Essays on Delegated Asset Management

Fabian Garavito

Ph.D. Thesis

Department of Finance

London School of Economics and Political Science

2010

A thesis submitted to the Department of Finance of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, June 2010

List of Tables

List of Figures

Contents

Declaration

I certify that the thesis I have presented for examination for the Ph.D. degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

My thesis explores issues concerned with delegated asset management and their implications for asset prices. The first chapter of my thesis documents how the organizational form of a mutual fund affects its investment strategies. I show that decentralized funds allocate a greater portion of their capital to soft (opaque) information companies than centralized funds. Consistent with the inability of centralized organizations to handle soft information, I find that decentralized funds are better at investing in soft information companies than hierarchical funds. Furthermore, I find that high levels of ownership by decentralized funds predict high future returns for soft information companies. The second chapter shows that while fund families may use mutual fund incubation (the creation and management of a fund before it is offered to the public) in an opportunistic fashion, they also seem to use it to foster innovation. I document that fund families tend to launch their incubated funds when their past performance is high, consistent with a behaviour that aims to exploit the convex relationship between past performance and current flows. However, I also show that, after their Initial Public Offerings (IPO's), incubated funds tend to hold more illiquid stocks, hold more concentrated portfolios, invest in less-popular securities and are better at purchasing stocks than non-incubated funds. This difference in investment strategies is due to the incubation period as it allows managers to explore different corners of the market without having to take into account performance-induced capital flows. I also present evidence

that, despite their outstanding pre-IPO track records, incubated funds attract a smaller share of new-fund capital flows than non-incubated funds. The third chapter (joint work with Dr. Mungo Wilson) shows that mutual fund families cater to investors' demand by offering funds with investment styles that are in vogue. I also show that this catering exacerbates the "dumb money" effect. In other words, fund IPOs have a positive effect on the persistence of investment style capital flows.

Para Sofía, con mucho amor.

Acknowledgements

I would like to thank my supervisors Christopher Polk and Michela Verardo for all their help and time throughout my studies. Their guidance, advice and help were instrumental during my PhD. I am specially grateful to the Financial Service Authority (U.K.) for providing funding during my PhD studies through the FSA PhD Scholarship in Economics and Financial Regulation. I would also like to thank my colleagues and friends Miguel Anton and Vincent Fardeau for their friendship, help and time spent in endless conversations about research, academia and football.

I would also like to thank my family for their unbounded love, encouragement and unconditional support. Without Laura and Sofia's love, smiles, support, patience and encouragement, I would have been unable to finish my PhD. I would also like to thank my parents and my brother for their love and support. Despite the distance, they have always been with me every step of the way. Special thanks to my aunt Leonor and my uncle Alonso who have always believed in me and who have been source of inspiration and pride.

Lastly, I would like to thank Antoine Faure-Grimaud. Antoine gave me the opportunity to join the programme and went out of his way to help me every time I needed a helping hand. Besides my profound admiration for his research work, I was deeply touched by his dedication to all students. He was very generous with his time despite his very tight agenda and multiple responsibilities and was very caring and helpful with all his students. He made a huge impact in my life and I will always be grateful for those long talks, his advise, his good sense humour, and his guidance. Merci Antoine.

Chapter 1

Organizational Diseconomies in the Mutual Fund Industry

1.1 Introduction

Information collection is a central part of investing. As suggested by several economists, the collection of information can ameliorate the adverse selection problem faced by investors[1]. Moreover, it has been argued that the collection of information may be influenced by organizational structure. For instance, it has been documented that decentralized organizations are better suited at processing soft information (i.e. information that is not quantitative in nature and that is difficult to summarize in a numeric score), since this type of information can not be transferable within a hierarchy. On the other hand, standardization and economies of scale make centralized organizations more efficient at producing and processing hard information (information that is quantitative in nature). The main objective of this paper is to study the effect of organizational structure on information collection and investing in the mutual fund industry.

[1]Stiglitz and Weiss (1980) show that asymmetric information may explain why capital does not flow to firms with positive net present value projects. Lelan and Pyle (1977), Campbell and Kracaw (1980), Diamond (1984), Haubrich (1989) Diamond (1991) describe how large institutional creditors can partially overcome the problem of adverse selection by producing information about firms and using it in their credit decisions.

Stein (2002) models the effect of organizational design (centralized vs decentralized structures) on the collection of information in firms where division managers compete for internal funds to finance their projects. In a centralized organization, where research and capital allocation are carried out by different agents, division managers (research) will know that the only information a CEO will use in allocating capital across divisions is information that can be credibly transmittable. This type of information, also known in the literature as hard information, is quantitative and objective in nature (i.e. sales growth rate of a company over the past 5 years. This, in turn, means that any effort exerted in collecting information that can not be credibly transmittable, or soft information (information that can not be easily agreed upon, i.e. honesty of a CEO), will go to waste. Division managers will know this ex-ante and will re-direct all their efforts to collect hard information. Furthermore, competition amongst division managers for limited internal funds will lead them to collect as much hard information as possible on their investment projects. This will, in turn, translate into the production of vast amounts of hard information. On the other hand, in a decentralized organization, research and capital allocation will be conducted by the same agent (division managers collect information on projects and decide how to allocate capital). In this case, division managers have the incentive to collect as much (hard and soft) information as possible on their investment prospects in order to minimize the adverse selection problem. Stein shows that if all information about investment projects is hard, centralized organizations have an advantage over decentralized firms in their ability to distribute capital across divisions within the organization. On the other hand, if all information available about investment opportunities is soft, decentralized organizations will have superior fund allocation across projects than centralized firms due to their ability to collect and use soft information.

My findings are consistent with Stein's model. I document that the level of centralization of a mutual fund's organizational structure positively covaries with the degree with which a fund invests in hard information companies (companies for which most of the information is hard). In other words, as the organizational structure changes from decentralized to centralized, so do the incentives to collect hard over soft information. This explains the larger portfolio tilt of centralized funds towards hard information companies. I also exploit the difference in the ability to collect and use soft and hard information of two different organizational forms (decentralized and centralized) by constructing buy-and-hold self-financing trading strategies that yield positive risk-adjusted returns. For instance, a self-financing buy-and-hold trading strategy that buys a basket of soft information stocks held by decentralized funds and short-sells a basket of soft information stocks held by centralized funds every quarter, produces risk-adjusted returns between 49 and 64 bps per month (6% and 7.95% per year respectively). Moreover, consistent with the hypothesis that decentralized organizations have a unique ability to collect and use soft information, I illustrate that high levels of ownership by decentralized funds predict high future returns for soft information companies. For instance a one standard deviation in the intensity of investment by decentralized funds in a soft information stock, predicts an additional future return of 17 bps per month.

The distinction between soft and hard information has been studied before in the banking literature, with particular emphasis on the incorporation of soft and hard information in different lending technologies (i.e. credit scoring, relationship lending) by different organizational forms (large vs. small banks)[2]. One of the main conclusions in this strand of the literature is that large banks tend to be at a disadvantage when lending to small businesses. The reason given is that large banks are very centralized and small businesses tend to be informationally opaque (they mostly produce soft information). The disadvantage emerges from the fact that centralized organizations are ill-suited to use soft information [3]. However, Berger Rosen and Udell (2007) argue that

[2]For a more detailed discussion on the subject see the papers surveyed in Berger Rosen and Udell (2007).

[3]However, Berger and Udell (2006) have pointed out that large banks (hierarchies) may have developed lending technologies that allow them to lend to opaque businesses (soft information companies). Examples of these lending technologies are small business credit scoring asset-based lending, factoring, fixed assets lending and leasing (See Berger and Udell (2006)).

past empirical research in this area is inconclusive since some variables of interest were not considered. Therefore the evidence in the banking literature on the effects of organizational form on the collection and usage of information is mixed.

Actively managed US equity mutual funds provide an ideal environment to analyze the effects of organizational diseconomies in information collection and capital allocation for several reasons. Firstly, investing is a task that is information intensive. Secondly, due to disclosure requirements, it is possible to measure fund organizational characteristics, and the information opaqueness (hard vs. soft information) of funds' holdings. In this paper, I construct information and organizational complexity scores to measure the softness of the information generated by stocks and the degree of centralization of mutual funds.

Chen et al (2004) look at the issue of organizational diseconomies in the delegated asset management industry. They examine a particular cross-section of the data, September 1997, and find that small and solo-managed funds are more likely to invest in and are better at choosing local stocks (companies whose headquarters are geographically close to the fund's main offices) than large and non-solo managed funds. They present this evidence as an indication that decentralized funds are better at collecting and using soft information companies. However, they do not look at the other implications of Stein (2002), namely whether centralized funds tilt their portfolios to hard information stocks. Moreover, their measure of organizational complexity overlooks the effect of fund families in the way member funds operate. For instance, Gaspar Massa and Matos (2006) and Cici Gibson and Moussawi (2006), document that fund families have the incentives and mechanisms to influence the capital allocation of its member funds. This creates the kind of separation between research and decision making found in centralized organizations. In a concurrent project, Massa and Zhang (2009) study how the organizational structure of an asset management company affects its strategies and performance. Using a sample of US fixed income mutual and insurance-managed funds, the authors show that more hierarchical structures invest less in firms located close to them and deliver lower performance.

The paper is organized as follows. Section 2 describes the hypotheses. Section 3 describes the data, section 4 presents the results and section 5 concludes.

1.2 Theory and Hypothesis Construction

1.2.1 Soft and Hard Information

Petersen (2004) presents a detailed characterization of hard and soft information in finance. Hard Information is the kind of information that can be easily reduced to numbers. Examples of hard information in finance are financial statements, credit history, and stock returns. On the other hand, one can think of soft information as information that can not be completely summarized in a numeric score. Examples of soft information in finance can be opinions and rumours. Due to its quantitative nature, hard information can be easily collected, stored and transmitted (these characteristics also make it difficult to contain). A second dimension used by Petersen to characterize information is the way in which it is collected. The collection of hard information need not be personal. Therefore, the collection process can be at arms length, automated and standardized. However, it places restrictions on what can be collected. With soft information, the context under which it is collected and the collector of the information are part of the information itself. For instance, if I say that the manager of a firm has great business acumen, the information depends on my definition of business acumen. One of the advantages of hard information is that it can lower production costs through standardization and automatization. Hard information is easy to store as the information does not depend on who collected it. This means that the information remains in an organization even if the agent who collected it leaves the firm. However, collection of hard information also leads to a loss of information which in some contexts can be quite important (i.e. venture capital, where most of the information about investments is soft). Moreover, the fact that hard information is difficult to contain can keep managers from fully collecting informational rents (i.e. in the case of equity investing, it reduces the ability an investor has to earn abnormal returns).[4]

1.2.2 Hypotheses construction

Motivated by Stein (2002), I conjecture that the organizational form of a mutual fund affects managers' incentives to collect information. This should be reflected in the kind of stocks that managers pick and in their ability to choose stocks with different degrees of information "softness".

Hypothesis 1. The level of mutual fund centralization should positively covary with the tendency to hold a greater proportion of hard information companies.

Hypothesis 1 follows from the fact that incentives to collect information are affected by organizational form. As stated before, fund managers that operate in centralized funds (where research and decision making are conducted by different agents), know that they will not be able to credibly transmit soft information. Ex-ante, they decide to steer their efforts towards collecting more hard information. Therefore, as the organizational structure becomes more hierarchical, funds increase the share of resources allocated to hard information stocks.

Hypothesis 2. Decentralized (centralized) funds are better than centralized (decentralized) funds at investing in soft (hard) information companies.

If decentralized funds are better suited than centralized funds at collecting and using soft information, they should have a superior ability to invest in soft information stocks than centralized funds. On the other hand, centralized funds should be better than decentralized funds at investing in hard information companies since they are better at gathering and incorporating hard

[4]Petersen also notes that soft information can be hardened and cites credit scoring as an example. In addition, he presents examples of hardening of information (Mercantile Agency, R.G.Dun, and Bradstreets in the 1840) and explains how the evolution of financial markets over the last forty years has been in part a replacement of soft information with hard information as the basis for financial transactions

information.

Hypothesis 3. Decentralized funds' ownership of soft information stocks should forecast these stocks' future returns

Since decentralized funds are able to collect and incorporate soft information in their investment decisions, high decentralized funds' ownership in soft information stocks should forecast higher expected returns since high ownership should reflect the collection of positive information on a stock.

1.3 Dataset Construction and Methodology

1.3.1 Information variables and score construction

The information score I construct aims at measuring the information "softness" of a firm and the extent to which information about a company has been hardened. The score is based on four variables. The first two variables, market capitalization (SIZE) and age (AGE), measure the information softness of a firm. These variables have been previously used in the banking literature and the main idea is that information available about older and larger firms tends to be harder than information generated by younger and smaller companies. The other two variables, number of analyst forecasts (NUM EST) and institutional ownership (OWN), measure the extent to which information about a company has been hardened. The basic premise is that these two variables measure the level of due diligence on a company. For instance, it is plausible to think that there is more hard information about a company followed by 50 analysts than by a company without analysts' coverage.

1.3.1.1 Information Score

The construction of the information score for each stock is based on the ranking of its information variables. To try and purge the size effect from the other information variables (all other variables are highly correlated with size), I orthogonalize institutional ownership, age and number of estimates with respect to size. I conduct this orthogonalization by regressing age, institutional ownership and number of estimates on size, similar to Hong et al (2000) and Nagel (2005). The orthogonalized values of these variables are the error term of these regressions. Since institutional ownership is bounded between 0 and 1, it is necessary to transform the variable so that it maps to the real line. I perform the following logit transformation,

$$Logit(OWN) = \log\left(\frac{OWN}{1 - OWN}\right) \tag{1.1}$$

where the values below 0.0001 and above 0.9999 are replaced with 0.0001 and 0.9999 respectively. The information variable residuals are calculated by regressing the information variables on Log Size and squared Log size. On average, across all quarters, I find the following relations:

$$Logit(OWN) = -7.31 + 1.68 Logsize - 0.09 Logsize^2 + \epsilon \tag{1.2}$$

$$NUMEST = 2.13 - 1.90 Logsize + 0.43 Logsize^2 + \epsilon \tag{1.3}$$

$$LogAGE = 4.44 - 0.25 Logsize + 0.035 Logsize^2 + \epsilon \tag{1.4}$$

The basic procedure to calculate the information score is as follows. Each quarter, I take the universe of NYSE stocks and rank them in 20 groups by size and residual age. I use the NYSE size and residual age rank cut-off points to rank stocks held by mutual funds. I also rank stocks held by mutual funds in 20 groups by residual institutional ownership and residual number of analysts' estimates. For each stock, I then calculate an aggregate information variable by summing up the ranks of the four variables for each stock. For instance, if a stock belongs to size group 1, residual age group 2, residual institutional ownership group 2 and residual number of analyst estimates group 10, the aggregate information variable equals 15. Next, I group these stocks by this aggregate information variable in deciles. The information score will be equal to the aggregate information variable decile.

1.3.2 Mutual Fund Variables and Hierarchy Score

The main objective of the hierarchy score is to measure the organizational complexity of a fund. As mentioned earlier, there are two dimensions to consider, the centralization of tasks within the fund and the organizational complexity of the fund family to which the fund belongs. To measure the level of organizational complexity within the fund, I use is the number of managers that control the asset allocation (NUM MGRS) and net assets under management (AUM). As far as number of mangers, the premise is that funds with many managers will tend to be team-managed. This, in turn, causes managers to decide on an asset allocation based on consensus by sharing, and thus transferring, information with other agents. Regarding net assets under management, the main idea is that larger organizations are more hierarchical, since large organizations tend to centralize activities. This variable has also been used in Chen et al (2004) to measure the organizational complexity of mutual funds. The third and fourth variables I use are the number of funds (NUM FUNDS) and the total assets under management in actively managed US equity funds (FAM SIZE) of the family a fund belongs to. These variables are motivated by papers in fund cross-subsidization and fund proliferation (Massa (2003) and Gaspar Massa Matos (2006)). The idea is that maximizing fee income at the family level is different from maximizing fee income at the fund level. This will lead families to cross-subsidize funds that are the most likely to benefit from the convex relationship between past performance and current net flows documented in Chevalier and Ellison (1997) and Sirri and Tufano (1998). Families will also try to enhance the performance of funds that maximize the positive spill over effect that a top-performing fund has on the its member funds' net flows (Nanda et al (2004)). Therefore fund families have incentives to cross-subsidize fund returns in other to maximize their own income fee. As such, it is reasonable to believe that the asset allocation of a fund will be influenced by its family. This creates the separation between research and fund allocation mentioned in Stein (2002).

1.3.2.1 Hierarchy Score

As indicated above, the variables Number of Funds, Family Size and AUM are highly and positively correlated. Therefore a hierarchy score based on the raw values of the hierarchy variables would not provide much extra information. For instance ranking funds by AUM would produce a very similar ranking if funds are ranked by Family size. I therefore, orthogonalize the number of funds with respect to AUM. I also orthogonalize Family Size with respect to AUM and Number of Funds. On average, across all quarters, I find the following relations:

$$LogNUMFUNDS = 1.64 + 0.074LogAUM + 0.013LogAUM^2 + \epsilon \qquad (1.5)$$

$$\begin{aligned} LogFAMSIZE &= 2.04 + 0.40LogAUM + 0.0004LogAUM^2 \\ &+1.81LogNUMFUNDS - 0.034LogNUMFunds^2 + \epsilon \end{aligned}$$

Each quarter, I rank funds by AUM, residual number family funds, and residual family size in quintiles. I calculate an aggregate hierarchy variable by summing up these rankings and the number of managers (which takes values from one to five). For instance, if a fund belongs to AUM group 1, residual number of funds group 2, residual family size group 2 and it has 3 managers, the aggregate hierarchy variable equals 8. Each quarter, I rank funds by this aggregate hierarchy variable in deciles. The hierarchy score will be equal to its aggregate hierarchy variable decile.

1.3.3 Data

I construct my sample of US actively managed equity funds in four steps using the CRSP Survivor-Bias Free U.S. Mutual Fund Database (fund characteristics) and from the Thompson Financial Mutual Fund Database (fund holdings). First, index funds are filtered out from the CRSP Survivor-Bias

Free U.S. Mutual Fund Database. Secondly, we eliminate funds (FUNDNO'S) that have reporting gaps of more than 12 months from the Thompson Financial Mutual Fund Database (S13)[5]. Thirdly, we exclude funds that on average do not hold at least 80% of NYSE-AMEX-NASDAQ common stocks from the Thompson Financial Mutual Fund Database (S13). Lastly, we merge these two datasets by using the MFLINKS provided by WRDS. The number of shares own by a fund on a particular issue are corrected and expressed as of the date of disclosure[6].

Data on stock returns and prices are from the Center for Research in Security Prices (CRSP) Monthly Stocks File for NYSE, Amex, and NASDAQ stocks. I eliminate closed-end funds, real estate investment trusts (REIT), American Depository Receipts (ADR), foreign companies, primes, and scores. I exclude stocks below the 20th NYSE size percentile from the tests that look at stocks returns due to the well-documented asset-pricing anomalies in small stocks (Griffin and Lemmon (2002)). Market capitalization is defined as the product between share price and shares outstanding. Age is defined as the number of months that a security is present in the CRSP Monthly File. Data on institutional holdings are obtained from the Thomson Financial Institutional Holdings (13F) database. I extract quarterly holdings from the first quarter of 1993 to the last quarter of 2006. I calculate the share of institutional ownership by summing the stock holdings of all reporting institutions for each stock in each quarter. Stocks that are on CRSP, but without any reported institutional holdings, are assumed to have zero institutional ownership[7]. The number of analysts' estimates is calculated using I/B/E/S. At the end of a company's fiscal year, I count the maximum number of one-year EPS

[5]These funds are suspect as Thompson usually recycles unique fund identifiers. See WRDS User's Guide to Thomson Reuters Mutual Fund and Investment Company Common Stock Holdings Databases (July 2008)

[6]Portfolio holdings are collected every quarter (vintages). If a fund has not disclosed holdings (because it was late or because it was not supposed to), the data provider fills in "missing quarters" by carrying forward the holdings of the prior quarter adjusting for a range of corporate events (i.e. Stock splits). For more information, see Wermers (1999) and Gompers and Metrick (2001)

[7]The number of shares owned by an institutional investor are also corrected by the late-filer problem described above.

estimates that were outstanding during the fiscal year in question. NUM EST is the maximum number of one-year EPS estimates in the most recent fiscal year.

As stock return predictors, I use book-to-market (B/M), firm-level volatility (VOL), and turnover (TURN). The book value of equity in the nominator of B/M is taken from the Compustat Database, and it is defined as in Daniel and Titman (2006). At the end of each quarter t, we calculate B/M as the book value of equity from the most recent fiscal year-end that is preceding quarter-end t by at least six months divided by the market value of equity at the end of quarter t. Consistent with Fama and French (1993), I exclude firms with negative book values. Returns on stocks are obtained from CRSP and are corrected for delisting biases as suggested by Shumway (1997) and Shumway and Warther (1999)

Organizational characteristics of mutual funds are taken from the CRSP Survivor-Bias-Free US Mutual Fund Database. I identify all share classes issued by a mutual fund using the MFLINKS provided by WRDS and calculate the mutual fund characteristics at the fund level, not at the share class level. The sample starts in the first quarter of 1993 and ends in the last quarter of 2006. I calculate the number of managers by counting the different names in the database's manager field. The funds in my sample have a maximum of 5 names in the database manager filed. However, for some funds, the manager field is set to "Team Managed". If this is the case, I set the variable "number of managers" equal to five. Assets under management equals the total TNA of the fund's share classes. The number of actively managed US equity funds per family is calculated at the end of each quarter. The family size variable is the sum of all actively managed US equity funds' TNA offered by a family.

1.3.4 Summary Statistics

Table 1.1 presents some summary statistics on the variables used to construct the information score (Size, Age, Number of Analyst Estimates and Institutional Ownership). It also contains summary statistics on stock return predictors book-to-market (B/M, its natural logarithm), firm-level volatility (VOL), and turnover (TURN). All statistics are calculated cross-sectionally each quarter and are then averaged across time. These statistics are calculated for stocks held by mutual funds. A few points a noteworthy. First, it is important to note the strong positive correlation between size and all other information variables. This is the reason why purging the size effect from the other information variables is important when constructing the score. Otherwise the information score would only capture variation in market capitalization of stocks in my sample. The information score (INFO) positively correlates with all the information variables as expected. The information score also covaries negatively with past volatility (VOL), and performance (RET12), and positively with liquidity (TURN) and book-to-market ratio (Log BM). In order to better understand the information score, I also analyze the returns of each of its deciles. Table 1.2 describes the risk-adjusted excess returns of value weighted portfolios of stocks in each information score group and their loadings on a set of factors. We can see that the main facts from the correlation structure in Table 1.1 are reflected in this table as well. For instance, we can see that the portfolio of stocks with information score equal to zero (INFO = 0) load more on SMB (smaller companies), UMD (momentum stocks) LIQ (Illiquid stocks) and ST REV (short-term reversal stocks) and less on HML (high book-to-market companies) than stocks in the information score equal to nine (INFO = 9). There also seems to be a negative difference between the risk-adjusted excess returns of the stocks in the bottom and top deciles of the information score, albeit the statistical significance is not very high.

Table 1.3 presents some summary statistics on the variables used to construct the hierarchy score and on other mutual fund variables of interest. We can see that the number of managers is not highly correlated with any other hierarchy variable. However, NUM FUNDS, AUM and Log FAM SIZE are highly correlated. The hierarchy score covaries positively with all hierarchy variables and with the value-weighted information score of each fund. This is the first sign that centralized funds tilt their portfolios towards hard information companies. It is also important to note that the correlation between the hierarchy score (HIERARCHY) and a concentration measure (W) is negative. In other words, centralized funds tend to be less concentrated that decentralized funds. The hierarchy score also covaries positively with fund age which indicates that young funds tend to be less hierarchical. In order top better understand the hierarchy score, I analyze the return characteristics of each hierarchy quintile. Table 1.4 describes the risk-adjusted excess returns of AUM-weighted portfolios of funds in each hierarchy score quintile. There is a negative difference between the risk-adjusted excess returns of the portfolio of funds in the bottom and top quintile of the hierarchy score distribution. Moreover, the excess return of decentralized funds (HIERARCHY = 1) seems to load more on the SMB factor than centralized funds (HIERARCHY = 5). Other than that, the returns of the bottom and top quintiles of the hierarchy score seem to load very similarly on other factors.

