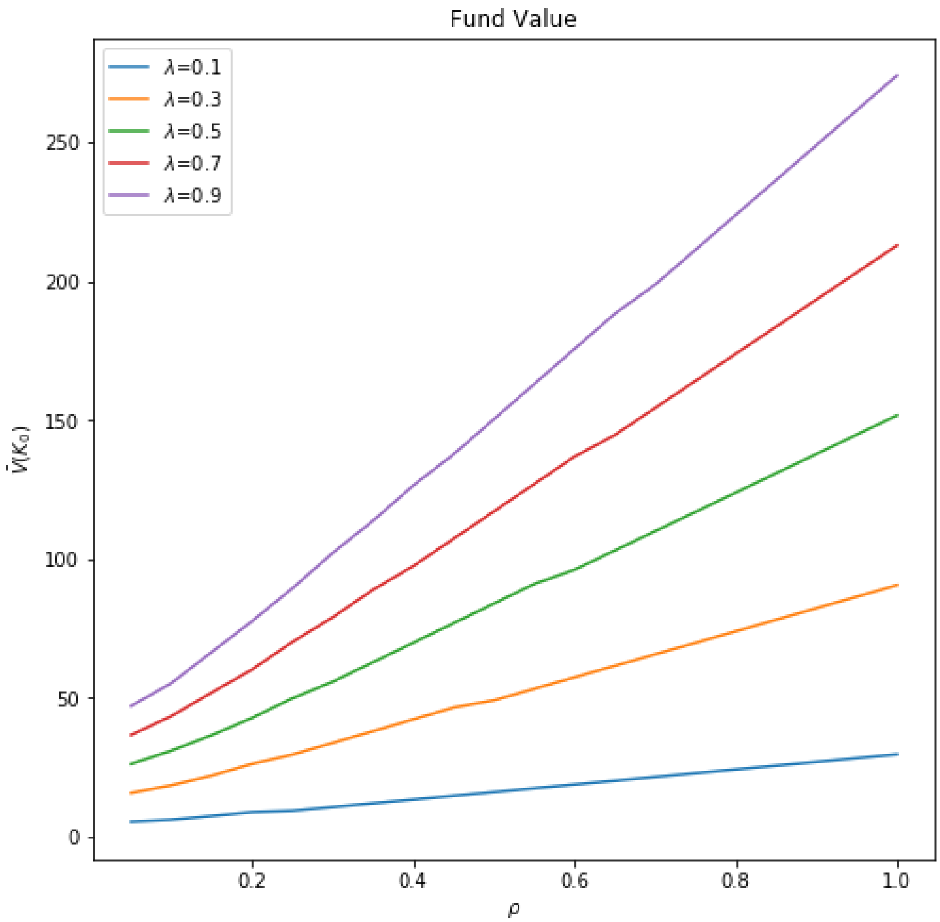


Figure 3.1. Initial Fund Value Here we plot $\bar{V}(K_0)$ under various combinations of ρ, arrival rate, and λ, signal precision. Not surprisingly, both greater deal flow and higher precision causes fund to have higher value. Moreover, we see that the same increase in ρ has a larger impact on value when λ is higher. This is because higher deal flows magnifies the benefit of better screening ability.

