Essays on Fund Families: Ties and Trade Offs

Author: Harold Dean Spilker

Boston College Electronic Thesis or Dissertation, 2017

Boston College

The Carroll Graduate School of Management

Department of Finance

ESSAYS ON FUND FAMILIES: TIES AND TRADE OFFS

a dissertation

by

HAROLD D. SPILKER III

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Finance

April 2017

ESSAYS ON FUND FAMILIES: TIES AND TRADE OFFS

by

HAROLD D. SPILKER III

Advisor: Ronnie Sadka

In the first essay of this dissertation, I study the impact that hedge fund manager connections have on investment ideas. I find that hedge fund managers who previously worked at the same prior hedge fund invest more similarly, hold more overlapping portfolios, and trade and overweight the same stocks relative to managers who do not share an employment connection. Overall, these results support theoretical prediction that networked managers share ideas that leads to price discovery for commonly held stocks.

The second essay analyzes the role of ETFs in mutual fund families and is joint work with Caitlin Dannhauser. We study mutual fund and ETF twins - index funds from the same family that follow the same benchmark. We find that mutual fund twins have lower overall tax burdens while ETF twins have higher long-term yields and unrealized capital gains, but are compensated with lower expense ratios. Fund families benefit because twin offerings generate higher flows than their non-twin peers. These results support previous research that mutual fund families use diversification and subsidization to benefit the overall family.

In the third essay, I study the use of latent factors in explaining hedge fund returns. Using an alternative latent factor estimator, asymptotic principal components (APC), I find explains more of the common variation of hedge fund returns on average and does so with greater efficiency than that found in the literature. I also identify an increase in the common variation across hedge fund excess return in the time-series via the extracted latent factors. My results suggest an impetus for future researchers to employ APC factors when characterizing hedge fund performance.

AKNOWLEDGEMENTS

This dissertation as benefited greatly from the guidance and assistance of my advisor Ronnie Sadka. My committee composed of Pierluigi Balduzzi, Alan Marcus, Jeffrey Pontiff, and Jonathan Reuter, has also been integral to the completion of my dissertation. I also appreciate the lessons and support provided by the remainder of the Finance faculty at Boston College, as well as my Ph.D. colleagues who proved invaluable to my intellectual growth. Lastly, I want to extend my deepest gratitude to the Finance Department at Boston College for supporting me throughout the Ph.D. process, and in particular, for supporting me during my successful battle with stage IV cancer.

I also want to thank my family, who made this and many other achievements possible. Thanks to my parents, Dean and Anita, for providing me the work ethic to succeed; my older sister, Janie, for my alter ego; my little sister, Mandy, for her positivity even in the end; my parents-in-law, Betty and David, for their humility; and my daughter, Finley, for teaching me unconditional love. Last, but always first, thanks to my wife, Kirsten, for agreeing to marry me with eyes wide open, for her inspiring resolve and optimism in the face of herculean obstacles during our journey thus far, and for being my best friend.

Hedge Fund Family Ties

Harold D. Spilker III*

April 13, 2017

Abstract

Social connections are an important determinant for investment decisions. Using a proprietary dataset of hedge fund manager employment backgrounds, I find that hedge fund managers who previously worked at the same hedge fund invest more similarly and hold overlapping portfolios up to 49% more than managers who do not share an employment connection. Furthermore, related managers contemporaneously herd more and over-weight the same stocks versus unconnected managers, providing evidence for the social exchange of investment ideas within a network. A long/short portfolio of overlapped-connected/unique-unconnected stocks generates alpha of 4.5% per year and permanently conveys private information to asset prices. These findings support the "quid-pro-quo" model of Stein (2008) and confirms that shared employment histories signal increased ex-ante correlations between connected portfolios.

*Special thanks to my dissertation committee (Pierluigi Balduzzi, Alan Marcus, Jeffrey Pontiff, Jonathan Reuter, and Ronnie Sadka (chair)) for their assistance on this article–the first essay in my doctoral dissertation. I am also grateful to Caitlin Dannhauser, Slava Fos, Matthew Osborn, Christopher Schwarz (discussant), Phil Strahan, and Jerome Taillard for useful comments. Thanks also to participants at the 2015 Ph.D. Conference in Finance at the University of Southern California, and at seminars at Boston College, Florida International University, Texas Tech University, the University of Hawaiʻi at Mānoa, and the U.S. Securities and Exchange Commission's Department of Economic and Risk Analysis. Spilker is from the Carroll School of Management, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467; Contact: tray.spilker@bc.edu or www2.bc.edu/tray-spilker.

"It is not the consciousness of men that determines their being, but, on the contrary, their social being that determines their consciousness."

– Karl Marx, A Contribution to the Critique of Political Economy (1859)

Introduction

Given the human condition is an inherently social one, as Marx points out, individuality is difficult to divorce from the social fabric from which it is collectively shaped. This sentiment is found in the English literature as early as the 17th century, when John Donne penned that "No man is an island entire of itself."[1] In more recent history, the popular press has noted the primacy of social networks over individual intelligence, coining the common saying – "It's not what you know; it's who you know."[2] Given the recent rise of social networking platforms, the market has ensconced social valuation in asset prices today. The combined market capitalizations of the three largest social networking platforms Facebook, Twitter, and Snapchat total $426 billion, or 2.6% of U.S. GDP, as of this writing.[3]

The financial literature identifies the economic consequence of social networks as well. Hong, Kubik, and Stein (2005) are among the first to establish word-of-mouth effects among mutual fund managers. The authors find that managers working in the same city buy more of the same stocks when their co-located peers do so. Pool, Stoffman, and Yonker (2015) find that mutual fund managers who live in the same neighborhood invest in the same stock more often than managers who are not neighbors. Hochberg, Ljungqvist, and Lu (2007) find more highly networked venture capitalists outperform lesser-networked competitors by virtue of their superior ability to exit investments profitably. Furthermore, Cohen, Frazzini, and Malloy (2008) show that mutual fund managers generate outsized returns when trading connected stocks, where the connection is between the manager and a corporate board member who attended the same university.