1.4 Results

1.4.1 Determinants of information score

My first hypothesis indicates that centralized mutual funds will allocate a greater share of their resources to hard information companies. In order to determine whether centralized funds tend to invest in the higher spectrum of the information score, I calculate the value-weighted average information score of all holdings in a fund's portfolio. This measures the extent to which funds are allocated to information companies in the top of the information score distribution.

Table 1.5 regresses the value weighted average information score on dummies for each hierarchy score quintile and on control variables. I control for (self-reported) investment styles and time heterogeneity by adding investment style and time dummies. Furthermore, net flows can also have an impact in the value-weighted average information score. For instance, if a fund experiences high net flows, it will have to allocate the new funds rather quickly as their investment styles may not allow the fund to hold cash reserves above a certain threshold. As pointed out in Pollet and Wilson (2006) and Lou (2009), funds tend to scale their holdings asymmetrically depending upon whether they experience inflows or redemptions. Therefore, net flows should affect portfolio holdings of existing positions and thus the value weighted average information score. I also control for current quarterly returns and portfolio concentration (average portfolio weight). The first column regresses the average information score on the hierarchy dummies based on the contemporaneous hierarchy score (LAG=0). We can see that as a fund increases its organizational complexity, the information score increases monotonically. For instance, the difference between the average information score of funds with hierarchy score equal to one (decentralized fund) and funds with hierarchy score equal to five (centralized fund) is 0.219 or 3.3% of the sample average. The effect is also statistically significant.

One of the potential problems of this specification is the difficulty to establish causality. Hypothesis 1 indicates that the degree of centralization of a fund (hierarchy score) should have an effect on the type of stocks held in its portfolio (i.e. soft vs hard information stocks). However the tests above do not rule out that the effect is the other way around, i.e. the average information "softness" of stocks held in a mutual fund portfolio has an effect on a fund's organizational complexity. For instance, one could argue that funds decide to concentrate in a sub-set of stocks (i.e. soft information companies) and then decide on a hierarchy structure. This seems unlikely for two reasons. Firstly, the tests are motivated by a theory based on the incentives to collect information provided by different organizational structures. In other words, theory supports the causal relationship between organizational structure and information collection explored here. Secondly, fund families cater to investors' demand by supplying funds with investment styles that are in vogue (Garavito and Wilson (2009)). This catering takes places regardless of the information score of an investment style. However, to partially address these concerns, I regress the contemporaneous average information score on hierarchy dummies based on past hierarchy score. The results are shown in the second (LAG = 6 Months) and third (LAG = 12 months) columns. The results are very similar to those in the first specification.

1.4.1.1 Robustness - New Funds Information Score

An alternative way to explore the robustness of the results above is to look at organizational variables and future information scores of funds that have been launched during the sample. This test clearly identifies the effect we want to study since, prior to the launch of a fund, no investments are made (i.e. no information score) but an organizational structure is put in place (i.e. number of managers running the fund) or inherited (i.e. number of funds in the family). These allows us to better identify the effect of organizational structure in a sub-sample of newly launched mutual funds.

To this end, I regress the information score of newly launched funds twelve months after their IPO's on two organizational variables [8]. These organizational variables are the number of investment managers running the fund (NUM MGRS, fund level variable), and the number of equity funds in the fund's family (LOG FUNDS, family level variable) just before their IPO. Further I add controls which are measured concurrently with the information score (12 months after fund's IPO). I do not include the hierarchy score due to data restrictions on the organizational structure of funds at the time of their IPO. Moreover, the size of newly launched funds is small and hence, the hierarchy score of these funds will tend to be very similar. The results are found in Table 1.6. As we can see both variables, positively covary with future observations of information score. In other words, the more hierarchical a fund is at inception, the more it will tend to allocate capital to stocks in the higher spectrum of the information score. These results corroborate the findings of Table 1.5.

[8] The data comes from Garavito (2008).

1.4.2 Difference in the ability to collect and use soft and hard information

1.4.2.1 Value-weighted trading strategy

Hypothesis 2 states that decentralized (centralized) funds should better at investing in soft (hard) information stocks than centralized (decentralized) funds. In order to test this hypothesis, I construct a self-financing trading strategy that exploits the difference in ability of decentralized and centralized funds to collect and use soft and hard information. In order to test the insights of Stein's model, I define soft information stocks as those that belong to the bottom decile of the information score distribution (i.e. information score distribution = 1). Stocks that are in the top decile of the distribution (information score = 10) are labelled as "hard information stocks". I apply a similar strategy with the organizational complexity of funds. Funds in the bottom quintile of the hierarchy score (hierarchy score = 1) are called decentralized funds and funds in the top quintile of the distribution (hierarchy score = 5) are called centralized funds. From the cross-section of stocks held by mutual funds, I identify the set of soft information stocks held by centralized and decentralized mutual funds respectively. The first self-financing trading strategy takes a long position in a portfolio made of soft information companies held by decentralized funds at each quarter-end date. This purchase is financed by short selling a portfolio composed of soft information stocks held by centralized funds. The portfolio weights of each portfolio trade (long and short portfolios) are proportional to the market capitalization of each stock. The return of this self-financing trading strategy is measured every month and rebalanced every quarter. The results are summarized in Table 1.7. Panel A shows the returns on this trading strategy and on its long and short portfolio trades. The risk-adjusted excess return for the long side of the trade is positive and for the short side is negative with low levels of statistical significance. The return on the strategy is positive and economically significance (49 bps per month or 6% per year). It is important to note that these returns are orthogonal to passive benchmarks (factors) that carry positive premia. I also include a soft-information factor, which tracks the excess return of a value-weighted portfolio of the soft information stocks held by mutual funds (SOFT). This ensures that the trading strategy is neutral with respect to the information score. These results show that there is a positive difference in the ability to collect and use soft information between decentralized and centralized funds.

Similarly, I implement a self-financing trading strategy that aims at exploiting the better provision of incentives to collect hard information that centralized funds offer. As before, the trading strategy consists of buying a value-weighted portfolio of hard information stocks held by centralized funds financed by short-selling a value-weighted portfolio hard information companies held by decentralized funds. I keep track of the monthly returns of the self-financing trading strategy, and I rebalance it every quarter. Panel B shows the results for the second trading strategy. As one can see, the strategy delivers returns that are very small and not statistically different from zero. The failure to find a positive alpha, and hence a positive difference in collecting and using hard information between these two organizational forms, can be due to several reasons. Firstly, hard information can be easily purchased. If outsourcing the production of hard information is cheaper than collecting it, then decentralized and centralized funds are likely to purchase the same information. In this case, there won't be any difference in the ability to generate hard information and, hence, a trading strategy as the one proposed above, would yield zero expected returns. Secondly, since hard information is easy to transmit, it is also difficult to contain. This feature reduces the likelihood that managers earn positive risk-adjusted returns by collecting hard information since, once a piece of information has been harden, it can be transmitted outside the fund (i.e. through employee turnover, communication with research analysts, etc.).

1.4.2.2 Investment-intensity trading strategy

The degree by which a fund loads on a holding should indicate the fund's managers beliefs about the holding's future performance. For example, if a manager strongly believes that IBM stock returns will be high, the manager will overweight this security in its portfolio. Likewise, if a manager is not very confident about the future performance of a long position, it is likely that the manager will underweight it. Therefore, a more direct test of relative performance of centralized and decentralized funds is to incorporate this information in the trading strategy. Consequently, I redefine the portfolio weights of the trading strategies to reflect the aggregate beliefs of managers in each organizational structure. The new weights are based on the average "investment intensity" that each organizational structure assigns to each stock. In the spirit of Cohen Polk and Silli (2010), I define the "investment intensity" of a fund on a holding, as the ratio between the fund's portfolio weight of that holding and the fund's average portfolio weight. A large and positive ratio indicates that the fund is very confident about the positive future performance of the holding. In contrast, a small ratio indicates that the fund's management team is not as bullish on a position. I calculate the aum-weighted average "investment intensity" for all the soft and hard information stocks held by decentralized and centralized funds respectively. I then normalize the average "investment intensity" of all the holdings in each of the trading strategies' long and short portfolios. I used the normalized aum-weighted "investment intensity" as the new portfolio weights. For instance, suppose that there are only two decentralized funds (A and B) with three and four million dollars of AUM respectively. Further, assume that funds A and B both invest in soft information companies C and D with investment intensity on C of 2 and 0.5 and investment intensity on D of 0.2 and 0.4 respectively. Therefore, the long portfolio trade of the soft information strategy would invest 0.78% and 0.22% of each dollar in stocks C and D respectively[9].

[9]The AUM-weighted intensity of investment for C and D are $\frac{3}{7}2 + \frac{4}{7}0.5 = 1.14$ and $\frac{3}{7}0.2 + \frac{4}{7}0.4 = 0.31$ respectively. Therefore, invest $\frac{1.14}{1.14+0.31} = 0.78\%$ in C and $\frac{0.31}{1.14+0.31} =$

Panel A of Table 1.8 presents the return of the trading strategy based on soft information stocks and its long and short side. The risk-adjusted excess return for the long and short side of the trade are positive and negative respectively. Compared to the previous table, the statistical and economical significance of the short side is larger. This shows the inability of centralized funds to deal with soft information. For the long side of the trade, the alpha is positive but the statistical significance is not strong. The risk-adjusted return on the strategy is positive and economically significance (64 bps per month or 7.95% per year). This provides further evidence that there is a positive difference in the ability to collect and use soft information of two organizational forms. Panel B, shows the trading strategy based on hard information companies. The results are very similar to those of Table 1.7. The risk-adjusted return of the trading strategy is very close to zero.

1.4.2.3 Investment-intensity trading strategy: Overlap holdings only

A better test to check whether there is a difference in the ability of using and collecting information of different organizational structures is to run the trading strategy above using holdings that overlap in both sides of the strategy. This will compare how different organizational structures invest in the same (hard or soft information) companies. If there is a difference as measured by the risk-adjusted return of the strategy, it will indicate that one organization structure is better that the other one in collecting and using a particular kind of information (hard or soft information). The results are presented in Table 1.9. Each portfolio trade has the same holdings but their portfolio weights depend on the average intensity of investment as defined in the previous sub-section. Panel A presents the results for the strategy that employs soft information companies only. The results are very similar to those from Table 1.8. We can see that the strategy delivers positive risk-adjusted returns of up to 72 bps per month (9% per year). We observe that the main contributor to this return is the short side of the trade (portfolio of soft information companies

0.22% and in D

held by centralised funds), as the risk-adjusted return of this portfolio trade is -57 bps per month (-7 % per year). The strategy does not load significantly in any of the risk factors and has a low R^2. Panel B shows the results for the strategy based on hard information companies. As before, we see a positive difference between centralized and decentralized funds when investing in hard information companies, albeit, the difference is very small and not statistically significant. This test shows that the results presented in Table 1.8 are not due to a few non-overlapping holdings, or the number of holdings in each portfolio trade.

1.4.2.4 Hard information trading strategies: Ability to purchase hard information and benchmarking needs

As stated before, if outsourcing the production of hard information is cheaper than its collection, then decentralized and centralized funds are likely to subcontract its collection. If this is the case, it is then likely that both types of funds get exposed to the same set of information. Hence, it would be expected to observe that centralized and decentralized funds held a large number of hard information companies in common. Outsourcing the collection of information can therefore diminish the competitive advantage that hierarchical organizations have in collecting hard information. This would explain why the alpha in the trading strategy for hard information companies, while positive, is very close to zero. Another reason may be that hard information companies are held for non-alpha generating reasons. For instance, suppose that funds invest part of their capital to track a benchmark (i.e. S&P 500). If hard information companies tend to be constituents of a widely held index, then it is natural to expect that the trading strategy proposed above for hard information companies, will measure the ability to collect and use hard information *and* benchmark tracking activities of some organizational forms. To address this two points, I modify the portfolio trades in Table 1.8. In order to see whether highly researched companies are driving the results in the hard information trading strategy, I eliminate them from the portfolio trades. I do this by calculating the median of the residual of number of analysts' estimates for each portfolio trade in the hard information strategy of Table 1.8. Next, I eliminate the holdings that are above their respective median from each portfolio trade and run the same strategy as in Table 1.8. The results are summarized in Panel A of Table 1.10. The alpha of the short side of the trade (DECENTRALIZED) is negative. This is in contrasts to previous tables, where the alpha was positive. The alpha of the spread is positive (12 bps per month or 1.5% annual) and statistically significant. This shows that if the strategy concentrates in hard information companies in the lower range of analyst coverage the strategy delivers positive risk-adjusted returns. This supports the idea that the outsourcing of information collection will affect the result of the strategy for hard information companies because it diminishes the advantage that centralized funds have. The other possible refinement to the strategy is by eliminating stocks that may be bought to track a benchmark such as a widely held index. I do this by eliminating S&P 500 constituents from each portfolio trade in the hard information strategy. Panel B of Table 1.10 presents the results. Similar to Panel A, the short side of the trade presents a negative risk-adjusted excess return. The spread of the strategy is positive and larger that in previous tables. In general, the results from this table explain why the hard information strategies from previous tables were not yielding positive alphas. Once refinements to control for outsourcing of information and benchmarking, the risk-adjusted return of the hard information strategy (SPREAD) turns positive with important levels of statistical significance. The evidence in this table suggests that centralized organizations are better at collecting hard information than decentralized funds.

1.4.3 Fund-by-fund trading strategy

One of the objections to the trading strategies implemented above could be that they do not control for investment style. For instance, if decentralized funds tend to be growth funds and growth companies tend to be soft information companies, then one could argue that the results above partially show that growth funds tend to be better at picking growth stocks. To address this concern, I implement a strategy based on holdings at the fund level. For decentralized and centralized funds (hierarchy score equal to one and five respectively), I re-calculate a fund-specific information score by ranking the stocks in a funds portfolio by the aggregate information variable calculated in section 1.3.1.1. I then group these stocks in four groups where the first and fourth groups have the stocks with the lowest and highest aggregate information score respectively (I exclude stocks below the 20th NYSE size percentile). I form a trading strategy for each fund by going long a value-weighted portfolio of companies in the fourth quartile of the fund-specific information score ("hard information" stocks) and short-selling a a value-weighted portfolio of companies in the first quartile of the fund-specific information score ("soft information" companies). Every month, I aggregate the trading strategies by averaging the returns for the universe of decentralized and centralized funds respectively. For every month, I calculate the difference between the average return of the centralized fund trading strategies and the average return of the decentralized fund trading strategies. I rebalance the fund-specific strategies at the end of each quarter. In other to rule out funds that only invest in hard or soft information companies (top and bottom deciles of the cross-sectional information score), I exclude funds that do not hold at least $\alpha\%$ of companies that are below the median of the cross-sectional information score.

This implementation controls for investment style since the trading strategy is based on the holdings of individual funds as opposed to the holdings of all funds of a particular organizational structure (i.e. goes long growth stocks, short-sells growth stocks). Since the trading strategy buys portfolios of "hard" information companies and short-sells portfolios of "soft" information companies, it is sensible to expect that the trading strategy for centralized funds yields positive returns whereas the strategy for decentralized funds delivers negative returns. Therefore the difference between both strategies (centralized minus decentralized) should be, on average, positive. To test this conjecture, I regress this difference on several pricing and benchmarking factors. I include factors that track the excess returns of all soft information companies (INFO1) and hard information companies (INFO10) . This ensures that the ALPHA of the difference is neutral with respect to the information scores. The results in Table 1.11 confirm my conjecture. Each panel in Table 1.11 presents the results for different levels of alpha. We can see that the difference between these strategies ranges from 0.18 % a month (2.18% year) to 0.37% a month (4.53% per year). This provides evidence that, at the fund level, centralized funds do much better on their "harder" information positions while decentralized funds do better on the their "softer" information holdings.

1.4.4 Cross-sectional regressions

Hypothesis 3 conjectures that decentralized funds are able to forecast future returns of soft information companies. In other words, since these funds are able to collect and incorporate soft information in their asset allocation process, their aggreagte actions should predict the average future returns of soft information companies. For instance, if the majority of decentralized funds overweights a particular soft information stock, it should be because the managers of these funds obtained positive (soft information) signals about the future performance of the stock. Therefore, I argue that the average "investment intensity" of decentralized funds should forecast the returns of soft information companies. In other words, high average intensity of investment of decentralised funds should forecast high future expected returns for soft information stocks.

To explore whether ownership of soft and hard information stocks by centralized and decentralized funds predicts the cross-section of stock returns, Table 1.12 presents a series of pooled regressions of returns of stocks held by mutual funds (I exclude stocks below the 20th NYSE size percentile). The sample goes from April 1993 to March 2007. The dependent variable is stock returns measured from the end of month t to the end of month t+3. These returns are regressed on stock characteristics measured at the end of month t. The stock predictors I employ are book-to-market ratio (B/M), the total individual stock return over the previous 12 months (RET12), the monthly trading volume scaled by the number of shares outstanding averaged across the previous three months (TURN), the standard deviation of monthly individual stock returns over the previous 12 months (VOL). I also include the aum-weighted intensity of investment for all hierarchy quintiles and dummies for all information deciles. Because some of the regressors do not have well-behaved distributions, I use their natural logarithm (INTENSITY, B/M, TURN). To control for time heterogeneity, I add time dummies and cluster standard errors by time across firms.

In the first column of Table 1.12 (Model 1), future returns are regressed on four predictors (B/M, RET12, TURN, and VOL). The results show the predictive power of these regressors and confirm previous findings in the literature. Model 2 adds dummies for all deciles of the information score. The average returns of stocks increases as their information score increases. For instance, we see that the average quarterly return of soft information companies (bottom decile of the information score distribution) is lower than that of hard information companies (top decile of information score). This is consistent with the findings in Table 1.2. Model 3 shows how future average returns covary with the aum-weighted investment intensity from different quintiles of the hierarchy score. We observe that, on average, ownership by centralized organizations tends to be negatively correlated with future returns. Model 4 shows a very similar picture when both, investment score dummies and aum-weighted investment intensities are added. Model 5 interacts the dummies for the lowest (SOFT) and highest (HARD) information score deciles with the aum-weighted investment intensities of decentralized (INTENSITY1) and centralized (INTENSITY5) funds. The interaction terms indicate the marginal effect of ownership by decentralized and centralized funds on the future return of soft and hard information companies. Of the four interaction terms, the only one that is economically and statistically significant is the one that measures the marginal effect of decentralized funds' ownership on the average future return of soft information companies (INTENSITY1 x SOFT). For example, a one standard deviation shock to INTENSITY1, leads to a marginal effect on the average future return of soft information stocks of 50 bps per quarter (2% per year). This provides evidence in favour of Hypothesis 3.

1.5 Conclusion

Loosely relying on Stein (2002), I test the effect of organizational structure on the collection and usage of information in a sample of actively managed US equity funds. I also develop scores that measure the information "softness" of stocks and the organizational complexity of mutual funds. I find that the level of organizational complexity of a fund positively covaries with its average information score of its holdings. I also document that there is a difference in ability to use soft information between decentralized and centralized funds. I show that centralized funds seem to also have an advantage in producing hard information, once one takes into account possible benchmarking needs and the outsourcing of information production. Furthermore, decentralized funds also seem able to pick soft information stocks with higher expected returns than their average.

The first set of results indicates that hierarchical funds tilt their portfolios towards hard information stocks. This confirms Stein's insight in that, since the only information that can be transmitted is hard information, hierarchical funds rely more on it and therefore tilt their portfolios to hard-information stocks. This relationship between organization and information helps explain the increase in demand for large stocks (usually hard information companies) documented in Gompers and Metrick (2001) and the consequent reversal of the small stock risk premia over the last 30 years. The surge of institutional investors and the growth of the delegated asset management industry, has given rise to complex hierarchical (centralized) organizations as investment managers. The rise of these hierarchical investment managers may also help explain the replacement of soft information with hard information as the basis for financial transactions documented in Petersen (2004).

One last important point is noteworthy. My results may help explain why concentrated funds tend to outperform diversified ones [10]. As we have seen, decentralized funds tilt their portfolios towards soft information stocks. Since soft information is not transferable, collectors of this information have a longer first mover advantage relative to collectors of more transferable information (i.e. hard information). Therefore, it is optimal for collectors of soft information companies to deviate from holding a diversified portfolio. Instead of holding a diversified portfolio, they choose to learn extensively about fewer stocks in hope of collecting informational rents in the future. From Table 1.3 we can see how the average portfolio weight is negatively correlated with the aum-weighted hierarchy score. Therefore, it is possible that the of concentrated funds outperform diversified ones due to the incorporation of large soft-information bets in their portfolios. The relationship between information production, hierarchical structure and portfolio concentration is therefore and interesting topic for future research in the field of delegated asset management.

[10] Kacperczyk Sialm and Zheng (2005) argue that concentrated managers outperform diversified ones and that the effect is more pronounced amongst managers that hold portfolios concentrated in few industries. Van Nieuwerburgh and Veldkamp (2008) derive conditions under which deviating and holding a concentrated portfolio is an optimal strategy. Bask Busse and Green (2006) discuss mutual fund performance and managers' willingness to take big bets in a relatively small number of stocks. They document that concentrated managers tend to outperform their diversified counterparts.

1.6 References

Baks, Klaas P., Jeffrey A. Busse, and T. Clifton Green, 2006, Fund managers who take big bets: Skilled or overconfident, Working paper, Emory University.

Berger, Allen N., Richard J. Rosen, and Gregory F. Udell, 2007, Does Market Size Structure Affect Competition? The Case of Small Business Lending, *Journal of Banking and Finance*, 31, 11–33.

Berger, Allen N. Gregory F. Udell, 2006, A more complete conceptual framework for SME finance, *Journal of Banking and Finance* 30, 2945–2966.

Campbell, Tim and William Kracaw, 1980, Information Production, Market Signaling and the Theory of Financial Intermediation, *Journal of Finance 35*, 863–881.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Cici, Gjergji , George S. Gibson and , Rabih Moussawi, 2006, For Better or Worse? Mutual Funds in Side-By-Side Management Relationships With Hedge Funds, Working Paper, Mason School of Business.

Cohen, Randolph, Christopher Polk, and Bernhard Silli, 2010, Best Ideas, Working paper, London School of Economics

Daniel, Kent and Sheridan Titman, 2006, Market Reactions to Tangible and Intangible Information, *Journal of Finance*, 61, 4, 1605-1643.

Diamond, Douglas, 1984, Financial Intermediation and Delegated Monitoring, *Review of Economic Studies* 51, 393-414.

Diamond, Douglas, 1991, Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt, *Journal of Political Economy* 99, 689–721.

Fama, Eugene F. and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Garavito, Fabian, 2008, Mutual Fund Incubation: Innovation or Marketing Tool?, Working paper, London School of Economics.

Garavito, Fabian, Mungo Wilson 2009, A Catering Explanation of Mutual Fund IPOs, Working paper, London School of Economics.

Gaspar, Jose-Miguel, Massimo Massa, and Pedro Matos, 2006, Favoritism in mutual fund families? Evidence of strategic cross-fund subsidization, *Journal of Finance* 61, 73–104.

Gompers, P.A., Metrick, A., 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 166, 229–260.

Griffin, M.J., Lemmon, M.L., 2002. Book-to-market equity, distress risk, and stock returns. *Journal of Finance* 57, 2317–2336.

Haubrich, Joseph, 1989, Financial Intermediation, Delegated Monitoring, and Long-Term Relationships, *Journal of Banking and Finance* 13, 9-20.

Hong, H., T. Lim, and J. Stein, 2000, Bad news travels slowly: Size, analyst coverage and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2012.

Leland, H.E., Pyle, D.H., 1977. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance* 32, 371–387.

Lou, Dong, 2009, A Flow-Based Explanation for Return Predictability, Working paper, London School of Economics.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance maximization is not the only game in town, *Journal of Financial Economics* 67, 249–304.

Massa, Massimo, Lei Zhang, 2009, The Effects of the Organizational Structure on Asset Management, Working paper, INSEAD.

Nagel, Stefan, 2005, Short sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309

Nanda, Vikram, JayWang, and Lu Zheng, 2003, Family values and the star phenomenon, Working paper, University of Michigan.