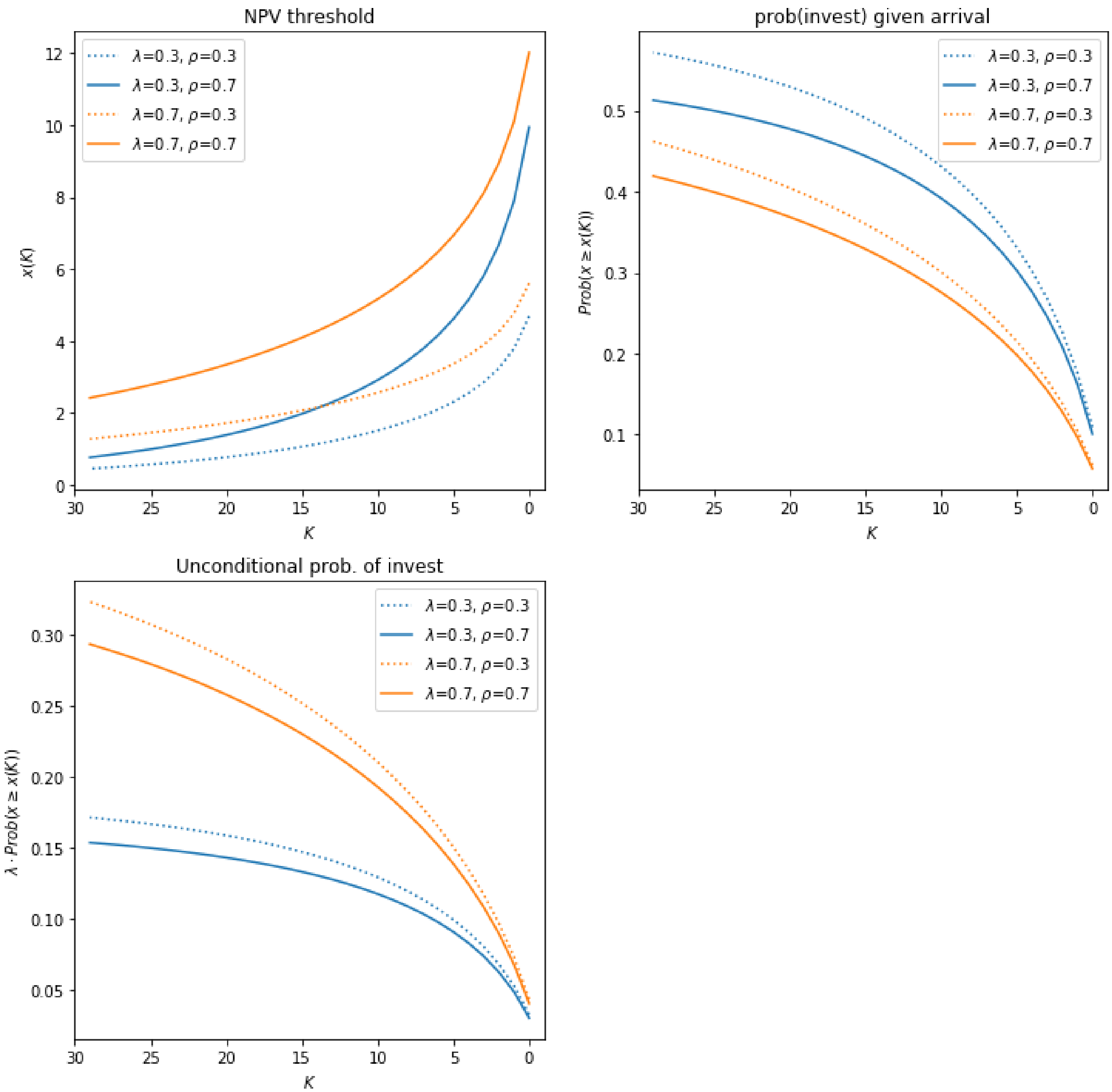


Figure 3.2. Investment Policy Here we illustrate the optimal investment policy under 4 different combinations of parameters. As shown in Proposition 3.1, policy varies with the current capital level K, which is on the axis. The values plotted are $x(K)$, the threshold used to determine whether the project is pursued; $P(x > x(K))$, the probability of investing conditional on the arrival of investment opportunity; and $\lambda P(x > x(K))$, the ex ante probability of invest.