In this study, I explore whether connected hedge fund managers – those who used to work for the same prior hedge fund – invest more similarly relative to managers who worked at differing

[1] Meditation 17, John Donne, Devotions upon Emergent Occasions XVII,1623

[2] 22 September, 1918, New York (NY) Tribune, "U. S. to Act to Oust Ship Work Slackers," pg. 9, col. 1

[3] LinkedIn was included in the original version of this paper. The final valuation of LinkedIn was roughly $26 billion prior to its acquisition by Microsoft in December 2016. See: https://www.wsj.com/articles/microsoft-closes-acquisition-of-linkedin-1481215151

hedge funds; i.e., unconnected managers. I bring to the literature a proprietary dataset of over 500 hedge fund manager employment histories. These hedge fund managers hail from the largest networks of hedge fund managers in the industry, thus providing many testable connections for statistical inference. I obtain quarterly portfolio holdings for each hedge fund using form 13F that is made available by the U.S. Securities and Exchange Commission (SEC). Unfortunately, 13F covered securities include only long positions in public equities. Hence, the analysis is confined to the long side of a hedge fund's portfolio; a common constraint in the hedge fund literature using 13F data.

In my first test, I use the pair-wise fund-level portfolio overlap measure of Pool et al. (2015) to measure quarterly portfolio overlap between each pair of hedge fund managers in my sample. I find portfolio overlap is 29.7% higher among connected hedge fund managers who share a common employment history. Since my identification relies on 13F disclosures, it follows that estimating portfolio overlap between managers pursuing strategies that are not well identified through long equity positions merely adds noise. Hence, when I constrain the test only to long/short equity hedge funds, portfolio overlap increases to 48.5% among connected managers. This baseline result shows that network connections are an important determinant for hedge fund manager investment decisions.

In addition to the social network channel, overlapping portfolios can also obtain from similar preference sets obtained from working for the same prior hedge fund manager. The manager origin literature identifies a learning channel that leads to manager fixed-effects. Bertrand and Schoar (2003) find corporate managers gain valuable portable skills from former workplaces that can be traced to the manager as they impact corporate actions in subsequent executive roles. Papageorgiou, Parwada, and Tan (2011) show that hedge fund managers with prior experience at hedge funds outperform those managers with training at other types of firms. Subsequently, hedge fund managers of the same background are likely to have learned to follow similar investment approaches.

In effort to disentangle the channels that lead to overlapping portfolios, I construct treatment and control groups to conduct stock-level tests in a differences-in-differences framework. The treatment group is formed by those managers who previously worked for one of the largest and most influential hedge funds in the industry's history; Tiger Management. Hedge fund managers belonging to this treatment group are known colloquially as "Tiger Cubs" and represent a sizeable fraction of the industry with over $250 billion in combined assets; roughly 13% of total industry assets in 2011 (Mallaby (2011)). I find the treatment group of managers concentrates more portfolio risk in fewer positions on average, and more actively trade their respective portfolios relative to the remaining sample of hedge funds.

These two investment decisions – portfolio construction and turnover – are shown to be important determinants of performance in the literature. Kacperczyk, Sialm, and Zheng (2005) find that managers who hold concentrated portfolios in industries where they possess informational advantages outperform more diversified portfolios. Pastor, Stambaugh, and Taylor (Forthcoming) show that hedge fund managers trade more in the presence of alpha producing trades, thus higher turnover generates superior excess returns. Hence, I posit that portfolio construction and turnover decisions are an important determinant in fund performance that arise from adopting techniques of a formative mentor and is shared across a network of connected hedge fund managers.

I create a control group by matching the treatment group on these learned characteristics that are observable in form 13F; average number of quarterly holdings, average position weighting, and turnover. In my first stock-level analysis, I find that connected managers conditionally herd more into stocks (Wermers (1999)), and over-weight these overlapped stocks by 12.4% relative to unconnected hedge funds. This finding is incremental to the 13% gain due to the word-of-mouth effect found by Hong et al. (2005), as I also match on the city where funds are headquartered.

Presumably, connected managers who share valuable investment ideas do so with an expectation of receiving credible investment leads from their network in the future. Stein (2008) models this "quid pro quo" behavior among fund managers, whom are in direct competition for fund flows, as an incentive-compatible equilibrium where the cost of lying in the model is high for the agent who does so, and where bad ideas never propagate beyond the incipient discussion stage. Supporting this, I find overlapped positions drive risk-adjusted return differences between the treatment and control groups. The treatment group derives 25% more of its alpha from commonly held positions in the network relative to the control group.

Importantly, this confirms information transfer to security prices, which Pool et al. (2015) also find among shared positions of mutual fund managers who are neighbors. The authors argue the impact on security prices is not merely endogeneity resulting from shared preference sets leading to similar trading conclusions as modeled by Froot, Scharfstein, and Stein (1992), but rather is indicative of valuable information transmitting through a network of neighboring fund managers. My findings support a similar hypothesis that a social network among hedge fund managers who share a common employment history is active and conveying information to security prices. It is hard to imagine a scenario where overlapped positions persist over the sample as a result of shared fixed-effects, particularly in the presence of contemporaneous herding and over-weighting behavior.

I buttress the network channel conclusion with results on cumulative average abnormal returns as used in Coval and Stafford (2007). I show that abnormal returns to stocks held by connected funds are higher than those held by unconnected funds. More importantly, this difference of 15% is statistically significant and is not mean reverting, indicating that information transmitted via connected managers carries private information to asset prices. If the shock were transitory, the alternative hypothesis of no information – where the network is merely exerting price pressure akin to a "pump and dump" strategy – would be more likely.