Pastor, Lubos and Robert F. Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy* 111, 642-685

Pollet, Joshua and Mungo Wilson, 2008, How Does Size Affect Mutual Fund Behavior?, *Journal of Finance* 63, 2941-2969.

Petersen, Mitchell. 2004. Information: Hard and Soft, Working Paper, Northwestern University.

Shumway, Tyler, 1997,The Delisting Bias in CRSP Data, *Journal of Finance*, 52, 1, 327-340.

Shumway, Tyler, and Vincent Warther, 1999, The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect, *Journal of Finance*, 54, 6, 2361-2379

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Stein, Jeremy, 2002, Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, *The Journal of Finance* 57, 1891–1921.

Stiglitz, Joseph, and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, *American Economic Review* 71, 393–410.

Van Nieuwerburgh, Stijn and Veldkamp, Laura, 2008, Information Acquisition and Under-Diversification, Working paper, New York University.

Wermers, Russ, 1999, Mutual Fund Herding and the Impact on Stock Prices, *Journal of Finance* 54(2), p.581-622.

Table 1.1: **Summary Statistics: Information Variables**

Panel A reports time-series averages of equal-weighted quarterly cross-sectional means and standard deviations of the information variables and the return predictors. Log B/M is the natural logarithm of the book-to-market ratio; Log AGE is the natural logarithm of the number of months a stock has been reported in CRSP; NUM EST is the number of analyst one-year EPS estimates outstanding for a stock; RET12 (momentum) is the total individual stock return over the previous 12 months; Log SZ is the log of market capitalization; TURN(turnover) is the monthly trading volume scaled by shares outstanding, averaged over the previous three months and divided by two for Nasdaq stocks; VOL (volatility) is the standard deviation of monthly individual stock returns over the previous 12 months; INFO is the information score. Panel B reports time-series averages of cross-sectional correlations. The sample period runs from the first quarter of 1993 to the last quarter of 2006. Only stocks held by mutual funds are included.

	Log AGE	Log BM	Log SZ	NUM EST	OWN	RET12	TURN	VOL	INFO
Panel A: Means standard deviations									
Mean	4.58	0.47	19.43	6.48	0.41	0.39%	0.25	0.14	5.67
Standard Deviation	1.19	0.33	1.85	7.81	0.26	0.04	0.32	0.09	2.86
Observations per quarter (average)	4786	4448	4749	4786	4786	4395	4677	4395	4749
Panel B: Contemporaneous correlations									
Log AGE		0.11	0.22	0.25	0.18	0.07	-0.09	-0.28	0.57
Log BM			-0.37	-0.16	-0.15	-0.41	-0.19	-0.05	0.16
Log SZ				0.74	0.57	0.23	0.23	-0.31	0.25
NUM EST					0.47	-0.01	0.25	-0.17	0.55
OWN						0.12	0.26	-0.20	0.58
RET12							0.12	-0.02	-0.13
TURN								0.31	0.13
VOL									-0.16

Table 1.2: **Information Score Returns - Regressions**

Dependent variable is the value-weigthed excess return of a portfolio made of all stocks with information score (INFO) equal to i at time t. Stocks below the NYSE 20th percentile are excluded. ALPHA is the risk-adjusted return of each portfolio. The three Fama-French factors are zero-investment portfolios representing the excess return of the market, MKTRF; the difference between a portfolio of "small" stocks and "big" stocks, SMB; and the difference between a portfolio of "high" book-to-market stocks and "low" book-to-market stocks, HML. The fourth factor, UMD, is the difference between a portfolio of stocks with high past one-year returns minus a portfolio of stocks with low past one-year returns. The fifth factor, LIQ, is the innovations in the aggregate level of liquidity in Pastor and Stambaugh (2003). The sixth factor, ST REV (Short-term Reversal), is the average return on the two low prior return portfolios minus the average return on the two high prior return portfolios. t-statistics are in parentheses.

	ALPHA	MKTRF	SMB	HML	UMD	LIQ	ST REV	R^2
INFO = 0	-0.26%	1.08	0.78	-0.56	0.20	0.11	0.08	0.89
	-(1.18)	(16.75)	(12.34)	-(6.96)	(4.38)	(3.30)	(1.51)	
INFO = 1	-0.53%	1.09	0.55	-0.49	0.42	0.04	0.16	0.84
	-(2.09)	(15.03)	(7.66)	-(5.41)	(8.12)	(1.14)	(2.50)	
INFO = 2	-0.09%	1.02	0.31	-0.23	0.29	0.02	0.01	0.85
	-(0.48)	(18.99)	(5.88)	-(3.40)	(7.55)	(0.58)	(0.31)	
INFO = 3	0.14%	1.01	-0.04	-0.24	0.00	-0.03	0.07	0.82
	(0.77)	(19.15)	-(0.82)	-(3.65)	-(0.05)	-(1.10)	(1.52)	
INFO = 4	-0.13%	1.02	-0.11	-0.20	0.02	0.00	0.09	0.86
	-(0.87)	(22.76)	-(2.51)	-(3.53)	(0.59)	(0.20)	(2.37)	
INFO = 5	-0.16%	0.91	-0.18	0.07	0.09	0.03	0.04	0.85
	-(1.23)	(25.35)	-(4.95)	(1.45)	(3.58)	(1.36)	(1.26)	
INFO = 6	-0.05%	0.91	-0.01	0.25	0.02	0.01	0.02	0.81
	-(0.33)	(22.51)	-(0.36)	(5.01)	(0.74)	(0.38)	(0.56)	
INFO = 7	-0.07%	1.01	-0.05	0.19	-0.07	-0.01	-0.03	0.88
	-(0.59)	(29.18)	-(1.40)	(4.30)	-(2.82)	-(0.45)	-(1.14)	
INFO = 8	0.17%	1.02	-0.02	-0.15	-0.11	-0.02	-0.02	0.90
	(1.35)	(27.58)	-(0.46)	-(3.31)	-(4.19)	-(1.10)	-(0.52)	
INFO = 9	0.14%	1.05	-0.04	0.21	-0.12	-0.01	-0.04	0.90
	(1.19)	(31.72)	-(1.28)	(5.15)	-(4.92)	-(0.61)	-(1.41)	
INFO 1 - INFO 9	-0.40%	0.03	0.82	-0.77	0.31	0.12	0.13	0.75
	-(1.46)	(0.36)	(10.54)	-(7.80)	(5.60)	(2.93)	(1.81)	

Table 1.3: **Summary Statistics: Hierarchy Variables**

Panel A reports time-series averages of equal-weighted quarterly cross-sectional means and standard deviations of mutual fund variables. Log AUM is the natural logarithm of t e total assets under management; NUM MGRS is the total number of managers in charge of the fund; Log FAM SIZE is the log of the total assets under management of all actively managed US equity funds in a family; NUM FUNDS is the number of actively managed US equity sibling funds; Log FLOW is the log flow of new funds and is defined as the difference between log growth rate for the AUM and the log return; Past Log FLOW is the Log Flow in the previous quarter; W is the average portfolio weight. Panel B reports time-series averages of cross-sectional correlations; HIERARCHY is the hierarchy score; INFO is the information score. The sample period runs from the first quarter of 1993 to the last quarter of 2006. Only actively managed US equity mutual funds are included.

	Log AGE	W	VWA INFO	Log AUM	Log FLOW	Log FAM SIZE	NUM FUNDS	NUM MGRS	HIERARCHY
Panel A: Means standard deviations									
Mean	4.35	0.02	7.32	4.87	0.05	7.37	10.45	2.10	2.97
Standard Deviation	0.97	0.01	0.88	1.88	0.29	2.31	12.45	1.37	1.39
Observations per quarter (average)	986	986	986	975	962	874	876	922	985
Panel B: Contemporaneous correlations									
Log AGE		-0.10	0.10	0.56	-0.24	0.19	0.02	0.00	0.15
W			0.08	-0.32	0.00	-0.31	-0.14	-0.10	-0.28
VWA INFO				0.08	-0.06	0.07	0.05	0.04	0.06
Log AUM					-0.07	0.59	0.21	0.06	0.45
Log FLOW						0.00	-0.01	0.00	0.00
Log FAM SIZE							0.67	0.04	0.76
NUM FUNDS								-0.08	0.40
NUM MGRS									0.49

Table 1.4: **Hierarchy Score Returns - Regressions**

Dependent variable is the quarterly excess return of a value-weigthed portoflio made of funds that belong to hierarchy schore (HIERARCHY) equal to i at time t.All other factores are explained in Table 1.2. t-statistics are in parentheses.

	ALPHA	MKTRF	SMB	HML	UMD	LIQ	ST REV	R^2
HIERARCHY = 1	-0.27%	1.00	0.26	0.00				0.98
	-(1.59)	(41.80)	(8.17)	-(0.13)				
	-0.31%	1.02	0.28	0.01	0.02	-0.01	-0.05	0.99
	-(1.58)	(39.88)	(8.92)	(0.23)	(1.04)	-(0.63)	-(1.73)	
HIERARCHY = 2	-0.14%	0.98	0.16	-0.09				0.98
	-(0.72)	(37.37)	(4.62)	-(2.80)				
	-0.15%	1.00	0.19	-0.08	0.03	-0.01	-0.07	0.99
	-(0.74)	(38.40)	(5.97)	-(2.58)	(1.15)	-(0.65)	-(2.75)	
HIERARCHY = 3	-0.37%	0.98	0.08	-0.06				0.99
	-(2.66)	(50.86)	(2.98)	-(2.50)				
	-0.41%	0.97	0.07	-0.06	0.01	0.01	0.01	0.99
	-(2.44)	(44.02)	(2.65)	-(2.15)	(0.46)	(0.40)	(0.50)	
HIERARCHY = 4	0.13%	1.00	0.10	-0.13				0.98
	(0.59)	(33.65)	(2.47)	-(3.55)				
	0.00%	1.03	0.13	-0.11	0.05	0.01	-0.06	0.98
	(0.01)	(34.25)	(3.42)	-(3.13)	(1.96)	(0.36)	-(1.98)	
HIERARCHY = 5	0.26%	1.00	0.10	-0.04				0.97
	(1.14)	(31.58)	(2.45)	-(1.09)				
	0.10%	1.04	0.15	-0.02	0.06	-0.02	-0.06	0.98
	(0.42)	(32.63)	(3.69)	-(0.42)	(2.02)	-(1.10)	-(1.69)	
H1 - H5	-0.53%	0.00	0.16	0.04				0.32
	-(2.99)	(0.01)	(4.75)	(1.26)				
	-0.41%	-0.02	0.14	0.02	-0.03	0.01	0.01	0.37
	-(1.95)	-(0.79)	(4.00)	(0.69)	-(1.37)	(0.68)	(0.35)	

Table 1.5: **Determinants of Information Score**

Dependent variable is the weighted average information score for each mutual fund portfolio in the sample at a quarter-end date. HIERARCHY i at time t - LAG is a dummy that equals one if the fund belongs to hierarchy score i, and equals zero otherwise. Log AGE is the natural logarithm of the number of months a fund has been reported in CRSP plus one. LOG FLOW is a quarterly measure of net flows defined as $flow_{i,t} = \log\left(\frac{TNA_{i,t}}{TNA_{i,t-1}}\right) - \log(1 + R_{i,t})$. PAST FLOW is the FLOW of the past quarter. RETURN is a fund's quarterly excess return. CONCENTRATION is the average portfolio weight of a fund. Robust t-statistics are in brackets. Standard errors are clustered at the fund level.

	LAG = 0	LAG = 6	LAG =12
INTERCEPT	7.030***	7.033***	7.010***
	[55.755]	[55.575]	[54.793]
HIERARCHY 2	0.098**	0.096**	0.096**
	[2.361]	[2.307]	[2.245]
HIERARCHY 3	0.125***	0.124***	0.137***
	[2.693]	[2.652]	[2.861]
HIERARCHY 4	0.160***	0.168***	0.177***
	[3.303]	[3.432]	[3.539]
HIERARCHY 5	0.219***	0.219***	0.237***
	[4.301]	[4.284]	[4.555]
AGE	0.126***	0.127***	0.127***
	[6.730]	[6.772]	[6.726]
LOG FLOW	-0.173***	-0.171***	-0.166***
	[-5.549]	[-5.519]	[-5.281]
PAST LOG FLOW	-0.118***	-0.114***	-0.114***
	[-4.303]	[-4.171]	[-4.166]
RETURN	-0.612***	-0.601***	-0.601***
	[-9.089]	[-8.978]	[-8.961]
CONCENTRATION	10.004***	10.028***	10.147***
	[4.898]	[4.892]	[4.906]
Investment Style Dummies	Yes	Yes	Yes
Year.Qtr Dummies	Yes	Yes	Yes
Observations	37733	37687	37597
R-squared	0.166	0.165	0.165

Table 1.6: **Determinants of Information Score - New Funds**

Dependent variable is the weighted average information score for each U.S. actively managed mutual fund portfolio that IPOs in my sample 12 months after they are launched. LOG FUNDS and NUM MGRS are defined as in Table 1.3. The other variables are defined as in Table 1.5. Robust t-statistics are in brackets.

	MODEL 1	MODEL 2
LOG FUNDS	0.069**	
	[2.104]	
NUM MGRS		0.032*
		[1.937]
LOG AGE	-0.128**	-0.138**
	[-1.965]	[-2.188]
LOG FLOW	0.006	-0.009
	[0.051]	[-0.074]
PAST LOG FLOW	-0.203*	-0.172*
	[-1.855]	[-1.764]
RETURN	-0.483**	-0.568**
	[-1.993]	[-2.567]
CONCENTRATION	-0.926	-1.466
	[-0.324]	[-0.553]
INTERCEPT	7.285***	7.415***
	[21.504]	[23.375]
Year Dummies	Yes	Yes
Investment Style Dummies	Yes	Yes
Observations	718	801
R-squared	0.294	0.294

Table 1.7: **Trading strategy: Value-Weighted Monthly Regressions**

Dependent variable is the monthly return of a portfolio trade. These four portfolio trades are the soft and hard information stocks held by the centralised and decentralised mutual funds in my sample in every quarter-end date. These portfolio trades are rebalanced every quarter. Panel A (B)shows regression coefficients for portfolios of soft (hard)information companies held by decentralized and centralized companies and their difference. These portfolios are value-weighted. The other variables (including SOFT (INFO = 1) and HARD (INFO = 10)) are defined as in Table 1.2. Robust t-statistics are in brackets.

Panel A: Soft Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (DEC-CEN)		
ALPHA	0.0036	0.0033	-0.0010	-0.0016	0.0047***	0.0050***	0.0049***
	[1.15]	[1.05]	[-0.33]	[-0.53]	[2.80]	[2.97]	[2.96]
MKTRF	1.25***	1.18***	1.21***	1.14***	0.04	0.05	0.07
	[15.42]	[15.21]	[15.33]	[14.33]	[1.00]	[1.15]	[0.75]
SMB	0.71***	0.69***	0.73***	0.71***	-0.02	-0.02	-0.01
	[4.70]	[4.59]	[5.41]	[5.32]	[-0.34]	[-0.35]	[-0.06]
HML	-0.78***	-0.82***	-0.73***	-0.77***	-0.05	-0.05	-0.06
	[-5.97]	[-6.21]	[-5.85]	[-6.00]	[-0.82]	[-0.86]	[-0.88]
UMD	0.33***	0.35***	0.26***	0.30***	0.06**	0.05*	0.06*
	[3.82]	[4.32]	[3.47]	[4.10]	[2.19]	[1.70]	[1.79]
LIQ		0.13**		0.12**		0.00	0.00
		[2.07]		[2.39]		[0.06]	[0.11]
ST REV		0.03		0.08		-0.05*	-0.05
		[0.29]		[0.80]		[-1.70]	[-1.65]
SOFT							-0.02
							[-0.22]
Observations	168	168	168	168	168	168	168
R-squared	0.84	0.85	0.85	0.86	0.04	0.05	0.05

Panel B: Hard Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (CEN-DEC)		
ALPHA	0.0013	0.0016	0.0012	0.0014	-0.0002	-0.0002	-0.0001
	[1.09]	[1.21]	[0.97]	[1.10]	[-1.65]	[-1.56]	[-1.47]
MKTRF	1.04***	1.05***	1.04***	1.05***	0.00	0.00	0.01
	[31.59]	[28.16]	[32.03]	[28.45]	[0.11]	[0.02]	[1.62]
SMB	-0.05	-0.05	-0.04	-0.04	0.01***	0.01***	0.01***
	[-1.42]	[-1.41]	[-1.19]	[-1.18]	[5.50]	[5.34]	[5.21]
HML	0.20***	0.20***	0.21***	0.21***	0.01***	0.01***	0.01***
	[4.38]	[4.40]	[4.70]	[4.71]	[4.33]	[4.06]	[4.43]
UMD	-0.10***	-0.11***	-0.10***	-0.11***	-0.00	-0.00	-0.00
	[-3.12]	[-3.52]	[-3.17]	[-3.58]	[-0.39]	[-0.36]	[-0.81]
LIQ		-0.01		-0.01		0.00	0.00
		[-0.40]		[-0.38]		[0.60]	[0.55]
ST REV		-0.04		-0.04		-0.00	-0.00
		[-0.85]		[-0.87]		[-0.32]	[-0.55]
HARD							-0.01*
							[-1.97]
Observations	168	168	168	168	168	168	168
R-squared	0.90	0.90	0.90	0.90	0.14	0.15	0.16

Table 1.8: **Trading strategy: Abnormal Portfolio Tilts Monthly Regressions**

Dependent variables are as in Table 1.7. However, the portfolio weights of each portfolio trade are calculated differently. For each fund in my sample, I calculate the investment tilt for each of its holdings. This tilt is defined as the fund's portfolio weight on a holding divided by the average portfolio weight of the fund. Each quarter-end date, I identify four portfolio trades as in Table 1.7. I calculate the asset-weigthed average portfolio tilt of each stock in each one of these portfolio trades. For each portfolio trade, I normalize the asset-weigthed average portfolio tilts so they sum up to one. I use these normalized tilts as portfolio weights in each of the four portfolio trades. All variables other variables are defined as in Table 1.7. Robust t-statistics are in brackets.

Panel A: Soft Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (DEC-CEN)		
ALPHA	0.0024	0.0022	-0.0042	-0.0044*	0.0066***	0.0067***	0.0064***
	[1.00]	[0.91]	[-1.65]	[-1.70]	[3.00]	[2.97]	[2.83]
MKTRF	1.23***	1.20***	1.31***	1.25***	-0.09	-0.05	0.08
	[19.77]	[20.13]	[20.38]	[18.41]	[-1.27]	[-0.68]	[0.64]
SMB	0.87***	0.86***	0.95***	0.93***	-0.08	-0.07	0.03
	[7.83]	[7.75]	[11.13]	[11.73]	[-1.01]	[-0.83]	[0.19]
HML	-0.38***	-0.40***	-0.45***	-0.49***	0.07	0.09	0.03
	[-4.06]	[-4.10]	[-4.47]	[-5.09]	[0.73]	[0.96]	[0.29]
UMD	0.18***	0.19***	0.26***	0.28***	-0.08	-0.09*	-0.07
	[3.08]	[3.27]	[4.67]	[5.40]	[-1.63]	[-1.83]	[-1.47]
LIQ		0.06		0.14***		-0.08*	-0.07*
		[1.28]		[3.10]		[-1.93]	[-1.68]
ST REV		0.01		0.01		-0.00	0.01
		[0.19]		[0.18]		[-0.04]	[0.11]
SOFT							-0.12
							[-1.15]
Observations	168	168	168	168	168	168	168
R-squared	0.88	0.88	0.89	0.90	0.09	0.12	0.13

Panel B: Hard Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (CEN-DEC)		
ALPHA	0.0001	0.0003	0.0003	0.0007	0.0003	0.0004	0.0004
	[0.05]	[0.21]	[0.27]	[0.57]	[0.60]	[0.98]	[0.87]
MKTRF	1.14***	1.15***	1.13***	1.14***	-0.01	-0.01	-0.05
	[28.48]	[27.46]	[30.08]	[30.53]	[-0.88]	[-0.67]	[-1.11]
SMB	0.35***	0.35***	0.34***	0.34***	-0.01	-0.01	-0.01
	[7.28]	[6.97]	[6.74]	[6.64]	[-0.27]	[-0.37]	[-0.27]
HML	0.58***	0.58***	0.51***	0.51***	-0.07***	-0.07***	-0.08***
	[10.68]	[10.39]	[9.35]	[8.93]	[-3.90]	[-4.00]	[-4.04]
UMD	-0.13***	-0.13***	-0.12***	-0.14***	0.00	-0.00	0.00
	[-3.57]	[-3.67]	[-3.39]	[-3.67]	[0.25]	[-0.20]	[0.24]
LIQ		-0.01		-0.01		0.01	0.01
		[-0.48]		[-0.28]		[0.86]	[0.89]
ST REV		-0.03		-0.06		-0.03	-0.03
		[-0.63]		[-1.19]		[-1.64]	[-1.60]
HARD							0.04
							[1.08]
Observations	168	168	168	168	168	168	168
R-squared	0.88	0.88	0.89	0.89	0.15	0.20	0.21

Table 1.9: **Trading strategy: Overlapping holdings Regressions**

Dependent variables are as in Table 1.7. Portfolio holdings of each portfolio trade are as in Table 1.7. However, I exclude non-overlapping stocks in each Panel. Portfolio weights are based on intensity of investment as in Table 1.8. All variables other variables are defined as in Table 1.7. Robust t-statistics are in brackets.

Panel A: Soft Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (DEC-CEN)		
ALPHA	0.0015	0.0015	-0.0057*	-0.0057*	0.0072**	0.0072**	0.0065**
	[0.58]	[0.55]	[-1.86]	[-1.81]	[2.34]	[2.26]	[2.08]
MKTRF	1.26***	1.24***	1.34***	1.27***	-0.08	-0.03	0.26
	[18.72]	[19.06]	[18.73]	[16.91]	[-1.05]	[-0.34]	[1.48]
SMB	0.87***	0.86***	0.97***	0.95***	-0.10	-0.09	0.12
	[7.84]	[7.77]	[11.01]	[11.84]	[-1.04]	[-0.83]	[0.59]
HML	-0.28***	-0.29***	-0.37***	-0.43***	0.10	0.14	-0.01
	[-2.74]	[-2.81]	[-3.19]	[-3.88]	[0.83]	[1.11]	[-0.10]
UMD	0.19***	0.20***	0.28***	0.31***	-0.09	-0.11*	-0.06
	[3.26]	[3.26]	[4.47]	[5.13]	[-1.55]	[-1.74]	[-0.94]
LIQ		0.05		0.17***		-0.12**	-0.09*
		[0.83]		[3.22]		[-2.19]	[-1.70]
ST REV		-0.01		-0.02		0.02	0.04
		[-0.09]		[-0.26]		[0.20]	[0.46]
SOFT							-0.27*
							[-1.72]
Observations	168	168	168	168	168	168	168
R-squared	0.86	0.86	0.84	0.85	0.07	0.11	0.14

Panel B: Hard Info Stocks	DECENTRALIZED		CENTRALIZED		SPREAD (CEN-DEC)		
ALPHA	0.0001	0.0003	0.0003	0.0008	0.0002	0.0004	0.0004
	[0.07]	[0.25]	[0.26]	[0.58]	[0.52]	[0.90]	[0.78]
MKTRF	1.14***	1.15***	1.12***	1.14***	-0.01	-0.01	-0.06
	[28.61]	[27.52]	[30.34]	[30.70]	[-1.02]	[-0.73]	[-1.17]
SMB	0.35***	0.35***	0.34***	0.34***	-0.01	-0.01	-0.01
	[7.21]	[6.90]	[6.64]	[6.55]	[-0.34]	[-0.43]	[-0.33]
HML	0.58***	0.58***	0.50***	0.50***	-0.08***	-0.08***	-0.09***
	[10.70]	[10.42]	[9.26]	[8.85]	[-4.12]	[-4.20]	[-4.26]
UMD	-0.12***	-0.13***	-0.12***	-0.13***	0.01	0.00	0.01
	[-3.41]	[-3.56]	[-3.19]	[-3.52]	[0.49]	[0.07]	[0.51]
LIQ		-0.01		-0.01		0.00	0.01
		[-0.45]		[-0.30]		[0.62]	[0.66]
ST REV		-0.03		-0.07		-0.03	-0.03
		[-0.67]		[-1.26]		[-1.63]	[-1.59]
HARD							0.04
							[1.13]
Observations	168	168	168	168	168	168	168
R-squared	0.88	0.88	0.89	0.89	0.16	0.21	0.22

Table 1.10: **Trading strategy: Hard Information Companies**

Dependent variables are as in Table 1.7. Panel A calculates the residual number of analyst estimates median of each portfolio trade. I remove holdings with residual number of analyst estimates *above* their median. Portfolio weights are based on intensity of investment of each holding in each portfolio trade as in Table 1.8. Panel eliminates S&P 500 constituents from hard information portfolios trades from Table 1.8. Portfolio weights are based on intensity of investment as in Table 1.8. Dependent variables are as in Table 1.7. All variables other variables are defined as in Table 1.7. Robust t-statistics are in brackets.