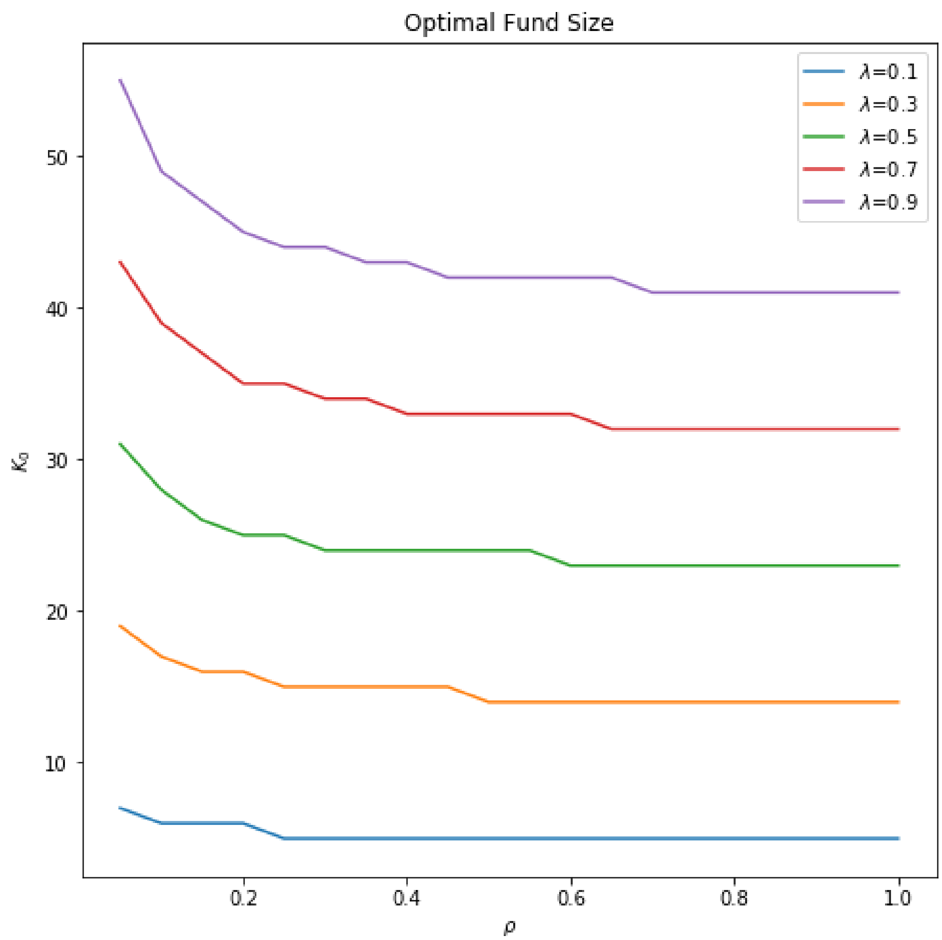


Figure 3.3. Optimal Size We plot K_0, the optimal initial size of the fund under different parameter valus for λ and ρ. Looking along each line, we see fund size is larger for smaller ρ.