A counterfactual where lesser informed funds within the network follow the holdings data of fellow funds they deem better informed is possible (Bikhchandani, Hirshleifer, and Welch (1992)). And hedge fund managers may also follow certain connected members' investing decisions intentionally to ensure they suffer no reputational degradation that may result from under-performing funds in the same network (Scharfstein and Stein (1990)). However, my results on herding behavior (Wermers (1999)), show that the treatment group conditionally herds more into and out of stocks relative to the control group. Thus, mimicking behavior within a network is likely a second order effect, giving primacy to the network channel.

In a contemporaneous paper, Gerritzen, Jackwerth, and Plazzi (2016) use a similar identification of prior employment histories to show that connected hedge fund managers have more similar risk exposures relative to unconnected managers. Thus, our findings are mutually supportive. My study reaches further, however, to identify *how* similar risk exposures manifest across connected managers. As such, I identify that the channel for performance similarities arises not only from shared investment behaviors via a learning channel,[4] but also through shared investment positions conveyed through a lively social network, which I attempt to disentangle. Furthermore, I show these connections have an impact on the cross-section of hedge fund returns, drive abnormal performance within a connected network of funds, and convey information to asset prices.

Lastly, my results contribute to the fund family literature. Brown, Fraser, and Liang (2008) note that the cost of initial due diligence on a prospective hedge fund manager exceeds $1 million and

[4]In this study, I use the term 'learning channel' to describe the acquisition mechanism that leads to manager fixed-effects found in Bertrand and Schoar (2003).

400 man hours. Further, Anson (2006) recommends spending 75 - 100 hours reviewing a hedge fund manager before investing. Thus, leveraging fund family structures is one avenue to reduce search related expenses. But, unlike mutual funds, managers who depart a brand name hedge fund often establish independent funds with no formal ties to their former employer. Several factors are likely to contribute to this difference. First, the "2+20" incentive structure of the hedge fund industry leads many portfolio managers within reputable hedge fund firms to strike out independently to capitalize on the cache of their training. Second, relative to the administrative and financial burden that would accompany a similar mutual fund launch, these managers face a lower regulatory burden when establishing a hedge fund management company. Therefore, mutual fund starts typically occur within families as a larger fund family can more easily handle this burden.

Unfortunately, investors who rely on family affiliation to source new funds face an inherent problem among independent hedge fund managers (Massa (2003), Gervais, Lynch, and Musto (2005)), as search becomes more costly when fund family affiliation is not readily apparent (Sirri and Tufano (1998)). However, when marketing their funds, hedge fund managers necessarily disclose their roles at prior employers as part of the due diligence process; with an expectation that their affiliation will offer a signal about their skill. Hence, by identifying connections between hedge fund managers, an informal fund family structure materializes, thereby providing a credible signal to reduce search costs. Additionally, given my results on connected managers, investors can use this "family" signal to more efficiently diversify portfolios of hedge funds by identifying hedge funds managers ex-ante whose performance is likely to be highly correlated (Elton, Gruber, and Green (2007)).

This paper is organized as follows. Section 1 discusses construction of the dataset and identifies the treatment group. Section 2 describes the methodological setting, including the matching process used to build the control group. Section 3 presents evidence that connected hedge funds have a higher percentage of overlapping portfolios. I attempt to dissentangle the channels that lead to overlapped portfolios in Section 4, by analyzing conditional herding tendencies and allocation decisions to commonly held positions. Section 5 provides evidence that these common positions drive abnormal returns, suggesting that information flows through a connected network and impacts asset prices, which I confirm later in the section with asset pricing implications. I conclude in Section 6.

1 Data and Sample Construction

I utilize a novel proprietary dataset of just over 500 hedge fund manager employment histories gleaned from the largest institutional investors and investment advisors in the U.S. As such, this selection of funds may produce differentiated returns and have more onerous lockup provisions relative to an unbiased sample, as these hedge funds are considered "best of breed" in the industry.[5] This list is cross-checked with internet searches and news wires to validate data quality. Reference sights include LinkedIn, Bloomberg, Insider Monkey, HF Alert, and MarketFolley, among others.

To this dataset, I hand match hedge funds to regulatory filings administered by the SEC, firstly by matching the name of the fund, then by the first filing date which corresponds most closely to the fund's inception date. From this, I obtain quarterly hedge fund portfolio holdings as reported in form 13F and made available in the Thomson Reuters Institutional Holdings Database (s34 master file).[6] Notably, 13F filings only identify long positions held in a given quarter-end portfolio, which truncates observations on the short side of a hedge fund's portfolio. Thus, my analysis is restricted to observed holdings within the long portfolios of my sample of hedge funds.

As identified in Aragon and Martin (2012), 13F filings contain information at the advisor level such that for larger asset management firms, holdings based data may incorporate multiple funds in the fund family, which may include portfolios other than hedge funds. Thus, I screen my sample to those funds whose line of business consists solely of hedge fund strategies (e.g., I exclude the likes of Merrill Lynch, GAMCO, etc.). Further, since I am using 13F holdings information to identify holding patterns across hedge fund networks, I restrict my sample to long/short hedge funds as this strategy is best identified by the information contained in 13F forms, resulting in 296 hedge funds in the base sample.

Thomson Reuters uses two date fields. RDATE (reporting date) represents the end of quarter date for which holdings information is valid, and FDATE, which is a vintage field created by

[5] Aragon (2007) shows that funds with lockup restrictions earn excess returns of 4-7% relative to funds without lockup clauses.

[6] Enacted in 1975, Section 13(f) of the Securities Exchange Act of 1934 requires institutional investment managers who manage over $100 million to file 13F forms within 45 days after the end of each quarter. These filings contain information on the fund's investment holdings at the end of the respective quarter that are defined as 13F securities. These include public equities, closed-end funds, exchange traded funds (ETFs), certain equity options and warrants, as well as certain convertible bonds.