Panel A: Number of Estimates	DECENTRALIZED		CENTRALIZED		SPREAD (CEN-DEC)		
ALPHA	-0.0009	-0.0006	0.0001	0.0006	0.0010*	0.0012*	0.0011*
	[-0.67]	[-0.45]	[0.06]	[0.43]	[1.73]	[1.94]	[1.89]
MKTRF	1.11***	1.12***	0.39***	1.08***	-0.04**	-0.03**	-0.09
	[31.68]	[31.36]	[8.12]	[31.69]	[-2.23]	[-2.09]	[-1.57]
SMB	0.39***	0.38***	0.57***	0.38***	0.00	-0.00	0.00
	[9.26]	[9.12]	[11.30]	[8.44]	[0.04]	[-0.03]	[0.07]
HML	0.67***	0.67***	-0.07	0.56***	-0.10***	-0.10***	-0.12***
	[14.03]	[13.28]	[-1.65]	[10.53]	[-4.56]	[-4.64]	[-4.75]
UMD	-0.06	-0.07**		-0.08**	-0.01	-0.01	-0.01
	[-1.61]	[-1.97]		[-2.19]	[-0.55]	[-0.98]	[-0.53]
LIQ		0.01		0.02		0.01	0.01
		[0.31]		[0.60]		[0.99]	[1.07]
ST REV		-0.05		-0.09*		-0.03*	-0.03*
		[-1.30]		[-1.94]		[-1.71]	[-1.73]
SOFT							0.05
							[1.12]
Observations	168	168	168	168	168	168	168
R-squared	0.88	0.88	0.87	0.88	0.16	0.19	0.20

Panel B: No S&P 500 Constituents	DECENTRALIZED		CENTRALIZED		SPREAD (CEN-DEC)		
ALPHA	-0.0007	-0.0005	0.0000	0.0006	0.0008	0.0010	0.0010
	[-0.51]	[-0.33]	[0.03]	[0.43]	[1.05]	[1.37]	[1.33]
MKTRF	1.13***	1.15***	1.12***	1.14***	-0.01	-0.01	-0.05
	[26.80]	[25.59]	[29.45]	[29.94]	[-0.73]	[-0.51]	[-0.63]
SMB	0.33***	0.34***	0.32***	0.32***	-0.02	-0.02	-0.02
	[6.33]	[6.06]	[5.81]	[5.83]	[-0.56]	[-0.66]	[-0.60]
HML	0.61***	0.62***	0.49***	0.49***	-0.13***	-0.13***	-0.14***
	[11.00]	[10.92]	[8.74]	[8.35]	[-4.45]	[-4.37]	[-4.15]
UMD	-0.04	-0.05	-0.06	-0.07*	-0.02	-0.03	-0.02
	[-1.05]	[-1.31]	[-1.40]	[-1.89]	[-0.96]	[-1.49]	[-1.21]
LIQ		-0.01		-0.01		0.01	0.01
		[-0.42]		[-0.24]		[0.51]	[0.53]
ST REV		-0.04		-0.09		-0.05	-0.04
		[-0.82]		[-1.59]		[-1.57]	[-1.56]
HARD							0.03
							[0.56]
Observations	168	168	168	168	168	168	168
R-squared	0.86	0.86	0.87	0.88	0.17	0.20	0.20

Table 1.11: **Fund-by-fund Self-financing trading strategy**

Each quarter I identify funds that belong to the top and bottom quintile of the hierarchy score (centralized and decentralized funds) that hold at least $\alpha\%$ of their assets in stocks that are below the information score median. For each fund, the information score is recalculated using the fund's holdings only. The self financing trading strategy is constructed by going long hard information companies (information score = 10) and short soft information companies (information score = 1) for each fund at the end of each quarter. I aggregate the return for the decentralized and centralized fund strategies respectively. The dependent variable in the regressions is the difference between the average return for the centralized fund trading strategies and the the average return for the centralized fund trading strategies. The strategies are held for a quarter and rebalanced at the end of each quarter. The other variables are defined as in Table 1.2. t-statistics are in parentheses.

	ALPHA	MKTRF	SMB	HML	UMD	LIQ	ST REV	INFO1	INFO9	R^2
Panel A: $\alpha = 5\%$	0.18%	0.02	-0.01	-0.01						0.01
	(2.12)	(0.93)	-(0.31)	-(0.16)						
	0.18%	0.05	0.02	-0.02	0.02	-0.02	-0.04	-0.03	0.04	0.08
	(2.06)	(0.58)	(0.53)	-(0.58)	(1.00)	-(1.67)	-(2.09)	-(1.06)	(0.60)	
Panel B: $\alpha = 10\%$	0.23%	0.02	-0.02	-0.01						0.01
	(2.56)	(1.01)	-(0.67)	-(0.26)						
	0.22%	0.05	0.02	-0.04	0.03	-0.02	-0.05	-0.05	0.06	0.11
	(2.45)	(0.56)	(0.63)	-(1.02)	(1.32)	-(1.61)	-(2.20)	-(1.55)	(0.95)	
Panel C: $\alpha = 15\%$	0.22%	0.03	-0.01	-0.01						0.02
	(2.15)	(1.27)	-(0.51)	-(0.25)						
	0.20%	0.12	0.03	-0.04	0.03	-0.02	-0.05	-0.06	0.02	0.10
	(2.00)	(1.19)	(0.81)	-(0.83)	(1.26)	-(1.57)	-(1.81)	-(1.71)	(0.27)	
Panel D: $\alpha = 20\%$	0.27%	0.05	-0.04	0.03						0.03
	(2.27)	(1.49)	-(1.23)	(0.65)						
	0.24%	0.13	0.01	0.00	0.04	-0.04	-0.04	-0.07	0.04	0.11
	(2.01)	(1.14)	(0.27)	-(0.01)	(1.48)	-(1.98)	-(1.26)	-(1.65)	(0.42)	
Panel E: $\alpha = 25\%$	0.35%	0.06	-0.06	0.00						0.03
	(2.43)	(1.44)	-(1.56)	(0.05)						
	0.29%	0.18	0.00	-0.03	0.07	-0.04	-0.05	-0.09	0.03	0.13
	(2.06)	(1.30)	-(0.06)	-(0.44)	(2.14)	-(1.95)	-(1.30)	-(1.68)	(0.29)	
Panel F: $\alpha = 30\%$	0.44%	0.05	-0.07	-0.05						0.03
	(2.71)	(1.06)	-(1.42)	-(0.82)						
	0.37%	0.06	-0.06	-0.06	0.09	-0.04	-0.05	-0.03	0.08	0.11
	(2.23)	(0.38)	-(0.87)	-(0.81)	(2.51)	-(1.51)	-(1.16)	-(0.51)	(0.68)	

Table 1.12: **Cross-sectional Regressions**

The Dependent variable is the stock return from month t to month t+3. B/M (book-to-market ratio), TURN (turnover), RET12 (momentum), and VOL (volatility) are defined as in Table 1.1. SOFT is a dummy that equals one if stock has the lowest information score and zero otherwise. HARD is a dummy that equals one if the stock has the highest information score. INTENSITY i is the aggregate investment intensity of funds belonging to quintile i of the hierarchy score on a stock. Aggregate investment intensity is defined as in Table 7. t-statistics are in brackets and errors are clustered by time across firms.

	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5
B/M	0.00874*	0.00817*	0.00841*	0.00724*	0.00719
	[1.861]	[1.755]	[1.868]	[1.674]	[1.655]
TURN	-0.00659	-0.00973*	-0.00612	-0.00867*	-0.00873*
	[-1.406]	[-1.932]	[-1.302]	[-1.724]	[-1.736]
RET12	0.09969	0.10426*	0.09864	0.10660*	0.10646*
	[1.632]	[1.716]	[1.577]	[1.705]	[1.702]
VOL	-0.08939**	-0.08380**	-0.09012**	-0.08619**	-0.08628**
	[-2.244]	[-2.140]	[-2.249]	[-2.178]	[-2.180]
INTENSITY1			0.00119	0.00067	0.00032
			[0.745]	[0.431]	[0.205]
INTENSITY2			0.00180	0.00126	0.00127
			[1.174]	[0.850]	[0.859]
INTENSITY3			0.00045	-0.00001	-0.00001
			[0.365]	[-0.010]	[-0.009]
INTENSITY4			-0.00128	-0.00185	-0.00183
			[-0.762]	[-1.055]	[-1.042]
INTENSITY5			-0.00260*	-0.00327**	-0.00351**
			[-1.869]	[-2.413]	[-2.584]
SOFT		0.00223		-0.01316***	-0.01495***
		[0.484]		[-2.998]	[-3.028]
HARD		0.00850*		-0.00580	-0.00646
		[1.917]		[-1.373]	[-1.597]
INTENSITY1 X SOFT					0.00787*
					[1.823]
INTENSITY1 X HARD					0.00025
					[0.112]
INTENSITY5 X SOFT					0.00113
					[0.223]
INTENSITY5 X HARD					0.00160
					[0.962]
Time Dummies	Yes	Yes	Yes	Yes	Yes
Info Score Dummies	No	Yes	No	Yes	Yes
Observations	170374	170374	170374	170374	170374
R-squared	0.166	0.166	0.166	0.167	0.167

Chapter 2

Mutual Fund Incubation: Innovation or Marketing Tool?

2.1 Introduction

Fund incubation (the creation and management of a mutual fund before it is offered to the public) is a very controversial topic in the money management industry. Those opposed to this practice argue that funds could be incubated with the sole purpose of generating outstanding track records for marketing purposes[1]. This claim seems to be supported by the high returns of incubated funds during their incubation period and their subsequent decline. On the other hand, it has been argued that fund incubation helps test relatively unknown investment strategies and fund managers. The challenges posed by fund incubation have been recognized by regulators, practitioners, and academics. My main objective is to study how the relaxation of certain financial constraints during the incubation period (i.e. capital redemptions) affects funds' investment strategies before and after their IPO's. Furthermore, I investigate whether mutual fund families are strategic when launching incubated funds and whether incubated funds attract a larger share of the new-fund capital flows than non-incubated funds.

[1] Sirri and Tufano (1998), Chevalier and Ellison (1997) document a convex relationship between past performance and current capital flows.

Mutual fund incubation is a tool that fund families employ to develop new fund offerings. Fund families start incubated funds with limited amounts of internally-raised capital. Fund families then have the option to offer these to the public or to terminate them at a later date[2]. If they choose to have an IPO for a mutual fund, the family can use the fund's past performance for marketing purposes. In the U.S., this practice is governed by a series of No-Action letters issued by the S.E.C. (See Evans (2007) for more details). The issue with fund incubation is that, if investors use past performance as a proxy for managerial skill, fund families may choose to launch their incubated funds when their past returns are unusually high. This will some lead investors to wrongly conclude that the funds being offered have above-average investment skills. However, the open empirical questions is whether mutual fund investors indeed chase pre-IPO performance of new fund offerings or whether there are other more important fund flow determinants for new funds. Additionally, it remains unclear whether the incubation period has a long lasting effect in the investment strategies followed by incubated funds or whether these funds are no different than other new fund offerings.

In this paper, I document that families are strategic when choosing IPO dates for their incubated funds. Incubated funds are offered to the public when their 12-month return is above the median return of the mutual fund industry. The 12-month return of incubated funds tends to stay above the industry median for about 15 months and reverts back to the industry mean after that. This is consistent with an opportunistic behaivor that tries to exploit the positive relationship between past performance and capital flows.

One of the findings in the literature is that the performance of incubated funds decreases after their IPO, reinforcing the idea that incubation is used as a marketing tool. However, the decrease in post-IPO performance of incubated funds is also consistent with the imposition of financial constraints

[2]Evans (2006) shows that incubated funds with high returns are more likely to be offered to the public whereas incubated funds with low returns are more likely to be liquidated. Evans(2007) documents that the inclusion of incubation returns in the CRSP Mutual Fund Database biases many of the results in the mutual fund literature, since these returns are unusually high.

associated with capital redemption rights. While during the incubation period funds have the commitment of seed funds from their sponsors, after their IPO's funds allow investors to withdraw their shares at the funds' current Net Asset Value (NAV) at any point in time. This feature imposes constraints on the type of investment strategies that can be implemented. For example, fund managers would be wary of high-return investments that required time to "mature" but that were very volatile in the interim. The reason is that investors may be unable to distinguish between bad interim returns of a high-return long-term investment strategy and bad performance due to managerial ability. This will induce some investors to redeem their capital forcing the fund manager to fire-sell assets in order to meet redemptions. These fire-sales would worsen the fund's performance, increase fund's redemptions and so on. This is exacerbated by the payoff complementarities amongst mutual fund investors (Chen Goldstein and Jiang (2007)). Prior studies provide evidence showing that investors who redeem their shares from a fund have a negative externality on investors who do not (see Eden(1999), Johnson (2004)). As a result, the expectation of future outflows increases the incentives of current investors to redeem their shares. Therefore, the insulation from outside investors' capital and the commitment of family funds in the incubation period eliminates the payoff complementarities and relaxes financial constraints. This allows funds to explore different "corners" of the market and to implement high-return investment strategies that may need time to materialize.

Consistent with these conjectures, I find that incubated funds hold more illiquid stocks with a lower degree of institutional ownership during their incubation period. I also find that the risk-adjusted returns of incubated finds decline after their IPO's. Furthermore, the characteristic-adjusted returns [3] of a strategy that buys and holds the portfolio disclosures of incubated funds, yields 0.39% per month more during the incubation period. However, when I analyze funds' trades of funds, which are active portfolio management deci-

[3]The characteristic adjusted returns are calcualted as in Daniel, Grinblatt, Titman, and Wermers (1997) and Wermers (2004). The DGTW benchmarks are available via http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm

sions and hence a better indication of managerial ability, I find that incubated funds are better at buying stocks in the incubation period and better at selling stocks after their IPO's. For instance, the difference in characteristic-adjusted returns between buy-and-hold strategies based on the purchases of incubated funds before and after the IPO is on average 62 bps (0.62%) per month. Additionally, the difference in characteristic-adjusted returns between buy-and-hold strategies that purchase stock disposals of incubated funds before and after the IPO is 54 bps per month (0.54%). This evidence shows that, while incubated fund managers are better at buying stocks during the incubation period, they are better at selling stocks after the IPO.

Despite the change in investment strategies of incubated funds before and after the IPO, I find that incubated funds follow strategies that are different from those of other new fund offerings. I document that, after their IPO's, incubated funds tend to hold more illiquid and concentrated portfolios than non-incubated funds. Van Nieuwerburgh and Veldkamp (2008) show that it is optimal for investors that first collect information systematically on a set of assets to hold concentrated portfolios. Therefore the difference in portfolio concentration between incubated and non-incubated funds can be attributed to the time spent learning about a set of stocks during the incubation period. I also show that incubated funds tend to hold stocks with lower institutional ownership. Since institutional investors (i.e. mutual funds) tend to invest in a very narrow set of stocks (Gompers and Metrick (2001)), fund incubation could be a way for fund families to test new investment strategies in lesser known securities. This evidence suggests that the incubation period helps funds explore a particular set of stocks. This "learning" process would not be feasible once the funds . Incubated funds can therefore be an important source of investment vehicles that span a different set of asset payoffs and that provide real diversification benefits to mutual fund investors. Moreover, I find that, while post-IPO risk-adjusted returns of incubated and non-incubated funds are very similar, incubated funds seem to be better at buying stock that non-incubated stocks. The difference in characteristic-adjusted returns between buy-and-hold strategies based on the post-IPO purchases of incubated funds and non-incubated funds respectively is 31 bps (0.31%) per month. This difference in managerial skill may be due to the fact that incubated funds can use their incubation period as a learning phase about a particular corner of the market and benefit from it long after their IPO's.

Furthermore, I study whether the high past returns of incubated mutual funds help them attract a larger share of fund flows than that of non-incubated funds. I document that, in the first 24 months after the IPO, the difference in relative fund flows between incubated and non-incubated funds is around -1%, after controlling for past (post-IPO) performance and other factors. This difference is mainly driven by the first post-IPO year of the new funds in my sample, where the difference is -2.5 %. These results illustrate that investors seem not be drawn to new funds offerings due to their pre-IPO track record.

The practice of fund incubation and the issues associated with it have been recognized in the academic literature. Most of the literature focuses on the biases induced by fund incubation. Elton, Gruber, and Blake (2001) suggest that track records of successful mutual funds may be backfilled with pre-IPO performance. The same issue is recognized in hedge fund databases by Park (1995) and Fung and Hsieh (2002). Elton, Gruber, and Rentzler (1989) suggest that the post-IPO underperformance of commodity funds could be due to the control that managers have over the reporting of pre-IPO performance. Evans (2007), quantifies the magnitude of the incubation bias in the CRSP Survivor-Bias Free U.S. Mutual Fund Database and proposes a methodology to account for it. Arteaga, Ciccotello, and Grant (1998) and Wisen (2002) examine the bias in returns for new fund offerings and its implications. Ackermann and Loughran (2007) examine the practice of mutual fund incubation in the U.S., the role of the S.E.C. and its potential unintended consequences[4].

Constraints imposed by redemption rights of investment vehicles and co-ordination problems amongst fund investors have also been studied in the

[4]For a legal treatment of fund incubation in the US see Ackermann and Loughran (2007), and Palmiter and Taha (2009), Franco (2009).

literature. The initial public offering of a fund allows us to precisely identify periods where funds operate with and without financial constraints and coordination problems amongst fund investors. This, in turn, helps us accurately determine the effects that these constrains have on the way funds operate. This relates to the literature on limits of arbitrage. In particular, it has been shown that limiting fund investors' redemption rights can have positive effect on fund returns and may affect the portfolio characteristics of funds. Stein (2005) illustrates how open-end investment vehicles may have serious impediments to exploit arbitrage opportunities. Edelen (1999) shows that the liquidity constraints imposed by redemptions rights in mutual funds can be quite substantial. Choria (1996) shows that the adverse effects of investor flows are more pronounced in funds holding more illiquid assets. Ippolito (1989) illustrates how the market-adjusted returns of load-type funds are higher than those of no-load funds. Ackermann, McEnally and Ravenscraft (1998) show a positive relation between hedge fund returns an lockup provisions. Aragon (2007) shows that there is a negative relation between the liquidity of hedge fund portfolios and their lockup periods. This paper also relates to another strand of the literature that has focused on the negative externalities that fund investors redeeming their capital have on the other fund's investors. Johnson (2004) shows that the liquidity costs imposed by short-term mutual fund investors are larger than that of long-term investors. Chen Goldstein and Jian (2007) present a model in which the payoff complementarities of investors in a mutual fund can cause investment fragility and provide evidence that conditional on low past performance, redemptions are higher for funds for which payoff complementarities are stronger.

The paper is structured as follows: Section 2.2 describes the data , section 2.3 presents the results and section 2.4 concludes.

2.2 Data

I obtain mutual fund data from the CRSP Survivor-Bias Free U.S. Mutual Fund Database. To ensure that the sample consists of actively managed U.S. equity funds, I exclude index funds and funds that on average do not hold at least 85% of CRSP common stocks. The percentage of funds invested in NYSE-AMEX-NASDAQ CRSP common stock is calculated using portfolio holdings obtained from the Thompson Financial Mutual Fund Database. I merge the holdings data and the mutual fund data using the MFLINK tables provided by WRDS. This filters out some non-US actively managed equity funds. Additionally, I screen out index funds by looking for the words associated with index funds and their abbreviations in the CRSP fund name variable. I also check that the MFLINK matches between the unique identifiers in the CRSP Mutual Fund Dataset and the unique identifiers in the Thomson Financial Mutual Fund Holdings database correspond to funds managed by the same management company. This eliminates erroneous MFLINK matches. As an additional check on the accuracy of the MFLINK matches, I eliminate funds for which the TNA reported in the TFN database is not between 1/1.3 and 1.3 of the TNA reported in the CRSP database.

To identify when a fund is officially offered to the public, I use NASD ticker creation dates. The creation date comes from annual snapshots of currently active tickers taken each January from 1999 to 2007. I match the CRSP Survivor-Bias Free US Mutual Fund and the NASD dataset[5] by matching their NASDAQ tickers, fund names and dates. If a mutual fund share class were terminated before 1999 or if a mutual fund share class were started and terminated between the January snapshots, the mutual fund share class would not be included in the NASD data. Following Evans (2007), I keep only those funds that have a CRSP start date that is greater than or equal to January 1 of 1996 [6]. The tickers are assigned to share classes of mutual funds (not to funds). I calculate the inception date of a mutual fund share class as the

[5] I thank Richard Evans for providing these data.

[6] This allows funds at the beginning of my sample to be incubated for at least 3 years.

minimum date between the first monthly return reported in the dataset and the "first offered date" variable from the CRSP dataset. I identify the share classes associated with each fund and calculate the inception date of the fund as the minimum inception date of its share classes' inception dates. Similarly, the IPO of a fund is the earliest IPO date of its share classes' IPO's.

Return and stock information data are from the Center for Research in Security Prices (CRSP) Monthly Stocks File for NYSE, Amex, and NASDAQ stocks. I eliminate closed-end funds, real estate investment trusts (REIT), American Depository Receipts (ADR), foreign companies, primes, and scores. To correct returns for delisting bias, I use the adjustment proposed in Shumway (1997) and Shumway and Warther (1999) for NYSE/AMEX and NASDAQ respectively. I calculate the share of institutional ownership by summing the stock holdings of all reporting institutions for each stock in each quarter. Stocks that are in CRSP, but without any reported institutional holdings, are assumed to have zero institutional ownership. Ownership greater than one are omitted as they could be a result of double-reporting by institutional investors.

I follow Evans (2007) to assess whether a fund is incubated or not. A fund is considered to be incubated if the difference between the IPO date and the fund's inception date is more than 12 calendar months. Funds that do not meet this criterion are considered to be non-incubated.

2.3 Results

2.3.1 Descriptive Statistics

My sample comprises 214 incubated funds and 591 non-incubated funds. Table 2.1 describes some summary statistics for the sample. The average period of incubation for incubated funds is 25 months. Non-incubated funds tend to be created 3 months before they are launched. The initial assets under management are very different for incubated and non incubated funds on the IPO date. Incubated funds tend to be smaller than non-incubated funds. The average TNA of incubated and non-incubated funds is 14.69 and 47.84 million dollars respectively. It is important to note that the difference in the average of the two type of funds is driven by some very large non-incubated funds (The difference in median TNA for both groups is only around 6 million dollars). Table 2.2 breaks down the investment objectives of new fund offerings for the sample period. The table shows a slight tendency for sector/specialized funds to be non-incubated. On the other hand, funds in more traditional investment objectives, like Growth and Income and Mid-Cap funds tend to be incubated. For instance, we can see that of all the new fund offerings that were incubated only 2.3% were Science and Technology funds compared to 8.0% of all the new offerings that were not incubated.

2.3.2 Fund Incubation and Fund Families' Incentives

It has been previously argued that the high returns of incubated funds in the incubation period are due to self-selection. For instance, suppose that a family of funds starts 10 funds, randomly buys stocks for each fund and holds the mutual fund portfolios for 3 years. At the end of the period, the family can decide to launch the funds with best performance histories (i.e. funds that beat their benchmarks) and hope to benefit from the convex relationship between past performance and fund flows [7]. In the following subsections, I document the extent of the difference between incubated fund returns before and after their IPO. I also study whether families are strategic when choosing IPO dates for their incubated funds.