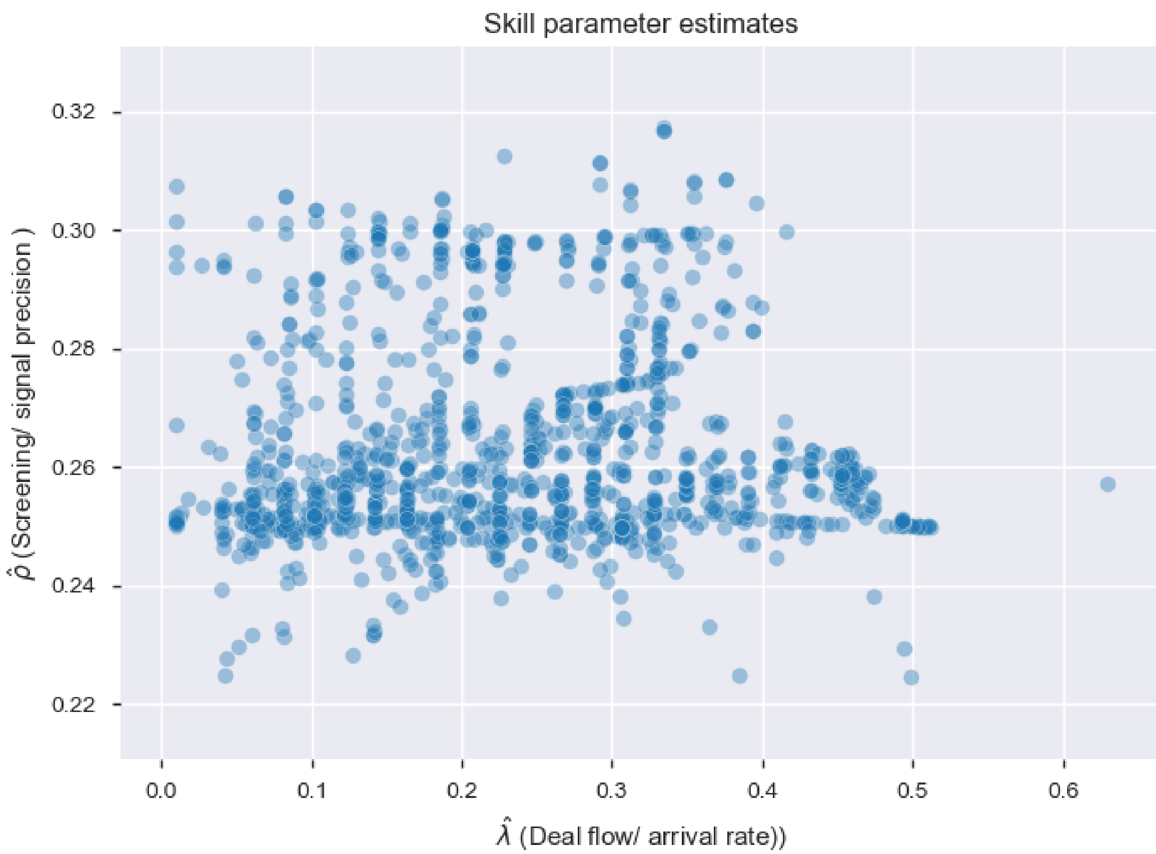


Figure 3.4. Estimation Results in Scatter Plots Here we plot $(\hat{\rho}, \hat{\lambda})$ for all funds in our sample.