Thomson Reuters to ensure continuity of holdings information when a fund is late in reporting to the SEC (i.e., RDATE is carried forward). Thomson Reuters notes that a "slight majority" of RDATEs coincide with the same FDATE for mutual funds found in the s12 master file, but that RDATE and FDATE are the same in a "large majority" of investment companies in the s34 master file. None the less, I follow standard practice and use only observations where RDATE is equal to FDATE to restrict stale data from my sample.[7]

Among the sample, one hedge fund employer stands out for generating the largest number of independent long/short hedge funds; Tiger Management. From its launch in 1980 to its effective closure in 2000,[8] Julian Robertson's Tiger Management generated annualized returns of roughly 19% and grew to $22 billion in assets, which was second only to George Soros' Quantum fund at the time. Over its history, Tiger Management spun out myriad portfolio managers who subsequently established their own competing hedge funds. In sum, I've identified some 69 "Tiger Cubs," who themselves sired 30 "Grand Cubs" and 4 "Great Grand Cubs," the sum of which oversee more than $250 billion in assets, or roughly 13% in total hedge fund assets in 2011(Mallaby (2011)). Due to the size of the connected funds in this network, I use them as the treatment group in my study.

Unfortunately, only 55 members of the treatment family have a history of 13F filings, which is further reduced to 46 after limiting the analysis to long/short hedge funds with a U.S. geographic focus. Lastly, I include only those funds with at least eight quarters worth of holdings of data resulting in a treatment group of 42 funds. The 13F filing requirement pertains to covered securities in the U.S., therefore hedge funds who invest mainly outside the U.S. will not be well identified by 13F holdings. Separating my treatment group from the remaining sample, and filtering the remaining sample by the above criterion results in 232 funds from which to select a control group. In sum, the data for this study covers quarterly observations from January, 2000 to December, 2013, containing 156,494 stock-quarter observations across 274 distinct hedge funds.

---

[7] See the Thomson Reuters User Guide found on the Wharton Research Data Service website: https://wrds-web.wharton.upenn.edu

[8] Julian Robertson closed the fund to outside capital and returned funds to investors after substantial losses. Tiger Management remains a family office known for its hedge fund seeding program.

2 Methodology

Allocation decisions in active fund management explain a large proportion of resultant performance (Markowitz (1959) and Sharpe (1970)). These decisions are typically embedded in portfolio construction and risk management techniques that govern what proportion of assets a manager allocates to an investment idea, an industry, net- and gross-market exposures, and hedging strategies, among others. Portfolio construction, specifically, dictates how much investment risk is concentrated in any single investment, across any groupings of positions (e.g., the total weightings of the largest ten investments, or to any single industry), how many total investments are held in the portfolio in a given cross-section, and upper bounds on weightings for the highest conviction investments. The discipline of portfolio construction balances the manager's desire to generate excess returns from over-weighting her best investment ideas, against putting the firm's survival at risk if the investment thesis fails.

To the extent that correlated portfolio construction behavior among connected hedge funds manifests, a logical channel to explain this phenomena may be shared influence from a formative mentor. Influence from other channels is unlikely to drive such a critical investment management decision, particularly as the heterogeneity of management styles does not predict which portfolio characteristics may follow. Hence, I argue that similarities in portfolio construction characterize a learned skill shared across connected managers, and thus I construct my control group by matching on these characteristics.

[**Table 1 here**]

Table 1 presents summary portfolio characteristics for the treatment group and the remaining unmatched sample. A discernible difference quickly emerges between the treatment and unmatched sample. The treatment family holds decidedly fewer positions during an average quarter, 39 versus the sample's 95. Further, the average number of positions is remarkably consistent among funds in the treatment group as shown by the standard deviation of 23 positions, which stands in contrast to the distribution of average number of holdings for the unmatched sample of 165 positions. Figure 1 shows the distribution of quarterly holdings for both the treatment and unmatched sample. The treatment group clearly holds fewer positions on average, resulting in more concentrated portfolios and larger allocations to each holding on average. The difference in weights assigned to the average position between the two groups is stark; the treatment group allocates 2.5 times more to an average position relative to the treatment group, 2.58% vs. 1.05%.[9]

[Figure 1 here]

Interestingly, the average holding size of investments in their predecessor's portfolio, Tiger Management, is smaller. Thus, it would be hard to conclude these funds transported a learned skill. Julian Robertson ran a much more diversified portfolio with average position size of around 1% of assets.[10] Yet, the summary statistics on the long/short treatment group indicate that these connected managers share a proclivity to run concentrated equity portfolios on the long side. This concentration of risk, combined with the fact that connected funds show a more narrow dispersion in average position size, suggests treatment funds employ similar investment techniques which likely originate from the learning channel.

2.1 Control Group Construction

Following my baseline specification, the remainder of tests in this paper rely on comparing position weights between the treatment group and a yet to be established control group. I thus attempt to control for the learning channel mechanism in effort to tease out ongoing social connections by identifying a matched control group based on similar portfolio construction. I perform a two nearest-neighbor match (without replacement) firstly on fund start date, then on fund longevity, average number of quarterly holdings, average position weighting, average turnover, and headquarter location. In this setting, a fund's start date is the first date for which a 13F report exists. Matching on this measure ensures both the treatment and control groups encountered similar market conditions at the incipient stage of their respective life cycles. Fund longevity is measured as the difference from the first 13F filing date to the last, and is used to ensure that I have a similar number of funds in the matched control group through the time-series. Since the unit of analysis relies on position level detail and portfolio construction behavior to define characteristics of hedge fund networks, I match on the average number of securities and average position weighting calculated from a fund's respective quarterly 13F filing history. Lastly, matching on location adjusts for the Hong et al. (2005) finding that fund managers in the same city exhibit word-of-mouth effects by trading the same stocks in a given quarter when their co-located peers do so.