2.3.2.1 Risk Adjusted Returns

Table 2.3 shows the extent to which the four factor risk adjusted alpha can be biased due to the backfilling of incubated fund returns in the CRSP Survivor-

[7] This practice is not as easy to implement as it seems. Firstly, under SEC rules, incubated funds belonging to the same family need to have distinct investment objectives. For instance, a family can not incubate 10 growth funds and pick the best one to be launched. Secondly, the S.E.C. rules indicate that any previous return history can be used for marketing purposes so long as it is not misleading for investors. In other words, the regulatory framework provides penalties for instances in which incubation is used is a misleading way.

Bias Free U.S. Mutual Fund Database. Similar to Evans (2007), I find that the apparent outperformance of incubated funds disappears once the backfilled return histories are removed. Panel A and B show the factor loadings and 4-factor alphas for incubated and non-incubated funds. I estimate these factors with 36 months of return data for each new fund. I then calculate the average factor loadings and alphas for incubated and non-incubated funds respectively. Panel A uses data since the inception of each fund (i.e with backfilled returns). This panel shows that, on average, incubated funds outperform non-incubated funds by 0.22% per month or 2.7% per year in risk-adjusted terms. It is also important to note that according to this panel, incubated funds tend to load less on the market factor and more on the HML factors than non-incubated funds. They also seem to implement less momentum strategies (higher UMD factor loading) than non-incubated funds at the 10% level of statistical confidence. Incubated funds also carry more idiosyncratic risk as the R^2 of their regressions is on average less than that of non-incubated funds. Panel B portrays a different picture. Panel B uses return history that has not been backfilled (i.e. return data after the mutual fund IPO). In this case, the outperformance of incubated funds is greatly reduced (9 bps per month at the 10% level). The difference in the market and HML loadings persists, although the statistical significance declines. There is no clear difference in the momentum factor loading and the level of idiosyncratic risk is very similar. This results are in line with what has been previous documented in the literature. The results also hint at some important characteristics of the incubation period. In this period, incubated funds seem to load less on well-known sources of risk and seem to implement strategies that have higher idiosyncratic risk. This suggests that risk-return profiles of incubated funds before and after their IPO are different.

2.3.2.2 IPO Date Selection for Incubated Funds

Figure 2.1 plots the average 12-month return decile of incubated funds. As we can see in Figure 2.1, the 12-month return of incubated funds is the highest at around the IPO date. We can see that the average 12-month return decile of incubated funds, peaks three months prior to the IPO. The return decile then remains somewhat constant for about 18 months and declines gradually after that. This illustrates that incubated funds are on average launched when their returns are above the median return of mutual funds. To formally test this conjecture, each month I group all public mutual funds according to their 12-month return decile. I then I regress their return deciles at the end of each month for all incubated funds in my sample on the number of months between the observation date and the IPO date (N). I further split N between negative and positive values. I use net returns as opposed to risk-adjusted returns as most of the evidence shows that the convex relationship between past performance and flows is based on non-risk-adjusted returns (See Del Guercio and Tkac (2002) and Del Guercio and Tkac (2007)) . Therefore, if families pursue funds by offering high past performance, they will focus on past returns in the incubation period. The regression is run for observations that are within 24 months of their fund's IPO date. The results (Table 2.4) confirm the conclusions drawn from Figure 2.1. As funds approach their IPO dates, their 12-month return decile increases (β_1 positive). After the IPO, the fund's 12-month return decile declines steadily (β_2 negative). I further control for the 12-month average assets under management (log of TNA), relative flows[8], investment style and style flows for every fund at time t. Adding this controls do not affect the main results. This evidence shows that families take their incubated funds public in an opportunistic fashion.

2.3.3 Investment Strategies and Fund Incubation

So far, we have seen that returns of incubated funds are unusually high before they are offered to the public. We have also seen that this may be premeditated as families try to exploit the positive relationship between past returns and capital flows by launching their incubated funds when their returns are high. This seems to corroborate concerns regarding fund incubation. If investors use

[8]Relative Flows are defined as $flow_{i,t} = \log\left(\frac{TNA_{i,t}}{TNA_{i,t-1}}\right) - \log(1 + R_{i,t})$

past performance as a proxy for managerial ability, investors in new fund offerings may wrongly infer that incubated funds have higher managerial ability that other comparable new funds. Indeed, Table 2.3 shows that, once funds have been launched, investors in incubated funds are only marginally better off than investors in non-incubated funds (9 bps per month or 1.08% per year at the 10% level of statistical confidence) as far as risk-adjusted returns are concerned. However, it is important to determine if investors are indeed attracted to incubated funds due to their pre-IPO performance. If this is so, fund incubation may be misleading for investors. Secondly, the drop in pre and post IPO returns of incubated funds could be due to costly constraints and coordination problems. During the incubation period, funds are shielded from investors capital. When allows them to take longer-term investments and to avoid coordination problems amongst investors. My data on fund IPO dates allows me to accurately determine when these constraints arise. This, in turn, helps us precisely estimate the effect of these financial constraints on the investment strategies followed by funds.

2.3.3.1 Portfolio Characteristics

Incubated Funds: Pre and Post IPO

The difference in returns of incubated funds before and after their IPO's could be due to financial constraints. As explained before, the threat of capital redemptions due to bad performance will induce manager to, ex-ante, take precautionary measures. To that extent, I study pre and post IPO incubated funds investment strategies by comparing their portfolio holdings in the two years preceding and in the 24 months after their IPO's. The first characteristic I study is the overall liquidity of the funds' portfolios. The results are presented in Table 2.5. In this test, I calculate the value-weighted average of the turnover of stocks held by mutual funds as a proxy for liquidity. Stock turnover is the average monthly turnover (trading volume / shares outstanding) in the last 3 months. Because Nasdaq is a dealer market with double counting of dealer buys and sells, the turnover of stocks traded on Nasdaq and

NYSE/Amex is not directly comparable (see, e.g., Atkins and Dyl (1997)). I follow Nagel (2005) and divide the turnover of Nasdaq stocks by two. I regresses this variable on INCUBATED, which is a dummy that equals one if the fund is incubated and zero otherwise and on PRE-IPO which is a dummy that equals one if the observation is before the IPO and zero otherwise. Atfer the IPO, incubated mutual funds hold securities that have less turnover than non-incubated funds. The difference in their value-weighted average turnover is -0.0142%. After controlling for the size of the fund (natural logarithm of assets under management), the average relative fund flows between the previous and the current disclosure date, the average relative fund flows to the investment style between the previous and the current disclosure date and the fund's investment style, this difference is equal to -0.0132 (10% of the sample average). This indicates that incubated funds hold more illiquid stocks during the incubation period which allows them to earn the liquidity premia associated with them. This is only possible thanks to the insulation from capital redemptions as interim bad performance related to liquidity provision will not lead to capital withdrawals. It is also interesting to note that the value-weighted average turnover of a fund positively covaries with its size, past flows and past investment style flows. As flows and size increase, funds opt to invest in more liquid assets. This is an indication that as funds grow, liquidity management becomes more important.

Next, I analyze the level of portfolio concentration of incubated funds before and after the IPO. To measure portfolio concentration I use the Herfindahl Index.[9] I find that there is no significant difference between the level of concentration between the pre and post IPO period of incubated funds. This points out that incubated funds do not change their level of focus after their IPO. Therefore, any difference in returns in both stages is not due to a change in portfolio concentration.

Another important portfolio characteristic is the value-weighted average of

[9] $herfindahl_{i,t} = \sum_{j \in K} w_{j,t}^2$ where K is the set of stocks held by mutual fund i at time t and w is the portfolio weight of j.

institutional ownership of stocks held by a mutual fund. It is widely documented that a large part of the stocks in the US is not held by institutional investors. It has also been shown that institutional investors are inclined to hold stocks with very narrow characteristics (Gompers and Metrick (2001)). Therefore, funds with portfolios that have stocks with low institutional ownership, will tend to be funds that innovate by researching and investing in stocks that are not in the radar of institutional investors. Furthermore, a fund whose portfolio has a lower value-weighted average of institutional ownership provides retail investors access to investment vehicles that explore different corners of the market. This, in turn, could be translated into greater diversification for retail investors. I define institutional ownership of a stock as the number of shares held by institutions (according to the 13F filings) divided by the total number of shares outstanding. Since institutional ownership is highly correlated with size (Gompers and Metrick (2001), Nagel (2005), Garavito (2010)), I orthogonalize the institutional ownership with respect to size as in Garavito (2010)[10]. This calculates the component of institutional ownership that is not correlated with size. Subsequently, I rank the residual institutional ownership from highest to lowest and then group each quarterly cross-section in 20 groups where the first group (Rank = 1) has the stocks with the lowest residual institutional ownerships and the last group (Rank = 20) has the stocks with the largest residual institutional ownerships. I calculate the value-weighted average of the residual institutional ownership ranking of the stocks held by a mutual fund at a disclosure date. I regress this measure on dummies for incubated funds and for whether the observation is before or after a fund IPO. I find that incubated funds tend to invest in companies with less institutional ownership than non-incubated funds. The difference in means of the incubated and non-incubated fund value-weighted average institutional ownership is -0.4461 (4.7% of the sample average). This shows that, during the incubation period, incubated funds tend to invest in stocks that are less popular with

[10] $Logit(OWNERSHIP) = -7.31 + 1.68\ Log\ size - 0.09\ Log\ size^2 + \epsilon$ where ϵ is the component of institutional ownership orthogonal to size.

institutional investors than after their IPO.

I also compare the levels of portfolio turnover and the expense ratios. Portfolio Turnover as defined by CRSP is the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month total net assets (TNA) of the fund. The expense ratio is defined as Ratio of total investment that shareholders pay for the fund's operating expenses, which include 12b-1 fees. For each fund, I obtain the portfolio turnover and expense ratio for the leading share class (First IPO, first inception date, largest TNA) of each fund. There is no clear difference between the portfolio turnover of an incubated fund during and after the incubation period. The expense ratio does not considerably change before and after the IPO. This helps rule out that the high pre-IPO returns of incubated funds is due to lower expense ratios or lower portfolio turnover.

The results above show that incubated funds change their investment behavior after going public. For example, incubated funds tend to hold more illiquid and lesser-known securities during the incubation. This provides evidence showing that, by removing financial constraints related to capital redemptions, the incubation period allows funds to take on investment strategies (i.e. liquidity provision) that would not be implementable otherwise.

After the IPO: Incubated vs Non-Incubated Funds

I also analyze the same portfolio characteristics as in the previous section for incubated and non-incubated funds and compare their post-IPO differences. The results are in Table 2.5. First, I study the liquidity of incubated and non-incubated fund portfolios. The difference in the value-weighted average turnover between incubated and non-incubated funds is -0.0065. After controlling for other factors, the previous difference changes to -0.0035 (2.7% of the sample average). This shows that incubated funds hold more illiquid stocks that non-incubated funds. The portfolio concentration of incubated and non-incubated funds is also different. The difference in Herfindahl Index between incubated and non-incubated funds in the post-IPO period is 0.0012 (4.7% of the sample average) and statistically different from zero. I also find that incubated funds tend to hold stocks that are less popular with institutional investors than non incubated funds. The difference in pre and post IPO means of the value-weighted average of institutional ownership for incubated funds is -0.098 (0.93% of the sample average). There does not seem to be significant portfolio turnover differences between incubated and non-incubated funds after their IPO's. Incubated funds do seem to have higher expense ratios than non-incubated funds (the difference is 0.07% or 4.6% of the sample average).

The evidence in this subsection shows that, even after incubated funds go public and capital redemption rights are imposed, they remain unique when compared to new fund offerings that were not incubated. After their IPO's, incubated funds tend to hold more illiquid and lesser-know stocks and more concentrated portfolios than non-incubated funds. Two points are noteworthy. Firstly, incubated mutual funds tend to change their investment strategies after their IPO's, but not to a point where they are indistinguishable from other new fund offerings. Secondly, the under-diversification of incubated funds, can be explained in Van Nieuwerburgh and Veldkamp (2008) framework where investors deviate from holding diversified portfolios if they can first collect information about assets systematically. In the case of incubated funds, the incubation period allows managers to collect and learn about assets without having to focus on their short-term performance. Consistent with this fact, I have shown that managers of incubated funds decide to learn about and invest in more marginal, lesser known firms where the benefits of collecting information are the greatest. This, in turn, leads them to concentrate their efforts in smaller set of stocks and hence to have more concentrated portfolios than non-incubated funds. In other words, the evidence suggests that incubated funds use the incubation period to explore different corners of the market and that this learning has a lasting impact well after the IPO.

2.3.3.2 Stock Picking Ability

Incubated Funds: Pre and Post IPO

As explained before, the decline in risk-adjusted returns of incubated funds after the IPO could be explained due to financial constraints. However, it is unclear whether this reduction is also due to a decline in their ability to choose stocks. In order to try to measure stock picking ability of these funds, I study buy-and-hold strategies that buy the portfolio of each fund at their disclosure dates. This method gives us a better indication of whether a funds are good at picking stocks rather than looking at net returns which are influenced by other factors (liquidity management, fees, etc.). More specifically, I implement a trading strategy that buys the portfolio holdings of a mutual fund at portfolio disclosure date t and holds it until portfolio disclosure date $t+1$. I calculate the characteristic-adjusted and excess returns of this strategy for all funds in my sample and compare their difference in means for incubated funds before and after their IPO's respectively. I do this by regressing the characteristic-adjusted (excess) return of this strategy for funds in my sample on INCUBATED, which is a dummy that equals one if the fund is incubated and zero otherwise and on PRE-IPO which is a dummy that equals one if the observation is before the IPO and zero otherwise. I employ observations that are within two years of the IPO date for each fund. Table 2.6 shows the results. As it can be seen, the difference in means of characteristic-adjusted (excess) returns (HOLDING-BASED) of incubated funds before and after their IPO's is 0.46% (0.72%) per month. After controlling for size, flows investment style and style flows, this difference in characteristic-adjusted (excess) returns is equal to 0.39 % (0.47%) per month. This finding corroborates the results from Table 2.3. It is also interesting to see that the return on this trading strategy is negatively correlated with size which is consistent with Chen et al (2004) and Berk and Green (2004). In other words, funds have a decreasing returns to scale technology and are unable to keep up with the same level of returns as they grow. Another interesting finding is that the return of this trading strategy negatively covaries with style flow. This is a sign that as more capital is going after the same set of stocks, current prices are pushed up and expected returns down.

A better test of the persistence of stock picking ability of incubated funds is to implement a buy-and-hold trading strategy based on the fund's purchases and disposals of assets[11]. This trading strategy is based on active decisions taken by the manager and hence is a better reflection of the managers beliefs. To analyze mutual fund trades, I form portfolios based on the buys (BOUGHT) and sells (SOLD) of each mutual fund during disclosure times *t-1* and *t*. I follow Kacperczy Sialm and Zheng (2005) to construct the weights of the BOUGHT and SOLD portfolios respectively. The weights in each portfolio are based on the intensity with which a manager sold/bought a stock between portfolio disclosures and are adjusted to control for momentum in the stock return. I hold these portfolios (BOUGHT, SOLD) until the next disclosure date (time *t+1*). These portfolios are then rebalanced as before. The differences between the mean characteristic-adjusted returns of incubated funds before and after their IPO's are 0.65% and 0.63% for BOUGHT and SOLD respectively. After controlling for size, flow, investment style flow and investment objectives, the differences in mean characteristic-adjusted and excess returns decrease but still remain economically and statistically significant (Characteristic-adjusted, BOUGHT: 0.62% SOLD: 0.54%; Excess-return, BOUGHT: 0.74%, SOLD: 0.55%). The positive difference for the BOUGHT trading strategy indicates that incubated funds are better at purchasing stocks in the incubation period. On the other hand, the positive difference between the pre and post IPO SOLD trading strategies, show that incubated funds are better at selling stocks in the post-IPO period[12]. These results show that there is a *transfer* of skill between the pre and post IPO periods. This is consistent with a shift in focus after the IPO. While during the incubation period funds could concentrate only on researching and buying stocks, after the IPO, funds

[11]Chen Jegadeesh and Wermers (2000) argue that the active decision of trading on a stock is a better indication of the manager's beliefs (and hence investment skill) than the passive decision of holding a stock, since the latter may be driven by non-performance related reasons such as concerns over transaction costs and capital gains taxes.

[12]A positive difference means that a portfolio based on pre-IPO disposals generates a larger return than that of a portfolio based on post-ipo disposals. This indicates that, during the incubation period, incubation funds sell assets to quickly relative to their post-IPO selling actions.

also have to focus on realizing capital gains and liquidity management of the fund in order to meet future expected redemptions.

After the IPO: Incubated vs Non-Incubated Funds

One of the main points in prior studies is that incubated funds do not deliver higher returns than non-incubated funds. It is therefore concluded that incubated funds do not have better managerial ability than non-incubated funds. As in the previous sub-section , I look at managerial skill by looking at the performance of a buy-and-hold strategies based on the holdings, purchases and disposals of new funds. The difference between the post-IPO trading strategies based on holdings of incubated and non-incubated funds is positive but not statistically different than zero (characteristic-adjusted 0.13%, excess return 0.18% per month). This seems to support the hypothesis of no difference in managerial ability between incubated and non incubated funds. However, when I look at trading strategies based on the portfolio changes, I obtain different results. The differences in means of post-IPO characteristic-adjusted returns of the trading strategies for incubated and non incubated funds are 0.35% and 0.27% a month for BOUGHT and SOLD respectively. After controlling for size, flows, flows to the investment style and investment style the differences are 0.31% and 0.21% for BOUGHT and SOLD respectively. The strategy based on the disposals of each fund (SOLD) is not statistically significant and therefore it is inconclusive. The positive mean difference between the characteristic-adjusted returns of the incubated and non-incubated BOUGHT strategies tells us that incubated funds are better at buying stocks than non-incubated stocks. This evidence shows that, when looking at active decisions of fund managers, there is a difference in managerial ability between incubated and non-incubated funds.

2.3.4 Fund Flows

As we have previously seen, families launch their incubated funds when their returns are above the industry median. If families engaged in this strategy to exploit the positive relationship between past performance and flows, it would be plausible to believe that, holding everything else constant, incubated funds would attract more relative inflows than non-incubated funds due to their pre-IPO returns. Figure 2.2 compares the relative fund flows to incubated and non-incubated funds and their TNA evolution for the first 24 months after the IPO. The figure shows two main patterns. Firstly, relative fund flows to non-incubated funds are larger than those of incubated funds in the first 12 months post-IPO. The net fund flows to non-incubated and incubated funds in the second year seem to be similar. Secondly the average TNA of non-incubated funds is larger than that of incubated funds. However, this gap shrinks with time. The difference between the rates at which incubated and non-incubated funds' assets grow, could be due to their fund families. Table 2.7 describes the average family size and number of funds decile of new fund offerings. These variables are calculated for each new fund at the IPO date. Panel A shows that non-incubated funds belong to larger families than Incubated funds. Panel B shows that non-incubated funds belong to more populated families than incubated funds. Therefore, the difference mentioned above could be due to the fact that large families are likely to have better distribution channels through larger broker networks and well-established relationships with institutional investors. Better distributions channels would ensure that new fund offerings reach wider audiences in a shorter period of time. This increases the likelihood that non-incubated funds attract more inflows than incubated funds very early on.

Table 2.8 formally studies the determinants of net fund flows for new fund offering in the first two years after their IPO's. I regress the monthly relative fund flows of each new fund offering on the past month return decile, the family size decile, investment style dummies, the natural logarithm of net fund flows to the investment style of each fund and a dummy that equals one if the fund was incubated and zero otherwise. Panel A shows the results using the observations within 2 years after IPO. As we can see from the univariate regressions, return, family size, relative flows to the investment style explain relative fund flows in the first two years of a fund. The coefficients on return decile and on relative flows to the investment style have signs consistent with previous results in the literature. The coefficient on family size is positive and highly significant. For instance a one decile change in family size decile explains approximately a 0.3% monthly increase in TNA due to relative capital flows. The multivariate regression shows a similar picture. As predicted in Figure 2.2, incubated funds tend to receive around 1% less in relative fund flows than non incubated funds. Panel B runs the same regressions for observations within one year after IPO. From this panel, we can conclude that the negative effect of incubation on funds is more pronounced (incubated funds receive on average around 2.3% less relative fund flows than non-incubated funds) during the first year. Panel C runs the same regression specifications using observations in the second year after the IPO. The results indicate that the negative effect of incubation on relative flows disappears while family size is less important in statistical and economic terms. We can conclude that economies of scale (larger distribution network) work very well in drawing relative fund flows to new fund offerings in the first year. However, this advantage seem to vanish in the second year of the fund. The negative coefficient on the incubation dummy shows that performance in the incubation period does not seem to help draw inflows to the fund. This can have several explanations. First, investors in incubated funds may be agnostic about high pre-IPO return histories. Investors know that returns in the incubation period were obtained without financial constraints imposed by redemption rights. It could also be that fund families cater to investors' demand. In this case, non-incubation could be an opportunistic attempt to satisfy short-lived investors' demand shocks for a particular investment style. If investors tend to chase investment styles that are in vogue to a greater extent than they pursue past performance, the catering hypothesis would help explain the difference in fund flows between incubated and non-incubated funds.

2.4 Conclusion

The evidence presented in this paper suggests that incubated funds behave differently in the pre and post IPO periods. Incubated funds seem to take advantage of the long-term commitment of investors and the lack of payoff complementarities in the pre-IPO period to implement investment strategies that may be risky in the interim but that have positive payoffs if held to "maturity" (i.e. liquidity provision). Incubated funds also take more concentrated bets and invest in stocks that are not part of the typical institutional investor stock universe. I also present evidence showing that incubated funds tend to behave strategically when choosing IPO dates. Incubated funds are offered to the public when their 12-month past performance peaks. Shortly after being launched, their performance declines back to the median performance. Furthermore, I showed that non-incubated funds tend to attract more investment dollars relative to incubated funds in the first post-IPO year. I also show that difference disappears after the first post-IPO year. Additionally, I document that families play a role in attracting investment funds to their new fund offerings.

This findings show that, the self-selection bias in the sample may not be the only reason that explains the high performance of incubated funds in the incubation period. During the incubation period, incubated funds pursue strategies that can only be implemented if there are no investors that can withdrawal funds on demand. Once funds go public, they must keep a certain level of liquidity to meet potential withdraw. My results also show that funds can use the incubation period to try and research new stocks. This is reflected in the levels of concentration of incubated funds. The findings also show that incubated funds invest in stocks that are less popular with institutional investors than stocks held by other new fund offerings. This is a very important feature for retail investors as some of them only have access to capital markets through institutional investors (pension funds, mutual funds, retirement accounts). By exposing investors to other stocks, incubated mutual funds help investors gain exposure to other sources of risk, thereby enhancing their portfolio diversification.

The differences between incubated and non-incubated funds illustrate that fund families utilize fund incubation as a way to offer something different to investors. This innovation could be a new fund manager with no prior track record, or a new investment strategy that has not been tested/well-understood. In other words, fund innovation is a way to "incubate" ideas in a controlled environment. Only after a fund investment strategies have proved their potential, they are offered to investors. Innovating while having to account for payoff complementarities that investors face and the costs associated with performance-induced outflows may not be feasible for managers of new funds.

The evidence presented also suggests that incubated funds do tend to time their IPO's. While fund incubation may be a conduit for innovation in the mutual fund industry, the incentives to launch funds when they outperform and to liquidate them when their returns are high remain. What it is interesting to see is that this decline does not happen automatically. Incubated funds are able to maintain their 12-month performance above the industry level for sometime after the first year. During this time incubated funds have an edge in stock picking ability and in total returns as previously shown. However, liquidity management and payoff complementarities become more important as the fund grows pushing the fund's performance back to the industry mean.

The last contribution of this paper also shows that incubated funds attract less relative fund flows than non-incubated funds. The obvious question is then, why do fund families incubate funds? As explained before, fund incubation is the perfect setting for fund families to try new ideas and managers. However, their strategic behavior in choosing IPO dates reveal the fact that they try to launch their incubated funds when it is optimal, i.e. when they are likely to maximize fund inflows. It remains interesting to understand what is special about *non-incubated* funds when it comes to attract relative fund flows. It could be that non-incubated funds are launched to cater to investors demands.