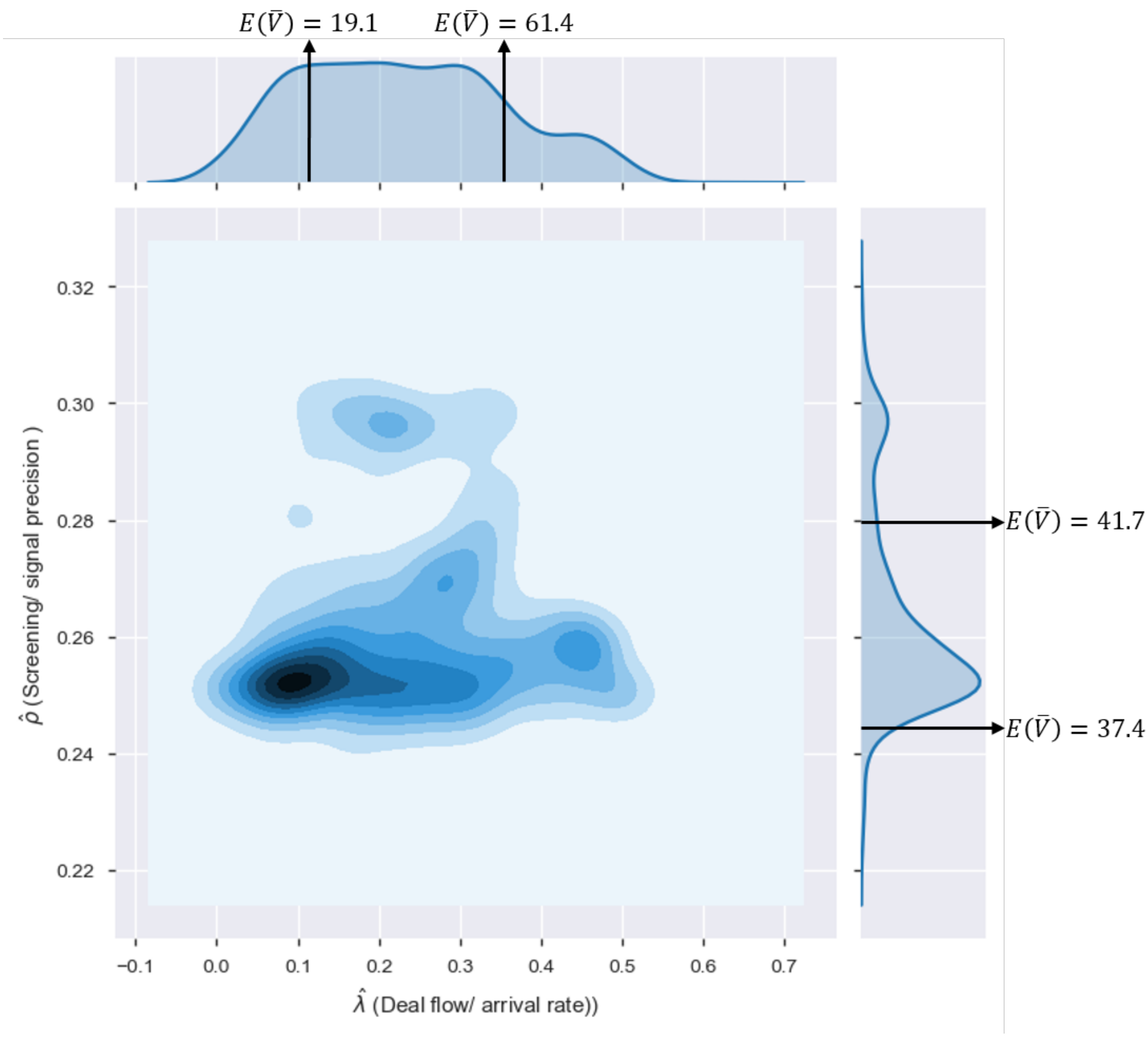


Figure 3.5. Density Plot We include both 2D density plot along with the marginal density associated with each $\hat{\rho}$ and $\hat{\lambda}$. On the marginal density plots we also compute and label $\bar{V}(K_0)$ at 4 sets of (λ, ρ): $\left(E\hat{\lambda}, E\hat{\rho} \pm \sigma_{\hat{\rho}}\right)$ and $\left(E\hat{\lambda} \pm \sigma_{\hat{\lambda}}, E\hat{\rho}\right)$. The positions of the labels correspond to mean $\pm$ 1 standard deviation for each ρ and λ