[9] Here, I define position weighting as fund j's dollar allocation to stock i at the end of quarter t, divided by the sum of all dollar allocations by fund j for that quarter.

[10] Thanks to Christopher Schwarz (discussant) for identifying this contradiction

The final panel of Table 1 contains summary information for the matched control group. As shown, the average number of 39 quarterly holdings is the same across the two groups, which results in average position weights of 2.6%. Matching on average life results in fund longevity of 32 quarters for each group. Notably, a wedge appears between the long assets invested between treatment and control groups; $1.56 billion versus $778 million. While it is a common identification strategy to include portfolio size (or assets under management, AUM) as a matching criterion for control group construction, I assert that deviations in AUM between connected managers is not a learned trait from the perspective of a manager fixed-effect. Just as portfolio management techniques is to alpha production, asset raising strategies are not guaranteed to result in flows. And since AUM is an outcome variable where the input is latent (ie., the manager's marketing strategy), it would not be appropriate to use as a matching variable in this setting.[11] Furthermore, larger asset flows enjoyed by connected managers in the treatment group likely arise from the variation that we are attempting to measure. For example, it's possible that as a result of their known pedigree–and hence, network connections–higher flows follow. This is an important characteristic because as assets accrue to the average hedge fund, typically the manager is compelled to reduce portfolio concentration and increase the average number of securities held in the portfolio. This suggests that connected managers in the treatment group might be more able to hew to their initial investment strategies of running more concentrated portfolios.

3 Network Connections and Portfolio Overlap

Do connected hedge funds invest in a more coordinated fashion on average? Further still, do they invest with more conviction when fund managers of the same ancestry invest similarly? A positive response to either would indicate a tie between connected funds, but a positive finding for the latter would suggest that connected funds share investment ideas through an active network.

In this section, I explore the first question in my baseline test; whether connected managers co-invest more often relative to unconnected managers. I use the *PortOverlap* measure offered by Pool et al. (2015) in my main analysis, with the notation adjusted slightly. The authors measure

[11]Matching on portfolio size reduces the close fit between the treatment and control group in terms of obervable control variables stemming from portfolio construction strategies.

portfolio overlap as a pairwise connection between funds j and k during quarter t as

$$PortOverlap_{j,k,t} = \sum_{i \in \mathcal{H}_t} min\{w_{i,j,t}, w_{i,k,t}\}. \tag{1}$$

where $w_{i,j,t}$ is fund j's portfolio weight in stock i during quarter t, $w_{i,k,t}$ is the same for fund k, and $\mathcal{H}_t$ is the set of all stocks held by funds j and k as reported at the end of quarter t. For example, if during quarter t fund j has 5% allocated to stock i, and fund k has 10% allocated to the same stock i, then *PortOverlap* is equal to 5% for this fund-pair during quarter t. Using this measure as my dependent variable, I estimate the following specification,

$$PortOverlap_{j,k,t} = \alpha + \beta SameNetwork_{j,k,t} + \delta SameCity_{j,k,t} + \Gamma' Controls_{j,k,t} + \varepsilon_{j,k,t}, \tag{2}$$

where, $SameNetwork_{j,k,t}$ is a dummy variable equal to one if fund j and fund k belong to the same hedge fund network, $SameCity_{j,k,t}$ is a dummy variable that is one if fund j and fund k are headquartered in the same city, and $Controls_{j,k,t}$ is a vector of control variables. As controls, I include (a) a set of dummy variables that are equal to one if funds j and k pursue the same hedge fund strategy (*Activist, Equity Long/Short, Equity Market Neutral, Event Driven, Macro, and Multi-Strategy*); (b) the AUM-based quintiles of funds j and k (*AUMQuintAvg*); and (c) the absolute value in the differences between AUM-based quintiles of funds j and k (*AUMQuintDiff*).[12]

Table 2 shows estimates and standard errors for various forms of the regression detailed in equation 3. Given the model is performed in a pairwise fashion, the same fund is present across myriad pairings, giving rise to a lack of independence across observations. Hence, standard errors are two-way clustered for each fund in the fund-pair.[13]. I provide results for this specification in two panels. The first includes results across all strategies, and the second reports estimates for only equity long/short funds.

In the first test (Column 1), I estimate whether the finding of Hong et al. (2005)–where managers exposed to the same media market by operating in the same city–holds among hedge fund managers. Interestingly, I find that it does not hold in this setting; hedge fund managers in my sample operating in the same city exhibit no differential propensity to hold overlapping portfolios. In the second test, I explore whether a manager's network connections influences their investment decisions. If so, then the loading on β will be positive,which I find. The significant

[12]Assets under management (AUM) in this setting are the sum of all long positions disclosed in each fund's quarterly 13F filing.

[13]This follows from Pool et al. (2015)

coefficient for *SameNetwork* of 1.75% implies that on average, connected managers have a 31.6% higher rate of overlapping portfolios than unconnected managers; moving from the average minimum portfolio overlap of 5.54% across all funds to 7.29% (= 1.75% + 5.54%). Strategy controls are included in Column 2 to ensure we account for investment strategies that may endogenously lead to overlapping investments. For example, hedge fund managers pursuing merger arbitrage (which is included under the event-driven strategy) will likely invest similarly in reaction to acquisition announcements regardless of common skill sets nor network affiliation. The same can be said of convertible arbitrage strategies (also covered by event-driven strategies) that are typically dependent on new security supply from corporate issuances.