For instance, suppose that investors are demanding growth funds. It would then be natural to expect fund families to fill the gap between demand and supply. This in turn would cause more flows to the growth investment style and would encourage more families to offer new growth funds. This could be exacerbated by the dynamics of competition of fund families for investment dollars[13]. Families that are able to offer investors a wide arrange of investment options that exploit investors time varying demand will be in a better position to attract investment funds. The catering to investors by asset managers and its implications remain an interesting area for future research.

[13]Massa (2003) argues that fund families exploit investors heterogeneity by offering them the possibility to switch between sibling funds at no cost. He also finds evidence showing that the more mutual funds families can differentiate from one another, the less they need to compete in terms of performance.

2.5 References

Ackermann Carl, Richard McEnally and David Ravenscraft, The Performance of Hedge Funds: Risk, Return, and Incentives, *The Journal of Finance* 54, 833–874

Ackermann, Carl and Tim Loughran, 2007, Mutual Fund Incubation and the Role of the Securities and Exchange Commission, *Journal of Business Ethics* 70, 33–37

Aragon, George O., Share restrictions and asset pricing: Evidence from the hedge fund industry, *Journal of Financial Economics* 83, 33–58

Arteaga, Kenneth R., Conrad S. Ciccotello, and C. Terry Grant, 1998, New Equity Funds: Marketing and Performance, *Financial Analysts Journal* 54, 4349.

Atkins, A.B., Dyl, E.A., 1997. Market structure and reported trading volume: Nasdaq versus the NYSE. *Journal of Financial Research* 20, 291304.

Berk, Jonathan and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy*, 2004, 1269–1295

Chen, Qi, Itay Goldstein and Wei Jiang, 2007, Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows, Working paper, University of Pennsylvania

Chen, Hsiu-Lang, Narasimhan Jegadeesh and Russ Wermers, 2000, The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers, textitJournal of Financial and Quantitative Analysis

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Chordia, Tarun, 1996, The structure of mutual fund charges, *Journal of Financial Economics* 41, 3-39.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Del Guercio, Diane, and Paula A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds, *Journal of Financial and Quantitative Analysis* 37, 523-557.

Del Guercio, Diane and Tkac, Paula A., 2007, Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow, *FRB of Atlanta Working Paper No. 2001-15.*

Edelen, Roger, 1999, Investor flows and the assessed performance of open-end mutual funds, *Journal of Financial Economics* 53, 439–466.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2001, A First Look at the Accuracy of the CRSP Mutual Fund Database and a Comparison of the CRSP and Morningstar Mutual Fund Databases", *Journal of Finance* 56, 2415-2430.

Elton, Edwin , Martin J. Gruber, Joel Rentzler, New Public Offerings, Information, and Investor Rationality: The Case of Publicly Offered Commodity Funds, *The Journal of Business* 62, 1–15.

Evans, Richard, 2006, Does Alpha Really Matter? Evidence from Mutual Fund Incubation, Termination and Manager Change, Working paper, University of Virginia Darden School of Business

Evans, Richard, 2007, The Incubation Bias, Working paper, University of Virginia Darden School of Business

Franco, Joseph A, 2009, A Consumer Protection Approach to Mutual Fund Disclosure and the Limits of Simplification,*Stanford Journal of Law, Business, and Finance*,15, 1–83

Fung, William, and David A. Hsieh, 2002, Hedge-Fund Benchmarks: Information Content and Biases, *Financial Analysts Journal* 58, 22-34.

Garavito, Fabian, 2008, Organization Diseconomies in the Mutual Fund Industry, Working paper, London School of Economics.

Gompers, P.A., Metrick, A., 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 166, 229–260.

Ippolito, Richard, 1989. Efficiency with costly information: a study of mutual fund performance, *Quarterly Journal of Economics* 104, 1-23

Johnson, Woodrow T., 2004, Predictable investment horizons and wealth transfers among mutual fund shareholders, *Journal of Finance* 59, 1979-2011.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2012.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance maximization is not the only game in town, *Journal of Financial Economics* 67, 249–304.

Nagel, Stefan, 2005, Short sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309

Park, James M., 1995, Managed Futures as an Investment Set, doctoral dissertation, Columbia University.

Palmiter, Alan R. and Ahmed Taha, 2009, Star Creation: The Incubation of Mutual Funds, *Vanderbilt Law Review* 62, 1485–1534

Shumway, Tyler, 1997, The Delisting Bias in CRSP Data, *The Journal of Finance* 52, 327–340

Shumway, Tyler and Vincent A. Warther, 1999, The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect, *The Journal of Finance* 54, 2361–2379

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Stein, Jeremy, 2005, Why are most funds open-end? competition and the limits of arbitrage, *Quarterly Journal of Economics* 120, 247-272.

Van Nieuwerburgh, Stijn and Laura Veldkamp, 2008, Information Acquisition and Under-Diversification, Working paper, New York University.

Wermers, Russ, 2004, Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence, Working paper, University of Maryland - Robert H. Smith School of Business

Wisen, Craig H., 2002, The Bias Associated with New Mutual Fund Returns, Working paper, University of Alaska Fairbanks.

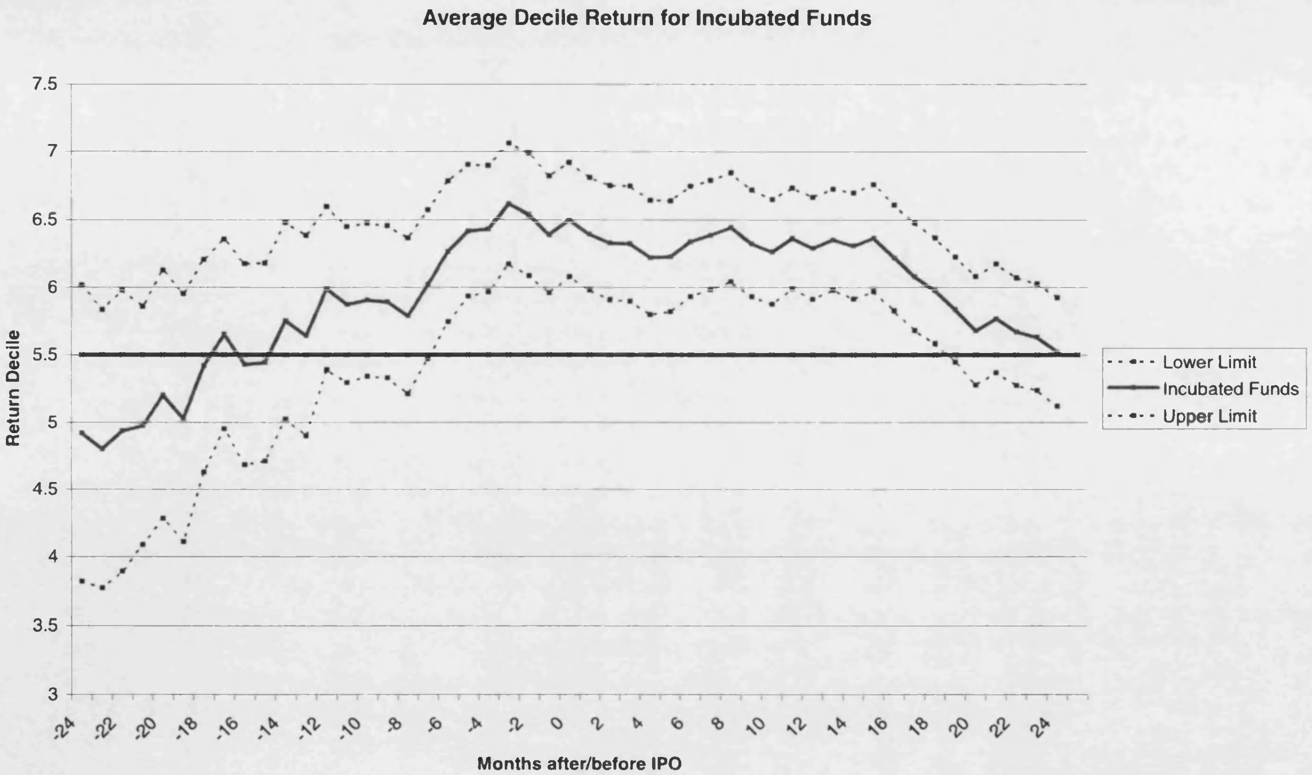

Figure 2.1: Average Decile Return for Incubated Funds

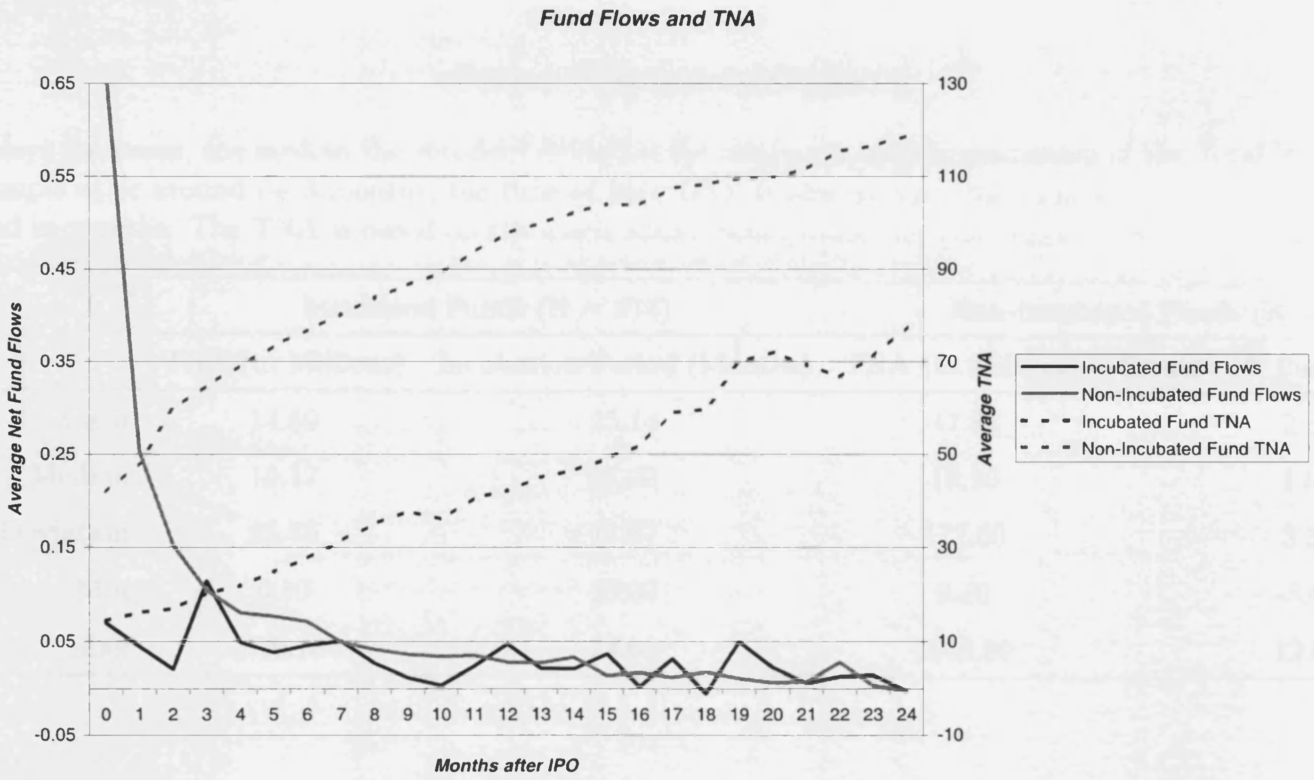

Figure 2.2: Fund Flows and TNA

Table 2.1: **Summary Statistics**

This table displays the mean, the median the standard deviation the minimum and the maximum of the Total Net Assets (TNA) for funds in the sample at or around (+ 3 months) the time of their IPO. It also displays the same statistics for the incubation period length measured in months. The TNA is based on the assets under management for the largest, oldest share class for each fund.

	Incubated Funds (N = 214)		Non-Incubated Funds (N = 591)	
	TNA (in Millions)	Incubation Period (Months)	TNA (in Millions)	Incubation Period (Months)
Mean	14.69	25.14	47.84	2.84
Median	10.17	23.00	16.10	1.00
Standard Deviation	25.76	11.67	172.60	3.35
Min	0.10	13.00	0.10	-3.00
Max	310.70	73.00	2645.80	12.00

Table 2.2: **Investment Objectives**

This table displays the Lipper investment objectives of funds in the sample as of 1999 or later. Frequency is the number of new fund offerings during the sample period. Percent is the percentage of new fund offerings with respect to the total new fund offerings in each category (incubated and non-incubated funds) over the sample period .

	Incubated Funds		Non-Incubated Funds		
	Frequency	Percent	Frequency	Percent	Difference
Capital Appreciation	11	5.1%	26	4.4%	0.7%
Equity Income	4	1.9%	11	1.9%	0.0%
Financial Services	3	1.4%	11	1.9%	-0.5%
Flexible Portfolio	1	0.5%	1	0.2%	0.3%
Growth	76	35.5%	227	38.4%	-2.9%
Growth and Income	35	16.4%	47	8.0%	8.4%
Health / Biotechnology	5	2.3%	23	3.9%	-1.6%
Mid Cap	30	14.0%	63	10.7%	3.4%
Micro Cap	3	1.4%	8	1.4%	0.0%
Specialty/Miscellaneous	2	0.9%	11	1.9%	-0.9%
Small Cap	38	17.8%	113	19.1%	-1.4%
Science Technology	5	2.3%	47	8.0%	-5.6%
Utility	1	0.5%	3	0.5%	0.0%

Table 2.3: **Average Alpha and Factor Loadings for New Funds**

This table displays the mean of regression coefficients and risk-adjusted alphas for funds in my sample. It also calculates the differences in factor loadings and 4-factor alphas for incubated and non-incubated funds. Panel A uses 36 months of data of monthly excess returns from the inception of the fund (the date the fund was created). Panel B uses 36 months of data from the IPO date onwards (after the fund applies and obtains a NASDAQ ticker). t-statistics are in parentheses.

	Intercept	Rm-Rf	SMB	HML	UMD	Average R
Panel A: 36 Months of Data Since Inception						
Non-Incubated Funds	0.02%	1.071	0.219	-0.047	0.042	88.03%
	(0.58)	(90.45)	(13.84)	-(2.48)	(5.12)	
Incubated Funds	0.23%	1.001	0.197	0.034	0.013	85.79%
	(4.44)	(69.17)	(8.08)	(0.98)	(0.91)	
Difference	-0.22%	0.070	0.021	-0.081	0.029	2.24%
	-(3.69)	(3.72)	(0.73)	-(2.06)	(1.79)	(2.78)
Panel A: 36 Months of Data Since IPO						
Non-Incubated Funds	-0.03%	1.065	0.224	-0.050	0.033	88.11%
	-(1.37)	(91.48)	(14.49)	-(2.70)	(3.96)	
Incubated Funds	0.06%	1.013	0.195	0.020	0.015	87.19%
	(1.36)	(68.23)	(8.29)	(0.63)	(0.70)	
Difference	-0.09%	0.052	0.029	-0.070	0.018	0.92%
	-(1.85)	(2.73)	(1.04)	-(1.91)	(0.75)	(1.15)

Table 2.4: **Return Decile Regression**

Dependent variable is the 12-month return decile of fund i at time t. N is the difference between date t and the IPO date. The regression is run for incubated funds where the difference between date t and the IPO date $\in [-24, 24]$. LOG TNA is the average log TNA of mutual fund i's TNA at time t. FLOW is the average relative log flow of fund i at time t. STYLE FLOW is the average log flow to the investment style of mutual fund i at time t. Averages are calculated between time t and time t-11.$\mathbb{1}$ is the indicator function. t-statistics are in parentheses.

$$RETURNDECILE_{i,t} = \alpha + \beta_1 N\mathbb{1}_{N_{i,t}<0} + \beta_2 N\mathbb{1}_{N_{i,t}\geq 0} + \gamma CONTROLS + \epsilon_{i,t}$$

	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6
INTERCEPT	6.55	1.88	6.06	6.17	3.73	
$N\mathbb{1}_{N_{i,t}<0}$	0.08	0.07	0.06	0.07	0.05	0.05
	(8.32)	(7.14)	(6.34)	(7.59)	(5.43)	(5.47)
$N\mathbb{1}_{N_{i,t}\geq 0}$	-0.04	-0.06	-0.02	-0.03	-0.03	-0.03
	-(6.94)	-(10.22)	-(3.60)	-(6.81)	-(5.36)	-(5.15)
LOG ASSETS		0.31			0.13	0.12
		(10.66)			(4.85)	(4.53)
FLOW			6.90		5.53	5.44
			(14.31)		(11.88)	(11.75)
STYLE FLOW				1.27	1.18	1.17
				(26.06)	(24.05)	(22.93)
Fund Styles	N	N	N	N	N	Y

Table 2.5: **Characteristics of Stocks Held by New Mutual Funds**

INCUBATED is a dummy equal to one if fund i is incubated and zero otherwise. PRE IPO is a dummy variable equal to one if observation at time t for fund i is before than the fund's IPO. INTERACTION is and interaction term between INCUBATED and PRE IPO date. TURNOVER is the value weighted average stock turnover of stocks held by mutual fund i at time t. Stock turnover is the average turnover (volume/shares outstanding) of a stock in the last 3 months (t-2, t-1, and t). HERFINDAHL is the Herfindahl index for mutual fund i at time t. INST OWN is the value weighted average rank of the orthogonalized logit institutional ownership of stocks held by mutual fund i at time t. The logit of the institutional ownership of a stock is defined as $\log\left(\frac{OWNERSHIP}{1-OWNERSHIP}\right)$. The orthogonalized logit of institutional ownership of stock i, is defined as the epsilon of the folloging equation: $Logit(OWNERSHIP) = -7.31 + 1.68\ Log\ size - 0.09\ Log\ size^2 + \epsilon$, where OWNERSHIP is the percentage of institutional ownership and size is the market capitalization of stock i at time t. PORT TURNOEVER, is the portfolio turnover of mutual fund i at time t. EXPENSE is the expense ratio of mutual fund i at time t. LOG ASSETS is the natural logarithm of fund i assets under management at time t. FLOW is the average monthly relative net fund flows of fund i between portfolio disclosures dates. STYLE FLOW is the average monthly net fund flows between fund's i portfolio disclosures dates of the investment style to which fund i belongs. Relative net fund flow is defined as in Pollet and Wilson (2007). Fund Styles are defined as in Table 2.2. t-statistics are in parentheses.

$$Y = \alpha + \beta_1 INCUBATED + \beta_2 PRE\_IPO + \beta_3 INCUBATED \times PRE\_IPO + \gamma CONTROLS + \epsilon$$

	TURNOVER	TURNOVER	HERFINDAHL	HERFINDAHL	INST OWN	INST OWN	PORT TURNOVER	PORT TURNOVER	EXPENSE	EXPENSE
INTERCEPT	0.137		0.025		11.919		1.564		0.015	
INCUBATED	-0.0065	-0.0035	0.0003	0.0012	-0.0490	-0.0986	-0.2500	-0.1343	0.0006	0.0007
	-(3.67)	-(2.09)	(0.64)	(2.75)	-(1.10)	-(2.38)	-(2.17)	-(1.17)	(2.63)	(2.95)
PRE IPO	-0.0027	-0.0034	0.0025	0.0005	-0.7213	-0.7555	-1.0487	-1.0262	0.0009	-0.0001
	-(0.59)	-(0.74)	(2.05)	(0.41)	-(6.15)	-(6.76)	-(1.21)	-(1.19)	(0.64)	-(0.06)
INTERACTION	-0.0115	-0.0099	-0.0006	-0.0004	0.3134	0.3095	0.9726	0.8770	-0.0006	0.0000
	-(2.16)	-(1.95)	-(0.41)	-(0.28)	(2.35)	(2.49)	(1.10)	(1.00)	-(0.45)	(0.00)
LOG ASSETS		0.0012		-0.0012		-0.0833		-0.0676		-0.0003
		(2.65)		-(10.83)		-(7.54)		-(2.02)		-(4.10)
FLOW		0.0054		0.0005		-0.1940		0.0411		0.0005
		(2.22)		(0.74)		-(3.27)		(0.10)		(0.82)
STYLE FLOW		0.1633		0.0011		-18.4539		-18.2857		-0.0189
		(2.52)		(0.06)		-(11.60)		-(3.09)		-(1.94)
$\beta_2 + \beta_3$	-0.0142	-0.0132	0.0020	0.0001	-0.4079	-0.4461	-0.0761	-0.1492	0.0002	-0.0001
$\beta_2 + \beta_3 = 0$ p-value	0.00	0.00	0.00	0.86	0.00	0.00	0.67	0.40	0.48	0.83
Fund Styles	-	Y	-	Y	-	Y	-	Y	-	Y

Table 2.6: **Stock Picking Ability of New Funds**

Dependent variable is the return of a HOLDINGS-BASED and TRADING-BASED buy-and-hold trading strategy. The HOLDINGS-BASED strategy buys the portfolio of mutual fund i at time t, and holds it until time t+1. The TRADING-BASED strategy implements a trading strategy between time t and t+1 that is based on the trades between time t-1 and t of fund i. Time t-1, t and t+1 are subsequent portfolio disclosure dates for mutual fund i. DGTW RET is the characteristic adjusted trading strategy return as in Daniel Grinblatt Titman and Wermers (1997). RET-RF is the excess return of the trading strategy. BOUGHT (SOLD) is a trading strategy that buys at time t the the purchases (disposals) of fund i between time t-1 and t. This portfolio is held until time t+1. The portfolio weights for BOUGHT and SOLD are as in Kacperczyk Sialm Zheng (2005). Trading strategies are only based on CRSP U.S. common stocks traded in NYSE, AMEX, and NASDAQ held by mutual funds. Independent variables are defined as in Table 2.5. Stock returns are corrected for the delisting bias present in CRSP following Shumway (1997) and Shumway and Warther (1999). t-statistics are in parentheses.

$$Y = \alpha + \beta_1 INCUBATED + \beta_2 PRE\_IPO + \beta_3 INCUBATED \times PRE\_IPO + \gamma CONTROLS + \epsilon$$

	HOLDINGS-BASED STRATEGY				TRADING-BASED STRATEGY							
					DGTW				RET-RF			
	DGTW		RET-RF		BOUGHT	SOLD	BOUGHT	SOLD	BOUGHT	SOLD	BOUGHT	SOLD
INTERCEPT	-0.91%		-1.05%		-1.05%	-1.10%			-1.15%	-1.20%		
	-(13.47)		-(11.70)		-(11.94)	-(12.47)			-(9.94)	-(11.16)		
INCUBATED	0.20%	0.13%	0.36%	0.18%	0.35%	0.27%	0.31%	0.21%	0.37%	0.33%	0.24%	0.21%
	(1.60)	(1.01)	(2.19)	(1.10)	(2.36)	(1.82)	(2.08)	(1.40)	(1.92)	(1.83)	(1.23)	(1.16)
PRE IPO	0.24%	0.11%	-0.57%	-0.97%	0.62%	-0.37%	0.49%	-0.63%	-0.53%	-1.33%	-1.12%	-1.98%
	(0.73)	(0.31)	-(1.31)	-(2.17)	(0.63)	-(0.37)	(0.49)	-(0.63)	-(0.40)	-(1.09)	-(0.86)	-(1.64)
INTERACTION	0.22%	0.28%	1.29%	1.44%	0.03%	1.00%	0.13%	1.17%	1.48%	2.12%	1.86%	2.53%
	(0.58)	(0.74)	(2.60)	(2.89)	(0.02)	(0.98)	(0.13)	(1.15)	(1.10)	(1.69)	(1.39)	(2.05)
LOG ASSETS		-0.001		-0.003			0.000	-0.001			-0.003	-0.004
		-(1.82)		-(6.26)			-(0.90)	-(2.93)			-(5.48)	-(6.86)
FLOW		0.001		0.000			-0.001	-0.003			0.007	0.003
		(0.61)		(0.08)			-(0.20)	-(0.59)			(1.02)	(0.50)
STYLE FLOW		-0.163		-0.340			-0.281	-0.330			-0.535	-0.513
		-(3.33)		-(5.33)			-(4.16)	-(4.90)			-(6.10)	-(6.33)
$\beta_2 + \beta_3$	0.46%	0.39%	0.72%	0.47%	0.65%	0.63%	0.62%	0.54%	0.95%	0.79%	0.74%	0.55%
$\beta_2 + \beta_3 = 0$ p-value	0.01	0.03	0.00	0.05	0.00	0.01	0.01	0.02	0.00	0.00	0.01	0.05
Fund Styles	-	Y	-	Y	-	-	Y	Y	-	-	Y	Y

Table 2.7: **Fund Family Characteristics**

This table presents differences in means of family characteristics of new funds. Panel A shows the difference in means of the number of member funds decile of families that have launched new funds. Panel B exhibits the difference in means of the assets under management decile of families that have launched new funds. t-statistics are in parentheses.