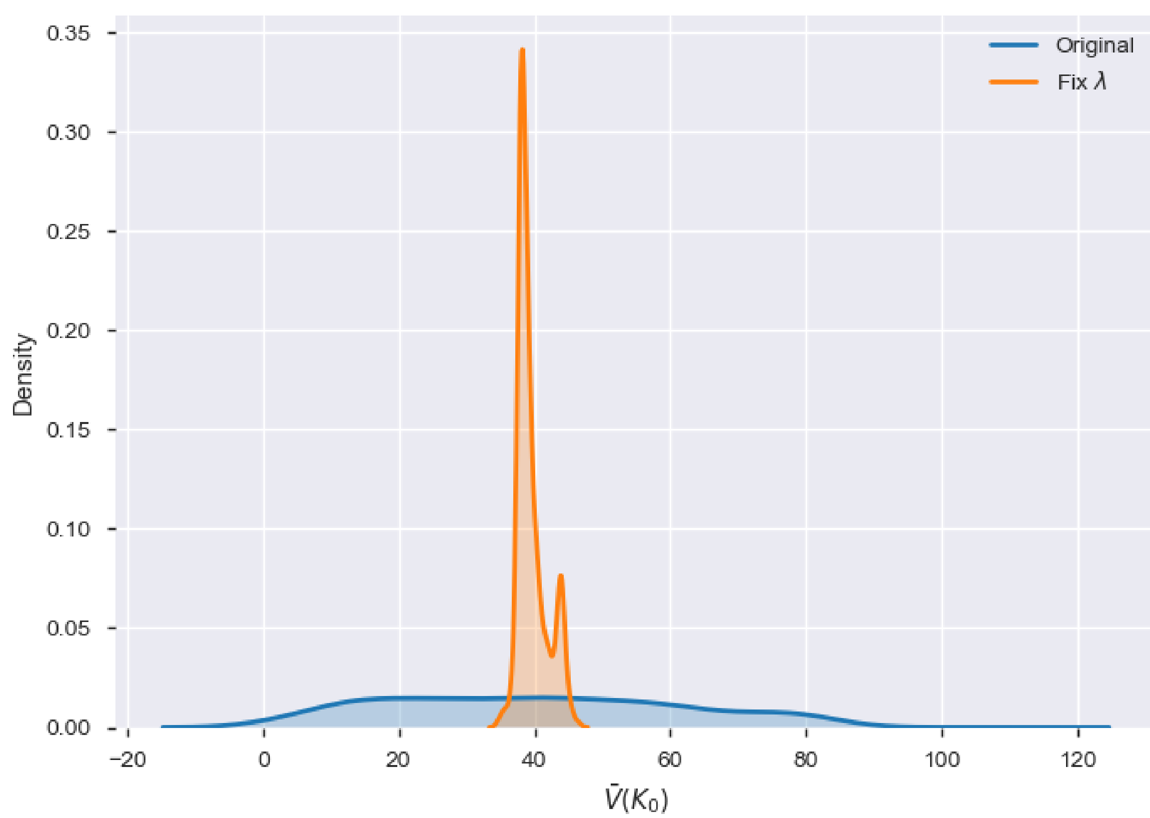


Figure 3.6. Density Comparison, fixing λ The blue density is the estimated fund values for all funds with variance of 472.7. The orange density correspond to fund values where λ is set to be mean($\hat{\lambda}$). The resulting fund values become much more concentrated, with a variance of 4.7

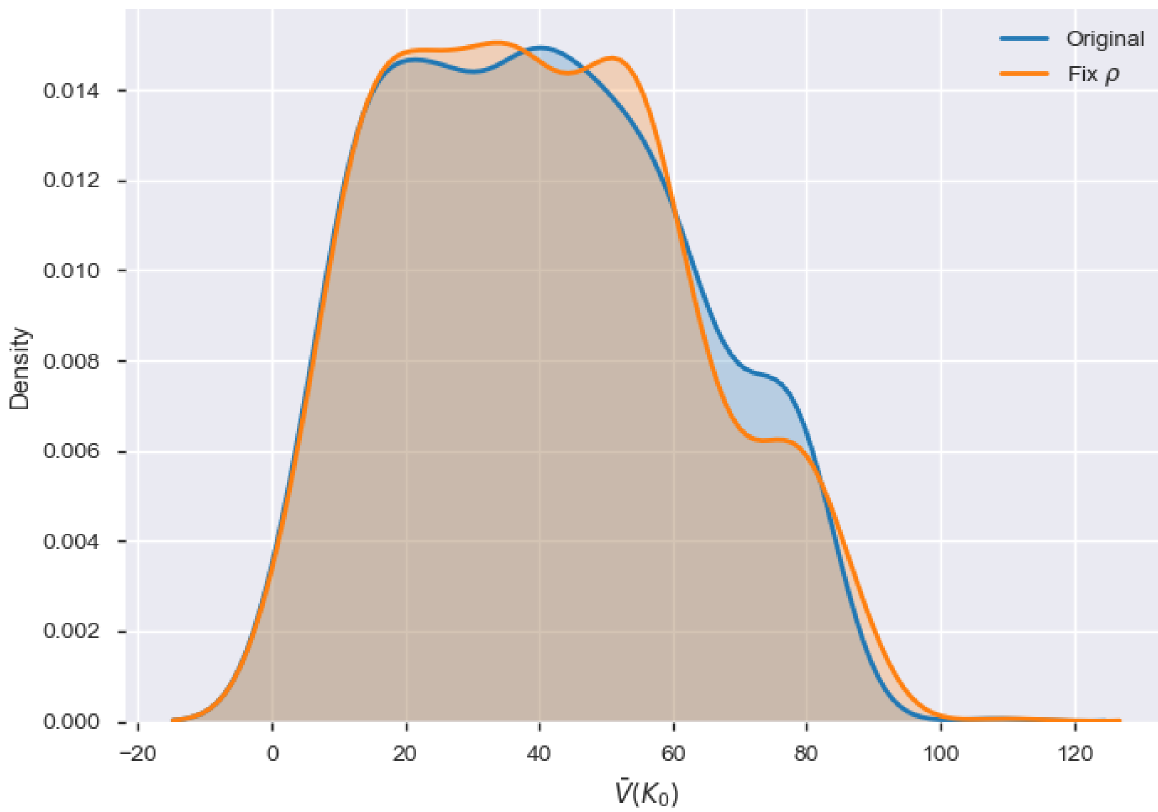


Figure 3.7. Density Comparison, fixing ρ The blue density is the estimated fund values for all funds with variance of 472.7. The orange density correspond to fund values where ρ is set to be mean($\hat{\rho}$). We observe virtually no loss in the spread in the resulting fund values, with variance of 475.9

BIBLIOGRAPHY

Admati, A. R. and Pfleiderer, P. (1987). Viable allocations of information in financial markets. *Journal of Economic Theory*, 43(1):76–115.