[**Table 2 here**]

Following the example of Pool et al. (2015), I calculate *AUMQuintAvg* as the average of fund size quintiles across funds j and k, which controls for the presumption that increases in fund-pair sizes results in a higher probability that an overlap can occur. Yet, this average size control cannot account for subtle differences *between* the fund-pair sizes. Consider for example two fund-pairs where in the first, funds j and k are of size quintile 5 and 1, respectively, resulting in an average size quintile of 3. Meanwhile in the second pairing, funds j and k are both in size quintile 3, resulting in the same average size quintile of 3. Comparing the average size quintiles in this example juxtaposes two fund-pairs with the same *AUMQuintAvg*, making it hard to determine whether a positive relationship in average AUM size among fund-pairs matters for portfolio overlaps, or whether the distance between AUM size for funds j and k is important. Hence, *AUMQuintDiff* accounts for this distance. A negative estimate for γ would indicate that as the difference in assets between funds j and k increases their portfolio overlap would decrease, which is exactly what I find. The estimate for *AUMQuintAvg* and *AUMQuintDiff* is 187 bps and -74 bps, respectively. Therefore, on average portfolio overlap increases by 113 bps (32.5% increase) after controlling for size. Notably, size controls explain almost half of the average portfolio overlap, with the constant coefficient, α decreasing by 37.2% to 3.48% in the final specification.

In the final specification, I include controls to account for variation in size across and between fund-pairs. In contrast to Pool et al. (2015) who find a higher overlap among fund-pairs following the same strategy, I find negative or insignificant loadings for all strategy controls. This finding suggests that hedge fund managers do indeed have more diversified portfolios within their respective strategies and relative to mutual funds. However, making cross-sectional inferences for strategies that are not well identified by 13F holdings information is problematic. For example, holdings data for the fixed income strategy is not likely to be informative as a majority of portfolio investments within this strategy are not covered securities required to be disclosed in 13F filings. Hence, 13F studies in the hedge fund industry are best suited to explain the long book of equity oriented strategies such as equity long/short.

Including all controls (Column 3), I find that portfolio overlap is 29.7% among hedge fund managers whom are connected through the same prior employer. The coefficient for *SameNetwork* 1.37% is incremental to the average fund-pair overlap of 4.61% (= 3.48% + 1.87% - 0.74%) whom do not share this connection, and after adding our control variables. Results for equity long/short funds is displayed in the second panel of Table 2. Here, I show that these funds have a *SameNetwork* estimate of 1.82%, which implies that portfolio overlap among connected long/short managers increases to 48.5%; from 3.75% (= 2.94% + 1.34% - 0.53%) to 5.57%. This magnitude is surprising as it implies that almost half of a connected manager's portfolio is common to portfolios within the network.

4 Social Influence on Holdings

In this section, I begin disentangling whether overlapping portfolios result from using an investment strategy learned from a common prior employer, or from idea sharing through connected networks. Whereas in the baseline specification I found that connected hedge fund manager hold overlapping portfolios more on average, here, I analyze contemporaneous stock-level activity among connected funds using the difference-in-difference framework. By controlling for manager fixed-effects such as portfolio construction and trading techniques that are likely learned from a formative mentor, I can highlight idea sharing that results from network effects.

4.1 Herding

The literature on herding tests whether excess trading, deemed "herding," occurs in a particular stock-quarter relative to expected trading activity when no herding occurs. If connected funds have more overlapped portfolios on average relative to the control group, then presumably co-investments accumulate through contemporaneous trading activity. To test whether the funds in my sample herd, I calculate herding measures for the full hedge fund sample, as well as for the treatment and control subgroups using the specification set forth in Lakonishok, Shleifer, and Vishny (1992). Specifically, the herding measure, HM, as expressed in Wermers (1999) is,

$$HM_{i,t} = |p_{i,t} - E[p_{i,t}]| - E\,|p_{i,t} - E[p_{i,t}]|\,, \tag{3}$$

where, $p_{i,t}$ is the proportion of funds buying stock i in quarter t scaled by all funds trading during the quarter. $E[p_{i,t}]$ is the expected proportion of purchases during the quarter, and is proxied by the proportion of all trades that are purchases over a given quarter. Similarly, the second term, $E\,|p_{i,t} - E[p_{i,t}]|$, serves as an adjustment factor to allow for random variation around the proportion of buyers when, in expectation, no herding arises in the full sample. The adjustment factor is calculated with respect to its subgroup (treatment or control groups), and I obtain the mean herding measure, $\overline{HM}$, by averaging across the subgroup of hedge fund managers.

Overall, I find no herding among hedge fund managers in my sample (Table 3). As pointed out in Lakonishok et al. (1999), this should not necessarily be surprising due to the market's need for each buyer to have a seller for clearing. The lack of average herding also holds for the treatment and control subgroups. However, in order to determine whether ideas are transmitted via a lively social network, the difference-in-difference comparison of conditional herding between treatment and control groups is of more interest. That is, we want to compare differences in the buy-herd measure, $\overline{BHM}$, and the the sell-herd measure, $\overline{SHM}$. Specifically, $\overline{BHM}$ is the average of the herding measure, $\overline{HM}$, conditional on $p_{it} > E[p_{it}]$, and conversely, $\overline{SHM}$ is the conditional average of $\overline{HM}$ when $p_{it} < E[p_{it}]$. These two measures identify the tendency of funds to herd while entering a trade (buys) or while exiting a trade (sells). Additionally, I also test whether a stronger herd obtains, calculating $\overline{HM}$ over instances in which successively more funds trade the stock-quarter (as suggested by Wermers (1999)).