	Mean
Panel A: Family Size	
Non-Incubated	8.09
Incubated	6.76
Difference	1.33
	(5.82)
Panel B: Number of Funds in Family	
Non-Incubated	8.49
Incubated	7.67
Difference	0.82
	(4.05)

Table 2.8: **Fund Flow Regressions**

Dependent variable, $FLOW_{i,t}$,is the relative net fund flows of fund i at time t . Relative net fund flow is defined as in Pollet and Wilson (2007). $RETURNDECILE_{i,t}$ is the monthly return decile of fund i at time t-1. $FAMILYSIZEDECILE_{i,t}$ is the assets under management decile at time t of the family to which fund i belongs. $LOGFLOWSTYLE_{i,t}$ is the relative fund flows at time t to the investment objective to which fund i belongs. INCUBATED is a dummy that equals one if fund i was incubated and zero otherwise. t-statistics are in parentheses.

	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6	MODEL7
Panel A: 1 - 24 Months After IPO							
INTERCEPT	0.021	0.013	0.034	0.040	-0.002	0.000	0.006
	(5.59)	(2.06)	(17.87)	(18.06)	-(0.33)	(0.01)	(0.74)
RETURN DECILE	0.003				0.003	0.002	0.002
	(4.99)				(4.98)	(3.38)	(3.43)
FAMILY SIZE DECILE		0.003			0.003	0.003	0.002
		(3.70)			(3.70)	(3.69)	(3.10)
LOG FLOW STYLE			1.259			1.195	1.198
			(10.55)			(9.89)	(9.92)
INCUBATED				-(0.01)			-0.010
				-(2.77)			-(2.22)
Panel B: 1 - 12 Months After IPO							
INTERCEPT	0.043	0.022	0.056	0.068	0.005	0.006	0.021
	(6.73)	(2.08)	(17.10)	(17.92)	(0.42)	(0.54)	(1.63)
RETURN DECILE	0.003				0.003	0.002	0.002
	(3.14)				(3.15)	(2.25)	(2.35)
FAMILY SIZE DECILE		0.005			0.005	0.005	0.004
		(3.69)			(3.70)	(3.67)	(2.77)
LOG FLOW STYLE			1.120			1.057	1.044
			(6.31)			(5.90)	(5.83)
INCUBATED				-(0.03)			-0.023
				-(4.03)			-(3.19)
Panel C: 13 - 24 Months After IPO							
INTERCEPT	0.002	0.002	0.014	0.014	-0.011	-0.008	-0.009
	(0.44)	(0.24)	(6.75)	(5.98)	-(1.37)	-(0.97)	-(1.11)
RETURN DECILE	0.002				0.002	0.002	0.002
	(3.81)				(3.80)	(2.52)	(2.51)
FAMILY SIZE DECILE		0.002			0.002	0.002	0.002
		(1.85)			(1.83)	(1.82)	(1.90)
LOG FLOW STYLE			1.215			1.146	1.141
			(7.83)			(7.29)	(7.24)
INCUBATED				(0.00)			0.003
				(0.65)			(0.59)

Chapter 3

A Catering Explanation of Mutual Fund IPOs

3.1 Introduction

Style valuations and style inflows predict mutual fund launches. The launches so predicted rarely outperform either their style benchmark or the universe of substitute funds and take place at their valuation peak . Our results imply that the marginal investor in equity mutual funds is precisely the least well-informed about investment opportunities.

The catering theory of Baker and Wurgler (2004) is applicable to a large range of corporate actions, including the decision to launch new mutual funds. In an efficient market, in which all investors are rational and informed, such launches should only take place in response to genuine investment opportunities where the previous range of products failed to span the space of asset payoffs given transactions costs. By contrast, in an inefficient market, shocks to the demands of poorly informed investors (Frazzini and Lamont's (2008) "dumb money" investors) will call forth a supply of essentially useless financial products designed to capture a share of the fees likely to be generated by these investors' relatively poorly informed decisions. The new funds so called forth are useless first because there are close substitutes already in existence and secondly because they proliferate the number of products, increasing search and other costs for all investors.

Mutual fund providers introduce new funds in two different ways. Funds either go from conception to birth very swiftly (non-incubated) or are first managed only with seed capital from the provider (incubated) until a subsequent IPO is arranged. Despite the obvious problems posed by incubation [1], incubation of funds is less likely to involve highly opportunistic provision of publicly investable funds solely in response to perceived demand by investors. We show that although our chosen demand indicators are very good at predicting mutual fund launches in general, and non-incubated fund launches in particular, they are much less able to predict the decision to go public for incubated investment funds.

If investors inflows are "sticky" (Choi, Laibson and Madrian (2007)) then opportunistic fund launches can permanently increase fee income to fund providers even in the absence of providing a genuinely useful service. In consequence, catering to short-lived investor demand shocks can permanently increase the population of mutual funds and can potentially explain the vast population of mutual funds relative to the relatively small variety of independent equity investment strategies that have been shown to produce high average returns. This explanation for the proliferation of mutual funds has obvious implications for hedge funds and structured financial products.

We also document that the dynamics of expected fund flows to new funds is conditional on investors' demand for investment styles. More specifically, we show that the positive relationship between past performance and current flows decreases as demand for the investment style at the time of the IPO increases. Additionally, the positive relationship between past investment style flows and flows to new mutual funds increases as investment style demand at the time of the IPO increases. These changes are stronger in the first year after the IPO. Therefore, when demand for an investment style is high (when catering

[1] Evans (2009) discusses how fund families have the incentive to launch only incubated funds with outstanding past performance so as to benefit from the convex relationship between past performance and current flows. Palmiter and Taha (2009) discuss the current regulatory framework for incubated funds

is more likely to happen), the importance of investor sentiment (as measured by flows into the style) in determining expected flows to new funds increases where as the significance of past performance declines.

Finally, we show that without such IPOs, mutual fund providers would miss out on capturing hot money style flows. IPOs increase the level and persistence of flows into hot styles, consistent with an attention-grabbing role for such launches. IPOs are perhaps the most important tool for mutual fund providers to attract and retain dumb money investment.

Our paper builds on the earlier work of Khorana and Servaes (1999) and Evans (2009), who use industrial organization arguments to explain the probability to launch a new fund by a family, but who do not consider demand shocks as the primary determinant of total new mutual fund offerings. Our paper is also related to the literature on IPO waves and IPO timing. Pastor and Veronesi (2005) construct a rational model where entrepreneurs time their IPO offerings when market conditions are favorable. Loughran and Ritter (1995) show that the number of stock IPO's in hot IPO markets are due to greater investment optimism. Baker and Wurgler (2000) find evidence of stock issuance timing by corporations and Lowry (2002) finds that fluctuations in investor sentiment explains the fluctuations in IPO issuance. Lee Shleifer and Thaler (1991) show that discounts of closed-end funds vary with investor sentiment and that the number of new closed-end fund offerings fluctuates with the closed-end fund discount.

Our paper is also related to the catering literature. Baker and Wurgler (2004) present evidence that managers cater to investors by paying dividends when investors most want to hold dividend-paying securities. Polk and Sapienza (2008) find evidence that stock market mispricing affects the level of firms set their investment policies since managers know that investors pay a premium for companies with high levels of investment. Baker Greenwood and Wurgler (2007) propose and test a catering theory of nominal stock prices whereby investors prefer low-price firms and therefore managers maintain share prices at lower levels.

This study complements earlier work in mutual fund proliferation. Massa (1998) argues that market segmentation and fund proliferation can be seen as marketing strategies used by the families to exploit investors' heterogeneity. Massa (2003) finds that the degree of product segmentation has a positive effect on mutual fund proliferation.

This paper is also linked to research in the effect of investor sentiment on mutual fund flows. Goetzmann Massa Rouwenhorst (1999) document that flows to different investment styles may be associated to investor sentiment about the equity premium. Brown Goetzmann Hiraki Shiraishi and Watanabe (2002) find evidence that daily mutual fund flows may be a proxy for investor sentiment. They also propose an index to measure market sentiment based on how investors move funds in and out of different styles. Frazzini and Lamont (2008) find evidence that investor sentiment for stocks, as measured by mutual fund flows, forecast future returns.

The rest of the paper is organized as follows: section 3.2 describes our data; section 3.3 describes our results; section 3.4 concludes.

3.2 Data

Our sample starts in January 1996 and ends in December 2005. We calculate the inception date[2] of a mutual fund by looking at the inception dates of its share classes[3]. We calculate the inception date of a mutual fund share class as the minimum date between the date of its first monthly return observation and the "first offered date" variable from the CRSP Survivor-Bias Free US Mutual Fund dataset. We identify the share classes associated with each fund and calculate the inception date of the fund as the minimum inception date of its share classes. As a proxy for mutual fund share class IPO date, we use

[2]This is the date on which the SEC has approved the N-8A and N-1A forms of a new fund.

[3]Mutual funds usually issue several share classes. Shares classes are claims to the same underlying portfolio with different fee structures.

NASDAQ mutual fund ticker creation dates[4]. We calculate the IPO date of a mutual fund as the minimum IPO date of its share classes. We match the CRSP Survivor-Bias Free US Mutual Fund and the NASD datasets by matching NASDAQ tickers within 24 month of the mutual fund share class IPO. For funds with more than one match[5], we manually choose the appropriate match. Following Evans (2009), we keep only those funds that have an inception date that is greater than or equal to January 1 of 1996[6]. A fund is considered to be incubated if the difference between the IPO date and the fund's inception date is more than 12 calendar months. Funds that do not meet this criterion are considered to be non-incubated.

We obtain mutual fund data from the CRSP Survivorship-Bias Free U.S. Mutual Fund Database and from the Thompson Financial Mutual Fund Database (S13). A sample of actively managed US equity funds is constructed in four steps. First, index funds are filtered out from the CRSP Survivor-Bias Free U.S. Mutual Fund Database. Secondly, we eliminate funds (FUNDNO'S) that have reporting gaps of more than 12 months[7] from the Thompson Financial Dataset (S13). Thirdly, we exclude funds that on average do not hold at least 80% of NYSE-AMEX-NASDAQ common stocks from the Thompson Financial Mutual Fund Database (S13). Lastly, we merge these two datasets (CRSP and Thompson) by using the MFLINKS provided by WRDS.

Return and stock information data are from the Center for Research in Security Prices (CRSP) Monthly Stocks File for NYSE, Amex, and NASDAQ stocks. We eliminate closed-end funds, real estate investment trusts (REIT), American Depository Receipts (ADR), foreign companies, primes,

[4]When a new fund is first sold to the public, the fund sponsor or family applies for a NASDAQ ticker. The NASD keeps a record of the date that each fund's ticker was created. Our sample consists of annual snapshots of currently active tickers taken each January from 1999 to 2006. Therefore, if a fund were terminated before 1999 or if a fund were started and terminated between the January snapshots, the fund would not have to be included in the NASD data. See Evans (2009) for more information.

[5]Some Nasdaq tickers are recycled, therefore it is possible to match a unique CRSP fund share class identifier (crsp fundno) with more than one ticker - ipo date pair

[6]This allows funds at the beginning of our sample to be incubated for at least 3 years.

[7]These funds are suspect as Thompson usually recycles unique fund identifiers. See WRDS User's Guide to Thomson Reuters Mutual Fund and Investment Company Common Stock Holdings Databases (July 2008).

and scores. To correct returns for delisting bias, we use the adjustment proposed in Shumway (1997) and Shumway and Warther (1999) for NYSE/AMEX and NASDAQ respectively. The book value of equity in the numerator of the book-to-market ratio (B/M) is taken from the Compustat Database as defined in Daniel and Titman (2006). At the end of each quarter t, we calculate B/M as the book value of equity from the most recent fiscal year-end that is preceding quarter-end t by at least six months divided by the market value of equity at the end of quarter t. Consistent with Fama and French (1993), we exclude firms with negative book values. The Book-to-Market ratio of the S&P 500 index is retrieved from Bloomberg.

We construct the monthly investment style book-to-market ratio for the styles in our sample by calculating the monthly value-weighted average book-to-market ratio of stocks that fall in each investment style. Table 3.4 explains how the universe of stocks for each investment style is defined.

Tables 3.1 through 3.3 document the characteristics of our sample. Table 3.1 shows the summary statistics for the investment styles in the sample. A couple of things are noteworthy. The correlation between investment style flows and return is positive. Since flows tend to be persistent, this shows a positive relation between performance and flows. The market capitalization of each investment style is negatively correlated with its book-to-market ratio. Large investment styles (i.e. Growth) will have companies that on average have low book-to-market ratios. Table 3.2 summarizes the mutual fund characteristics of funds in our sample. It is important to note that investment style flows and fund returns are correlated. Since fund returns are correlated with the return of their investment styles, the correlation between investment style returns and fund flows shows that investors send their money to styles that have done well (Warther (1995)).

Moreover the age of a fund is negatively correlated with its relative flows and its investment style flows. This show early evidence that funds tend to be launched when sentiment is high and that as time goes by and sentiment changes, so change the dynamics of expected flows. Table 3.3 breakdown the number of new fund offerings in our sample by investment style. As we can see, growth has the most number of new offerings during the sample. This is to be expected as growth was a very popular investment style during part of our sample. It is also interesting to note that most of the funds in our sample are non-incubated (73%). Therefore, non-incubation is by far the prefered method that families use to launch new mutual funds.

3.3 Results

3.3.1 Determinants of the Number of Fund IPOs

We want to establish whether measures of investors' demand and investment style valuation can explain the average number of new fund offerings. As we can see in Figure 1, the number of new funds offerings covaries positively with the Market-to-Book ratio of the market portfolio (S&P 500 Index). In other words, when valuations are very high, the number of new funds that come to market is higher than average and vice versa. In order to explore this fact, we run Poisson regressions on the number of funds that are launched each month in our sample on valuation metrics and investor demand. As a valuation metric, we use the book-to-market ratio of each investment style in our sample. The relative investment fund flows relative[8] to each investment style in our sample serves as a proxy for investors' demand for each investment style. The results are presented in Table 3.5. In the first column, we regress the number of *all* new funds in each month on lagged book-to-market ratios for the market portfolio after controlling for net dollar flows into the mutual fund industry. The average number of total funds that IPO covaries negatively with previous book-to market ratios. A one standard deviation increase in the S&P 500 B/M ratio (valuation decreases), leads to a 33% decline in the expected number of new funds. In other words, when valuations are high (the book-to-market-ratio of the market portfolio is low) the number of expected new funds that

[8] Relative flows are defined as $flow_{i,t} = \log\left(\frac{TNA_{i,t}}{TNA_{i,t-1}}\right) - \log\left(1 + R_{i,t}\right)$

IPO is high and vice versa. The same picture emerges when we disaggregate the data and examine it at the investment style level. This controls for the heterogeneity found in the data (i.e. whether the result depends on a particular investment style). Columns two and three show that the number of expected new funds that are launched, depends on the valuation and the demand for each investment style. In column four, we control for other covariates (relative return ranking, style size and the overall valuation of the market). We see that after controlling for other variables, we still see the same covariation pattern between lagged valuation of and demand for an investment style and the expected number of funds that IPO. The effect is statistically and economically significant. A one standard deviation increase in the investment style book-to-market ratio leads to a 17% decrease in the number of expected number of funds that are offered to the public in a particular month. A one standard deviation increase in the relative net flows to an investment style leads to an increase of 21% in the number of new funds that IPO.

Catering to investors will depend on how fast new funds can be launched. This could be a lengthy process due to registration requirements[9]. Therefore, it is natural to think that most of the funds that are offered to uninformed investors are non-incubated. Next, we disaggregate the data by whether funds were incubated or not and investigate if the catering hypothesis is stronger for non-incubated funds. Columns 5 and 6 of Table 3.5 present the results. We see that the number of non-incubated funds depends on past levels of valuation of and demand for investment styles. The result is also economically significant. A one standard deviation increase in the investment style book-to-market ratio decreases the expected number of new non-incubated funds by 20.5%. On the other hand a one standard deviation increase in the style flow leads to a 27% increase in the number of new non-incubated funds. The past levels of

[9]In the U.S., once a family has decided to launch a fund, they have to notify the Securities and Exchange Commission (S.E.C.) of the registration (Form N-8A). The family has to file the registration statement (Form N-1A) with the SEC within three months after the filing of the N-8A form. Upon receiving these forms, the SEC will examine them. Once they feel that all the registration requirements have been met, the S.E.C. issues a Notice of Effectiveness which effectively registers the mutual fund as an open-end investment vehicle regulated by the S.E.C.

demand (investment style flows) forecasts the number of incubated funds but its significance and economic importance is smaller. This indicates that non-incubated funds seem to be the preferred mechanism utilized by fund families to cater to investors. This makes sense as catering with incubated funds would involve forecasting long-term investor demand. Moreover, if catering is mainly provided via non-incubated funds, catering seems to be a profitable strategy for families as the relative flows to non-incubated funds are higher than those to incubated funds (Garavito (2008)). We will further explore this point later in the paper.

3.3.2 Evolution of mutual funds' book-to-market ratios after IPO

From Table 3.5, we see that when valuations are high, we observe a high number of new fund offerings. Our conjecture is that investors in new mutual funds are on average buying a portfolio of securities that is overvalued. Table 3.6 shows that funds tend to IPO when their valuations are at their highest and quantifies the extend of the overvaluation. The table regresses the new funds' book-to-market ratios on the age of the fund. The hypothesis is that as time goes by, the valuation level of the stocks held, holding everything else constant, decreases. We define the book-to-market ratio of a mutual fund as the value-weighted average of its stocks' book-to-market ratio natural logarithms. We define age of a mutual fund as the natural logarithm of one plus the number of quarters between the IPO date and the observation date. We consider all mandatory portfolio disclosures for the first 12 quarters of existence of funds in our sample. Therefore, each observation corresponds to the value-weighted average book-to-market ratio of a mutual fund in our sample reported on a mandatory portfolio disclosure date. The first column shows that the book-to-market of a fund increases (valuation of portfolio decreases) with time. However, there are many other covariates that affect the book-to-market ratio of a fund. It is important to control for the book-to-market ratio of the investment style given the mean-reverting nature of book to market ratios. In other words, if a fund is launched when the book-to-market ratio of its investment style is low, it is natural to expect that this book-to-market ratio reverts back to its mean. It is also important to control for past flows since it has been documented that funds tend to imperfectly scale their holdings when facing flows (Pollet and Wilson (2008), Lou (2009)). We also control for fund size as funds' liquidity management becomes more important for larger funds. Additionally, we control for past fund returns as equity returns and book-to-market ratios covary. After controlling for these variables, we still see that the average book to market ratio of mutual funds covaries positively with age. To control for fund and time heterogeneity we add fund and time fixed effects. The results are very similar. The economic significance is also important. For instance, the difference in book-to-market ratio of an average fund when launched and 10 quarters after is 0.028. That is 10 % of the sample average. To show that the result is not driven by the rise and collapse of the growth investment style in the late nineties, we run the same regression specification without growth funds in column 6 of Table 6. The results are economically and statistically similar.

3.3.3 New Funds' Risk-Adjusted Returns

Next, we investigate whether investors that rush to invest in these new fund offerings obtain higher risk-adjusted returns. We test this hypothesis by comparing the risk-adjusted return of a value-weighted portfolio of new funds and a control portfolio at each month-end date in our sample. Each portfolio of new funds is made of the new funds that were launched during each month. The control group consists of a value weighted portfolio of funds that are below of the 20th percentile of the mutual fund size distribution at each month-end date. This is an appropriate control as new funds are small and as mutual fund returns covary with size (Chen et al (2004)). We track the return of these portfolios for two or three years and calculate their Carhart alphas and betas. We repeat this process for each month in our sample. After obtaining

Carhart coefficients for each month-end portfolio, we calculate average alpha and betas for funds in the IPO and CONTROL portfolios respectively. The results are reported in Table 3.7. Panel A and B track these portfolios for 24 and 36 months after their formation respectively. Two facts are noteworthy. First, none of the portfolios has an alpha that is significantly different for zero. Therefore, new fund offerings do not achieve positive risk-adjusted returns. In other words, investors in these funds could replicate their performance by buying and holding passive investment instruments (i.e. buying and holding indices that replicate the Carhart factors). The added benefit with this passive strategy is the smaller fees associated with them. Moreover the difference in alphas of both portfolios is also not different from zero and both portfolios have very similar loadings on the Carhart factors. This indicates that new funds have very close substitutes in existence. Hence investors in new funds do not benefit from adding these new instruments to their portfolios as they do not provide different risk-return profiles than funds currently in offer.

3.3.4 New Funds' Expected Flows

So far we have seen that new mutual fund investors rush into investment styles that are in vogue. Fund families actively engage in providing funds that follow these styles. However, these new funds are overvalued and do not outperform comparable funds. The question that arises next is whether catering to investors is a profitable strategy for fund families. Funds derive their income from fees that depend on the level of assets under management. Therefore, funds that seek to increase their income will try to attract new flows. Next, we study the determinants of expected fund flows to new funds. This will help us determine the extend to which catering benefits new funds and their families. More specifically, we want to see if funds that are launched when demand for their style is high, i.e. when catering is more likely to happen, experience different dynamics in their expected fund flows than otherwise. To this end, we regress monthly (relative) fund flows on covariates that determine flows (past return, past investment style flow) and on a proxy for demand.

As a proxy for investors' demand for an investment style, we use the average monthly relative investment style flow during the past 6 months. We also add time and style fixed effects. Our sample consists of monthly observations during the 24 months after the IPO. The results are summarized in Table 3.8. The first column shows that the expected fund flows depend on past returns and past investment style flows. Demand for the investment style at the time of the IPO does not affect the average fund flows. Column 2, however, shows that this demand affects the dynamics of expected fund flows to new funds. When demand for an investment style is high at the time of IPO, the positive relation between past returns and flows weakens (DEMAND X RET) whereas the positive relationship between past style flows and fund flows gets stronger (DEMAND X STYLE FLOW).

The significance of these results is different for flows in the first and second year after a fund's IPO. In columns 3 and 4, we split the sample in two and run the same regression specification. The first sub-sample (Column 3) is the set of observations that takes places during the first year of existence of new fund in our sample. The second sub-sample (Column 4) are observations from the second year of the funds in our sample. We observe that the marginal contribution of past performance on capital flows to a mutual fund is lower in the first 12 months of a fund (DEMAND X RET coefficient is lower in the first part of the sample). On the other hand, the marginal contribution of investment style flows on net fund flows to new funds in stronger in the first year of a fund (DEMAND X STYLE FLOW coefficient is higher in the first part of the sample). A more direct test is to create a dummy variable (POST) that equals one for observations in the second year of a mutual fund and zero for observations in the first year. We see that the marginal contribution of past style flow conditional on investment style demand is higher in the first part of the sample (DEMAND X STYLE FLOW X POST is negative). The marginal contribution of past performance conditional on investment style demand is higher in the second year after the IPO (DEMAND X RET X POST). However, the latter effect is not statistically distinguishable from zero.

This indicates that the marginal effect of past investment style flows on relative flows to new funds is increased when investor sentiment at the time of the IPO is high. On the other hand, the marginal effect of past returns on flows (the performance-chasing behaivor of mutual fund investors) is lower when investor sentiment is high at the time of the IPO.