Antweiler, W. and Frank, M. Z. (2004). Is All That Talk Just Noise ? The Information Content of Internet Stock Message Boards. *Journal of Finance*, LIX(3).

Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *RAND Journal of Economics*, 40(2):209–233.

Aydogdu, M. and Wellman, J. W. (2011). The Effects of Advertising on Mutual Fund Flows : Results from a New Database. *Financial Management*, pages 785–809.

Bailey, W., Kumar, A., and Ng, D. (2011). Behavioral biases of mutual fund investors. *Journal of Financial Economics*, 102(1):1–27.

Bergstresser, D., Chalmers, J. M. R., and Tufano, P. (2009). Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry. *Review of Financial Studies*, 22.

Berk, J. B. and Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, 112(6):1269–1295.

Berk, J. B. and van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1):1–20.

BERK, J. B., van BINSBERGEN, J. H., and LIU, B. (2017). Matching Capital and Labor. *The Journal of Finance*.

Beshears, J., Choi, J. J., Madrian, B. C., and Beshears, J. (2009). How Does Simplified Disclosure Affect Individuals' Mutual Fund Choices?

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1):57.

Carlin, B. I. (2009). Strategic price complexity in retail financial markets. *Journal of Financial Economics*, 91(3):278–287.

Choi, J. J., Laibson, D., and Madrian, B. C. (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, 23(4):1405–1432.

Cronqvist, H. (2006). Advertising and Portfolio Choice.

FAMA, E. F. and FRENCH, K. R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, 65(5):1915–1947.

French, K. R. (2008). Presidential address: The cost of active investing. *Journal of Finance*, 63(4):1537–1573.

GENNAIOLI, N., SHLEIFER, A., and VISHNY, R. (2015). Money Doctors. *The Journal of Finance*, 70(1):91–114.

Gerakos, J., Linnainmaa, J. T., and Morse, A. (2016). Asset Managers: Institutional Performance and Smart Betas.

Goldstein, I. and Yang, L. (2015). Information Diversity and Complementarities in Trading and Information Acquisition. *Journal of Finance*, 70(4):1723–1765.

Grossman, S. and Stiglitz, J. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3):393–408.

Hillert, A., Niessen-ruenzi, A., and Ruenzi, S. (2014). Mutual Fund Shareholder Letter Tone - Do Investors Listen ?

Hortacsu, A. and Syverson, C. (2004). Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S & P 500 Index Funds. *The Quarterly Journal of Economics*, 119(2):403–456.

Hwang, B. H. and Kim, H. H. (2017). It pays to write well. *Journal of Financial Economics*, 124(2):373–394.

Kacperczyk, M., Nieuwerburgh, S. V., and Veldkamp, L. (2014). Time-varying fund manager skill. *Journal of Finance*, 69(4):1455–1484.

Kacperczyk, M., Van Nieuwerburgh, S., and Veldkamp, L. (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica*, 84(2):571–626.

Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791.

Lewellen, J. (2011). Institutional investors and the limits of arbitrage. *Journal of Financial Economics*, 102(1):62–80.

Loughran, T. and Mcdonald, B. (2011). When Is a Liability Not a Liability ? Textual Analysis , Dictionaries , and 10-Ks. *Journal of Finance*, LXVI(1).

Loughran, T. and Mcdonald, B. (2016). Textual Analysis in Accounting and Finance : A Survey. *Journal of Accounting Research*, 54(4):1187–1230.

Manconi, A. and Spalt, O. G. (2012). Learning by Doing and the Value of Experience for Mutual Fund Managers. *SSRN Electronic Journal*.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2017). Portfolio Liquidity and Diversification: Theory and Evidence.

Pedersen, L. H. and Garleanu, N. B. (2015). Efficiently Inefficient Markets for Assets and Asset Management. *NBER Working Paper*, XXXIII(2):81–87.

Veldkamp, L. L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.