[Table 3 here]

As summarized in Panel D of Table 3, the treatment group conditionally herds more than the control group of funds when entering and exiting a trade. For example, when three or more funds enter (exit) a position, connected managers herd more than unconnected managers by 0.55% (3.91%). Furthermore, the differences in conditional herding between the two groups increases as the number of funds herding into or out of the same stock-quarter increases. This is particularly true for funds entering a position, where the difference in conditional means for $\overline{BHM}$ increases monotonically as the number of funds who buy-herd during the stock-quarter increases from 3 to 20 funds. This indicates that on average, conditional herding is more pronounced among the treatment group where connected members not only engage in more common trades, but do so in relatively larger herds.

4.2 Differential Allocation to Overlapped Stocks

To further disentangle whether overlapping portfolios result from using an investment strategy learned from a shared employer or from sharing investment ideas through the connected network, I test the hypothesis that over-weighting commonly held positions does not happen randomly. That is, if fund managers in a connected network consistently herd into and over-weight commonly held positions, they likely do so in response to differential information. I posit these common investment ideas are transmitted through a connected network, which not only inform the manager's decision about which investment to make, but by how much additional capital to allocation to the position.

To test this hypothesis, I construct a measure for adjusted relative weighting, *AdjustedWeight*, to identify the sensitivity to position allocations between the treatment and control groups. Specifically, I measure *Adjusted Weight* as

$$AdjustedWeight_{i,j,t} = \left(\frac{\$Allocation_{i,j,t}}{\sum_{i=1}^{N} \$Allocation_{i,j,t}} \right) - \left(\frac{MktCap_{i,j,t}}{\sum_{i=1}^{N} MktCap_{i,j,t}} \right), \tag{4}$$

where the first term represents fund j's weighting to stock i in quarter t, and the second term adjusts this raw position weight by the market capitalization of stock i relative to the sum of market capitalizations of all stocks held in fund j at the end of each quarter t. I impose this adjustment to account for the increased likelihood that subgroups of funds will likely experience more stock-quarter overlaps among large capitalization stocks (as noted in Wermers (1999)). Notably, this measure moves the analysis from the fund-level, as measured in my baseline test, to the stock-level. I do this in effort to obtain more granularity in the investment decision between connected managers, and as a robustness test to the asset-weighted portfolio overlap results in the main specification.

To measure stock level weighting sensitivity between the treatment and control groups, I estimate the following fixed-effects model:

$$\begin{aligned} AdjustedWeight_{i,j,t} &= \alpha + \beta Overlap_{i,j,t} + \delta Treatment_j \\ &\quad + \gamma (Overlap_{i,j,t} * Treatment_j) + \tau_t + \varepsilon_{i,j,t}. \end{aligned} \tag{5}$$

Here, *Overlap* is dummy variable equal to one for quarters in which two or more funds of the respective subgroups hold the same stock-quarter. A positive β on *Overlap* indicates that hedge funds over-weight stock i during quarter t when other funds also hold the same stock. *Treatment* is a dummy variable equal to one for connected funds in the treatment network and is zero for all funds in the control group. Given that I constructed the control group by matching on the average number of quarterly positions (among other controls), we should expect to see no significance on this coefficient. Lastly, I include the interaction effect of being a member of the treatment group and holding a common stock-quarter, *Overlap*Treatment*. A significant coefficient for this regressor, γ, would indicate that members of the same network over-weight positions conditional on whether other connected members also hold the position, thereby suggesting an active social network among connected fund managers.

Findings for this test are shown in Table 4, where the last two columns highlights results adjusted for common cross-sectional shocks using quarter fixed effects. Focusing on the fully adjusted model in column 4, I find that conditional on holding the same stock, fund managers allocate 66 bps more to the position, a 29.5% increase over the average adjusted weight of 2.24% indicated by the constant term. This suggests that fund managers increase allocations to commonly held positions, but cannot explain why they act in tandem. For example, its possible that managers find comfort in numbers and over-weight a commonly held position because other funds also hold the stock. It's also possible that fund managers pursuing the same investment style reach independent conclusions that result in over-weightings on the same stocks. As expected, the coefficient for the treatment dummy is not differentiable from zero. This result also indicates that the position adjustment factor (right hand term of equation 2) is similar between the two groups, suggesting both groups allocate similar proportions across relative market capitalizations.

[**Table 4 here**]

The interaction coefficient, γ, of the treatment and overlap dummy variables can help control for common investment models within the same investment style (equity long/short in this case). The positive coefficient on γ indicates that conditional on a stock-quarter overlap, connected fund managers allocate 36 bps more to the position, which is a 12.4% increase in position weight relative to funds who do not share a similar connection. Having already constructed the control group by also matching on location, this represents an incremental response to peer allocations among hedge fund managers who work in the same media market (Hong et al. (2005)).

5 Asset Pricing Impact from Manager Networks

5.1 Risk-Adjusted Portfolio Returns

Thus far, I have shown that funds of a common ancestry have more correlated holding risk relative to those funds that do not share a history. So, why do connected funds co-invest at higher rates? Presumably, they do so in expectation that larger payoffs will follow. As discussed in the "quid-pro-quo" model of Stein (2008), the author theorizes that investment ideas will be exchanged among socially connected fund managers–whom are in direct competition for flows–under the expectation that investment ideas will be reciprocated in the future. Furthermore, the author posits that only good investment ideas will make it past the incipient discussion phase, given that socially connected managers will push back on poor investment ideas received from a trusted peer.