3.3.5 Effect of Fund IPOs on Investment Style Fund Flows

So far, we have documented that investors invest in new funds with investment styles in high demand and that this strategy does not pay off for them. This is the dumb-money effect, i.e. money follows hot investment styles that do not deliver high risk-adjusted returns. However, the most important question is whether catering to investors' demand, by launching pre-packaged portfolios of stocks that follow hot investment styles, is an important dumb-money channel. After all, one could argue that dumb-money would flow to hot investment styles even if there were no new fund IPO's. Table 3.9 explores this issue. The idea is to see the effect that IPO's have on flows to investment styles. We regress style flows on the past number of new funds (natural logarithm of number of new funds plus one) in each style and controls (past performance, past investment style flows, time and fixed effects). From the first column we see that the number of funds launched in the previous 2 months covaries positively with the flows into their investment style. A one standard deviation increase in the number of funds leads to a 0.006% increase in the expected relative fund flows to an investment style. This increase equals 12% of the sample average. Therefore new fund offerings often determine the level of expected fund flows to investment styles. This could be due to frictions in the market for mutual funds (i.e. switching costs between mutual funds, mutual fund brokers' not offering mutual funds in vogue, contracts between fund families and mutual fund brokers, investors' search costs, etc) as it may not be easy for investors to access existing funds that follow popular investment styles. In the second column, we interact the past number of new funds and past investment style flows. We see that the marginal effect of past investment style flows on current investment style flows is higher when the number of fund launches in that style is higher. In other words, the number of IPOs not only increases the expected flows into a style, but also increases its "stickiness"

3.4 Conclusion

We show that the mutual fund industry understands its customers pretty well, even if it does not understand how to generate high risk-adjusted average returns. In contrast to the previous literature, we focus on the willingness of fund providers to create new funds in order to capture investor demand for temporarily hot styles. We show that such willingness predicts fund launches, except for non-opportunistic incubated funds. It is possible that a catering theory of mutual fund provision can explain the vast number of available funds, provided that some mutual fund investors never remove their money. If so, the population of mutual funds is the outcome of a random series of temporary investor demand shocks.

From a policy perspective, our results suggest that providing new funds is too easy, and that marketing considerations may reduce the welfare value of mutual funds, which primarily benefit investors by offering low-cost diversification strategies within a given style. In addition, the preponderance of uniformed investor flows may impose externalities on all other economic agents that is exacerbated by easy mutual fund provision. These externalities may come in the form of payoff complementarities. If investor sentiment changes, uninformed investors will reallocate funds across investment styles. Since open-end mutual funds allow investors to redeem their shares at the funds current Net Asset Values (NAV) at any point in time, funds in investment styles experiencing redemptions will have to conduct unprofitable trades to meet redemptions. These unprofitable trades are externalities imposed by uninformed investors on other investors. Informed investors will know this ex-ante, and will try to anticipate the fund flow of uniformed investors[10]. This will induce higher volatility in flows to investment styles and thus will exacerbate any temporary price deviations away from fundamentals.

[10]This is related to the literature on bank runs and currency crises. The basic idea is that the last agent pulling out of a bank or a currency will bare the brunt of the losses. Ex-ante, all agents know this and will rush to withdraw causing financial fragility. Chen Goldstein and Jiang (2007) model this externalities in the context of mutual funds.

3.5 References

Baker, Malcom and Jeffrey Wurgler, 2000,"The Equity Share in New Issues and Aggregate Stock Returns", Journal of Finance, 55, 5, 2219-2257

Baker, Malcom and Jeffrey Wurgler, 2004,"A Catering Theory of Dividends", Journal of Finance, 59, 3, 1125-1165.

Baker, Malcom, Robin Greenwood and Jeffrey Wurgler, 2008,"Catering Through Nominal Share Prices", NBER Working Paper

Brown, Stephen J., William N. Goetzmann, Takato Hiraki, Noriyoshi Shiraishi and Masahiro Watanabe, 2003,"Investor Sentiment in Japanes and U.S. Daily Mutual Fund Flows", NBER Working Paper

Chen, Joseph, Harrison Hong, Ming Huang, Jeffrey D. Kubik, 2004, "Does fund size erode mutual fund performance? The role of liquidity and organization", American Economic Review 94, 1276–1302.

Chen, Qi, Itay Goldstein and Wei Jiang, 2007, "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outows", Working paper, University of Pennsylvania

Choi, James, David Laibson and Brigittie Madrian, 2007, "The Flypaper Effect in Individual Investor Asset Allocation", working paper, Yale University.

Evans, Richard, 2009, The Incubation Bias, Journal of Finance, forthcoming.

Daniel, Kent and Sheridan Titman, 2006, "Market Reactions to Tangible and Intangible Information ", Journal of Finance, 61, 4, 1605-1643.

Fama, Eugene and Kenneth R. French, 1993, "Common risk factors in the returns on stocks and bonds ", Journal of Financial Economics, 33, 1, 3-56

Frazzini, Andrea and Owen Lamont, 2008, "Dumb money: mutual fund flows and the cross-section of stock returns", Journal of Financial Economics, 88, 2, 299-322.

Garavito, Fabian, 2008, "Mutual Fund Incubation: Innovation or Marketing Tool?", working paper, London School of Economics.

Goetzmann, William N., Massimo Massa and K. Geert Rouwenhorst, 1999, Behavioral Factors in Mutual Fund Flows, Yale School of Management, Working Paper

Khorana, Ajay and Henri Servaes, 1999,The Determinants of Mutual Fund Starts , Review of Financial Studies, 12, 1043-1074.

Lee, Charles M. C.,Andrei Shleifer, Richard H. Thaler 1991, Investor Sentiment and the Closed-End Fund Puzzle, Journal of Finance, 46, 1, 75-109

Lou, Dong, 2009, A Flow-Based Explanation for Return Predictability, Working paper, London School of Economics.

Loughran, Tim and Jay R. Ritter 1995,The New Issues Puzzle, The Journal of Finance, 50, 1, 23-51

lowryLowry, Michelle 2003, Why does IPO volume fluctuate so much?, Journal of Financial Economics, 3–40

Massa, Massimo 1998, Why so many mutual funds, INSEAD, Working Paper

Massa, Massimo 2003, How do family strategies affect fund performance? When performance-maximization is not the only game in town, Journal of Financial Economics 67, 249–304

Palmiter, Alan and Ahmed Taha, 2009, Star Creation: The Incubation of Mutual Funds, Vanderbilt Law Review, 62, 5, 1485-1534

Pastor, Lubos and Pietro Veronesi, 2005, Rational IPO Waves, The Journal of Finance, 60, 4, 1713-1757

Polk, Christopher, and Paola Sapienza, 2009, The Stock Market and Corporate Investment: A Test of Catering Theory , Review of Financial Studies, 22, 187-217

Pollet, Joshua and Mungo Wilson, 2008,"How does size affect mutual fund behavior?", Journal of Finance, 63, 6, 2941-2969.

Shumway, Tyler, 1997,"The Delisting Bias in CRSP Data", Journal of Finance, 52, 1, 327-340.

Shumway, Tyler, and Vincent Warther, 1999,"The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect", Journal of Finance, 54, 6, 2361-2379

Warther, Vincent, 1995, "Aggregate Mutual Fund Flows and Security Returns", Journal of Financial Economics, 39, 2-3, 209-35

Figure 3.1: **New Funds vs S&P 500 Market-to-Book Ratio**

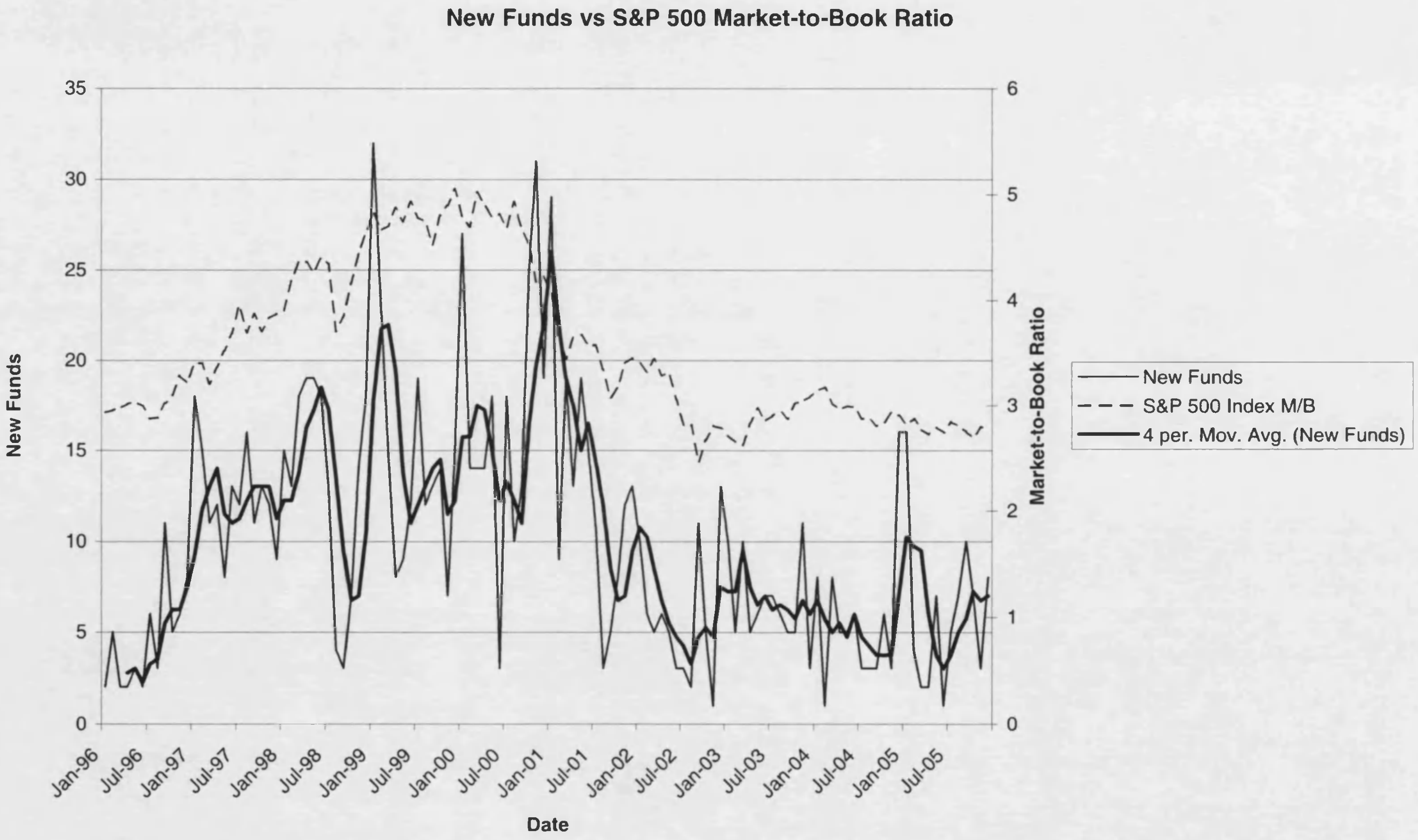

Table 3.1: **Summary Statistics - Investment Styles**

INVESTMENT STYLE B/M is the lagged average book-to-market ratio of the cross-section of stocks that belong to investment style. INV STYLE FLOW is investment style i's lagged monthly fund flow. Flows are defined as $flow_{i,t} = \log\left(\frac{TNA_{i,t}}{TNA_{i,t-1}}\right) - \log(1 + R_{i,t})$. SP 500 B/M is the lagged S&P 500 index's book-to-market ratio. INV STYLE RET is the lagged value weighted average decile of funds that belong to investment style i. INV STYLE SIZE is the natural logarithm of the assets under management of style.

	SIZE	INV STYLE RET	S&P 500 B/M	INV STYLE FLOW	INV STYLE BM
Panel A: Means and standard deviations					
Mean	110878	0.79%	0.30	0.005	0.70
Standard deviation	152030	6.00%	0.06	0.018	0.56
Number of Observations	1320	1320	1320	1320	1320
Panel B: Correlations					
SIZE		-0.03	-0.11	-0.04	-0.39
INV STYLE RET			-0.06	0.38	-0.01
SP 500 BM				-0.02	-0.04
INV STYLE FLOW					0.00

Table 3.2: **Summary Statistics - Mutual Funds**

VWA BM is the value-weighted average book-to-market of a fund at the portfolio disclosure date. AGE is the number of quarters between a portfolio disclosure date and the IPO date of the fund plus one. FLOW is the average monthly net fund flows to fund between disclosure dates. Flows are defined as in Table 3.1. INV STYLE FLOW is the average monthly net fund flows between portfolio disclosure dates. RET is the average monthly return of fund between disclosure dates. SIZE is the natural logarithm of the assets under management.

	VWA BM	LOG FLOW	INV STYLE FLOW	RET	AGE	SIZE
Panel A: Means and standard deviations						
Mean	0.28	0.0676	0.0039	0.56%	6.77	17.63
Standard deviation	0.12	0.2720	0.0089	4.23%	3.37	1.62
Number of Observations	9533	9406	9534	9419	9534	9438
Panel B: Correlations						
VWA BM		-0.05	-0.17	-0.09	0.08	-0.13
LOG FLOW			0.16	0.06	-0.26	0.01
INV STYLE FLOW				0.31	-0.21	0.10
RET					-0.04	0.08
AGE						0.21

Table 3.3: **Summary Statistics - New Funds' Investment Styles**

	New Funds	Non-Incubated Funds	Incubated Funds
Capital Appreciation	56	41	15
Equity Income	26	23	3
Financial Services	17	13	4
Growth	434	337	97
Growth and Income	131	96	35
Health / Biotechnology	33	26	7
Mid Cap	142	105	37
Micro Cap	18	13	5
Small Cap	239	193	46
Science and Technology	62	53	9
Total	1158	900	258

Table 3.4: **Construction of the Investment Style Book to Market Ratio**

INVESTMENT STYLE	OBJECTIVE	STOCKS
Capital Appreciation	Aggressive Growth Companies	Below the 17th percentile of the book-to-market ratio distribution
Equity Income	High Dividend Yield	Above the 66th percentile of the dividend yield distribution
Financial Services	Finance-related Companies	SIC Codes between 6000 and 6999
Growth	Growth Companies	Below the 33th percentile of the book-to-market ratio distribution
Growth and Income	Companies with high dividend yield and good growth Prospects	Below the 33th percentile of the book-to-market ratio distribution and above the 76th percentile of the dividend yield distribution
Health / Biotechnology	Health and Pharmaceutical Related Companies	SIC Codes 8011 and 8099 or between 2833 and 2836
Mid Cap	Mid-size Companies	Between the 33th and 66th percentile of the market capitalization distribution
Micro Cap	Very small Companies	Below the 17th percentile of the market capitalization distribution
Small Cap	Small Companies	Below the 33th percentile of the market capitalization distribution
Science and Technology	Technology Companies	First three SIC Codes: 357 365 366 367 382 386 381 481 482 484 489 737 or SIC Codes 3844 and 3845

Table 3.5: **Determinants of Number of Funds' IPO's - Poisson Regressions**

Dependent variable is the number of Fund IPO's that follow investment style i at time t. INVESTMENT STYLE B/M is the lagged average book-to-market ratio of the cross-section of stocks that belong to investment style i. INV STYLE FLOW is investment style i's lagged monthly fund flow. Flows are defined as $flow_{i,t} = \log\left(\frac{TNA_{i,t}}{TNA_{i,t-1}}\right) - \log(1 + R_{i,t})$. SP 500 B/M is the lagged S&P 500 index's book-to-market ratio. INV STYLE RET is the lagged value weighted average decile of funds that belong to investment style i. INV STYLE SIZE is the natural logarithm of the assets under management of style i at time t. Robust z-statistics are in brackets.

	FUNDS	$FUNDS_i$	$FUNDS_i$	$FUNDS_i$	NON-INCUBATED$_i$	INCUBATED$_i$
Panel A: 6-Month lag						
INV STYLE B/M		-0.409***		-0.330**	-0.409**	0.051
		[-2.76]		[-2.06]	[-2.33]	[0.15]
INV STYLE FLOW			10.604***	9.430***	11.982***	-1.608
			[4.71]	[3.62]	[4.39]	[-0.23]
INV STYLE RET				-0.027	-0.042**	0.034
				[-1.47]	[-2.05]	[0.92]
INV STYLE SIZE				0.465**	0.672***	-0.416
				[2.33]	[3.14]	[-0.81]
SP 500 B/M	-6.711***			3.657*	3.533*	3.982
	[-7.03]			[1.87]	[1.70]	[0.86]
Panel A: 7-Month lag						
INV STYLE B/M		-0.443***		-0.341*	-0.439**	0.086
		[-2.98]		[-1.95]	[-2.32]	[0.23]
INV STYLE FLOW			11.655***	9.707***	9.783***	10.936*
			[5.26]	[3.90]	[3.60]	[1.80]
INV STYLE RET				-0.021	-0.026	-0.001
				[-1.22]	[-1.42]	[-0.02]
INV STYLE SIZE				0.441**	0.700***	-0.760
				[2.20]	[3.29]	[-1.37]
SP 500 B/M	-6.953***			-1.367	0.178	-7.962
	[-8.52]			[-0.70]	[0.09]	[-1.56]
Panel A: 9-Month lag						
INV STYLE B/M		-0.421***		-0.299*	-0.433**	0.248
		[-2.85]		[-1.74]	[-2.31]	[0.66]
INV STYLE FLOW			11.781***	10.746***	9.948***	15.693***
			[5.46]	[4.45]	[3.96]	[2.98]
INV STYLE RET				-0.037**	-0.041**	-0.024
				[-2.21]	[-2.10]	[-0.68]
INV STYLE SIZE				0.405**	0.675***	-0.904
				[2.16]	[3.47]	[-1.60]
SP 500 B/M	-6.419***			-1.033	0.592	-8.191
	[-7.55]			[-0.55]	[0.30]	[-1.64]
Investment Style Dummies		Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes
Additional Controls	Dollar Flows					

Table 3.6: **Evolution of Funds' Average Book to Market Ratios after IPO**

Dependent variable is the value-weighted average book-to-market of stocks held by fund i at portfolio disclosure time t. AGE is the number of quarters between the current portfolio disclosure date and the IPO date of the fund plus one. INV STYLE BM is the average book-to-market ratio of the cross-section of stocks that belong to investment style i. FLOW is the average monthly net fund flows to fund i between the current and previous portfolio disclosure dates. Flows are defined as in Table 3.5. STYLE FLOW is the average monthly net fund flows to in fund's i investment style between the current and previous portfolio disclosure dates. RET is the average monthly return of fund i between the current and previous portfolio disclosure dates. Robust t-statistics are in brackets.

	ALL FUNDS	ALL FUNDS	ALL FUNDS	ALL FUNDS	NO GROWTH
AGE	0.00279***	0.00256***	0.00411***	0.00288***	0.00308***
	[7.528]	[6.521]	[11.276]	[5.986]	[4.486]
INV STYLE BM		0.04575***	0.04393***	0.02204***	0.02060***
		[19.809]	[8.930]	[4.328]	[3.362]
FLOW		0.00081	-0.00311	-0.00358**	-0.00403
		[0.186]	[-1.548]	[-2.073]	[-1.527]
STYLE FLOW		-1.78229***	-0.89638***	-0.61672***	-0.64183***
		[-11.725]	[-7.709]	[-5.780]	[-5.487]
RETURN		-0.06400*	-0.16177***	-0.10143***	-0.07517***
		[-1.743]	[-8.980]	[-5.392]	[-3.170]
TNA		-0.00834***	-0.02517***	-0.01081***	-0.01416***
		[-11.126]	[-12.877]	[-6.268]	[-5.719]
Constant	0.25961***	0.38686***	0.67203***	0.51202***	0.57804***
Observations	9118	8900	8900	8900	5146
Year.Quarter Dummies	No	No	No	Yes	Yes
Fund Fixed Effects	No	No	Yes	Yes	Yes

Table 3.7: **Average Alpha and Factor Loadings for New Funds**

This table displays the average regression coefficients and risk-adjusted alpha for portfolios of funds. IPO is a value weighted portfolio of funds that IPO during month t. CONTROL is a value weighted portfolio of funds that are below the 20th percentile of the mutual fund size distribution at time t. Panel A tracks portfolios for 24 months after formation time t. Panel B tracks portfolios for 36 months after formation time t. The three Fama–French factors are zero-investment portfolios representing the excess return of the market, Rm-Rf; the difference between a portfolio of "small" stocks and "big" stocks, SMB; and the difference between a portfolio of "high" book-to-market stocks and "low" book-to-market stocks, HML. The fourth factor, UMD, is the difference between a portfolio of stocks with high past one-year returns minus a portfolio of stocks with low past one-year returns. t-statistics are in parentheses.

	ALPHA	MKTRF	SMB	HML	UMD
Panel A: 24 Months after IPO					
IPO	0.0097%	1.020	0.246	-0.019	0.083
	(0.27)				
CONTROL	0.0197%	1.012	0.251	0.007	0.069
	(0.86)				
IPO-CONTROL	-0.0100%	0.009	-0.004	-0.026	0.013
	-(0.30)	(0.53)	-(0.18)	-(0.95)	(1.06)
Panel B: 36 Months after IPO					
IPO	-0.0032%	1.020	0.242	-0.038	0.078
	-(0.12)				
CONTROL	0.0029%	1.015	0.230	-0.003	0.069
	(0.21)				
IPO-CONTROL	-0.0062%	0.005	0.012	-0.035	0.009
	-(0.24)	(0.31)	(0.54)	-(1.32)	(0.82)

Table 3.8: **Expected Flows and Catering**

Dependent variable is the monthly net fund flow to fund i at time t. Flows are defined as in Table 3.5. RET_{t-1} is fund i's monthly net return. STYLE $FLOW_{t-1}$ is the relative monthly fund flows to fund's i investment style. DEMAND is the average monthly STYLE FLOW in the 6 months preceding fund i's IPO. POST is a dummy variable equal to one if t is between the 13th and the 24th month after the IPO of the fund and zero if t is between the 1st and 12th month after the IPO. Robust t-statistics are in brackets. Standard errors are clustered at the fund level

	FLOW	FLOW	$FLOW_{POST=0}$	$FLOW_{POST=1}$	FLOW	FLOW
RET_{t-1}	0.20359***	0.29428***	0.35875***	0.20206***	0.20269***	0.28303***
	[3.036]	[4.869]	[5.049]	[3.440]	[3.054]	[4.738]
STYLE $FLOW_{t-1}$	0.07609***	0.03560	0.04997	0.04703**	0.06398***	0.04645**
	[4.134]	[1.600]	[1.359]	[2.247]	[3.642]	[2.117]
DEMAND	-0.00002	-0.01311	-0.01151	0.01305	0.02047	-0.00154
	[-0.002]	[-0.909]	[-0.470]	[0.766]	[1.508]	[-0.105]
DEMAND X RET_{t-1}		-0.72013***	-0.77473**	-0.48277***		-0.61735**
		[-3.543]	[-2.323]	[-2.856]		[-2.056]
DEMAND X STYLE $FLOW_{t-1}$		0.31328***	0.28620**	-0.17394		0.30393***
		[3.101]	[2.518]	[-0.683]		[3.410]
POST					-0.04452***	-0.04263***
					[-17.433]	[-16.576]
DEMAND X RET_{t-1}X POST						0.04194
						[0.131]
DEMAND X STYLE $FLOW_{t-1}$ X POST						-0.50760**
						[-2.279]
Style Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year.Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27210	27210	13548	13662	27210	27210
R-squared	0.014	0.016	0.019	0.023	0.025	0.027

Table 3.9: **Effect of Fund IPO's on Investment Style Fund Flows**

Dependent variable is the monthly fund flows to style investment *i*. Flows are defined as in Table 3.5. STYLE FLOW is the monthly fund flows to investment style *i*. NEW FUNDS is the number of new funds in investment style *i*. RETURN is the monthly value-weighted average net return of funds that belong to investment style *i*. Robust t-statistics are in brackets.

	STYLE FLOW$_t$	STYLE FLOW$_t$
RETURN$_{t-1}$	0.122***	0.123***
	[5.158]	[5.015]
STYLE FLOW$_{t-2}$	0.256***	0.194***
	[4.851]	[3.609]
NEW FUNDS$_{t-2}$	0.001*	-0.000
	[1.789]	[-0.163]
STYLE FLOW$_{t-2}$ X NEW FUNDS$_{t-2}$		0.086**
		[2.119]
RETURN$_{t-1}$ X NEW FUNDS$_{t-2}$		-0.002
		[-0.494]
Year.Month Dummies	Yes	Yes
Investment Style Dummies	Yes	Yes
Standard Errors	White	White
Observations	1200	1200
R-squared	0.332	0.340