To examine whether performance is differentiatied for overlapped stock-quarter pairs, I compare risk-adjusted returns for the treatment and control groups using a difference-in-difference framework. As I do not have access to fund returns in my sample, I proxy for monthly returns using the long only holdings reported in 13F filings. Returns are constructed from aggregating weighted returns to all stocks held by fund j at the end of quarter t, multiplied by returns to stock i over the ensuing three months, $k = 1, 2, 3$; $R_{j,t+k} = \sum_{i=1}^{N} w_{i,j,t} r_{i,t+k}$. From this constructed time-series of returns, I subtract the risk-free rate and use these excess returns as the dependent variable. Since I construct fund returns from long-only positions, I select my factors from the mutual fund literature and run fund-level panel-regressions on the three-factor model of Fama and French (1993), as well as the momentum factor of Carhart (1997) and the liquidity factor of Pástor and Stambaugh (2003).[14]

Table 5 shows statistically significant excess monthly returns of roughly 40 basis points for the treatment and control groups across all specifications, along with strong positive responses to changes in the market factor. Interestingly, the literature to date has shown no positive excess returns on hedge fund portfolios constructed from 13F holdings information.[15] As noted in the

[14]Thanks to Ken French and Robert Stambaugh for providing these factors on their respective websites. I do not use the Fung and Hsieh (2004) four-factor model for hedge fund returns here given that the return series is constructed from long-only positions and does not exhibit option-like behavior found in hedge fund returns.

[15]See for example Agarwal, Fos, and Jiang (2013), Brown and Schwarz (2013), Griffin and Xu (2009), and Aragon and Martin (2012).

introduction, this dataset is compiled from proprietary sources whom have access to "best of breed" hedge fund managers. Hence, the sample likely has an upward bias in performance, for both treatment and control groups, which is confirmed with statistically significant excess return in Table 5. [16]

[Table 5 here]

5.2 Risk-Adjusted Returns to Overlapped Positions

One inference we can derive from excess returns to treatment group funds is that overlapped positions convey information to asset prices by way of a network of connected hedge fund managers. To pin down this conclusion, I explore differences in risk-adjusted returns between overlapped and non-overlapped–or unique–portions of grouped manager portfolios by dividing the dataset into overlapping and non-overlapping stock-quarters for the treatment and control groups.

The results of these tests are shown in Table 6. In Panel A, I find the difference in the coefficient for excess returns is larger and more statistically significant for overlapped stocks between the two groups relative to their overall respective portfolios. In Panel B, non-overlapped positions contribute little value to the treatment group's risk-adjusted performance, whereas for the control group, non-overlapped stocks actually creates negative alpha of -5 basis points a month. Taken together, overlapped positions drive excess return for both groups, but accounts for 14% more of the excess return among connected funds (70% versus 56%). Put differently, overlapped positions drive 25% more of the abnormal return for the treatment group over that of the control group. This suggests the treatment group may have an informational advantage over the control group regarding certain investments that permeates through a connected network.

Supporting Pool et al. (2015) and Stein (2008), these results infer a necessary condition that manager networks are generally active and are conveying information to asset prices. In fact, a long/short portfolio comprised of a long position in overlapped holdings in connected portfolios (overlapped-connceted) and a short position in non-overlapped holdings in unconnected portfolios (unique-unconnected) generates alpha of 4.5% per year.

[Table 6 here]

[16]Differences in the constant term between treatment and control groups is marginally insignificant at the monthly level, but decidedly significant at the quarterly and annual horizon.

5.3 Risk-Adjusted Returns During the 2008 Financial Crisis

On the heals of the global financial crisis of 2008-2009, liquidity risk borne by hedge funds has witnessed a significant increase in scrutiny. However, as Aragon (2007) points out, funds with more stringent lockup periods are better able to manage illiquid positions, thereby extracting liquidity premia for investors. Thus, funds that offer longer lockups should be in a better position to manage redemption pressure during extreme market environments.

As many of the funds in this proprietary sample are not found in the commercial databases, I cannot confirm whether differential lockup terms arise between the treatment and control groups. However, given the institutional quality of the funds found in this sample, there is little reason to suspect sophisticated investors would accept more onerous liquidity terms among two funds who exhibit similar signals ex-ante. Acceding to such terms, itself, is a signal that a) the investor is informed about a manager's distinguishing background, or b) they prefer to leave liquidity management decisions to a manager they deem better adept at managing liquidity vis-à-vis the performance tradeoff than they themselves could do. Further, even though a fund may have more stringent lockup provisions, the manager need not necessarily adhere to them, particularly when "haircut" fees allow investors to redeem funds prematurely. The ability to maintain full discretion over a portfolio by adhering to lockups during market down turns, particularly among volumes of early redemption requests, highlights liquidity management skill even in the presence of lockup terms.

Furthermore, an extreme market shock, particularly a liquidity driven event such as the financial crisis, provides a unique setting to analyze deviations in manager behavior within a network. To the extent that investment ideas communicated through a connected network contain no valuable information, then adverse market conditions provide incentive for connected funds to divest from commonly held, but non-informative, positions. Widely-held positions in a network makes them susceptible to liquidity induced drawdowns, particularly for information-lacking positions. Accordingly, the strength of a network can be tested amid an economic shock such as the financial crisis; if the network is robust, socially communicated investment ideas in overlapped portfolios should not exhibit liquidity weakness in the cross-section.

To analyze this setting, I employ the same risk-adjusted performance specification used before, and add indicator variables for each quarter during the crisis (3Q 2008 - 1Q 2009). Whereas the control group experiences a significant draw-down attributable to the market turmoil of the third quarter of 2008 (-2.9% per month during 3Q 2008), the treatment group exhibits no statistically significant losses attributable to the quarter (Table 7). This result attests to differentiated liquidity management during the period of largest liquidity shock in the crisis, and is reinforced by differences in liquidity loadings between the two groups. Similar cross-sectional differences remain when we progress into the fourth quarter of 2008 and first quarter of 2009; the treatment group continues to outpace the control group by generating statistically significantly different excess returns.

[**Table 7 here**]

Disentangling learning and network channels, Table 8 shows regression results of overlapping (Panel A) and non-overlapping (Panel B) portfolios between the treatment and control groups. Across overlapped investments, difference between the groups remains valid in each quarter tested. For example, during the third quarter of 2008, overlapped positions drive losses for the control group whereas they remain insignificant for connected managers.

Further, as markets rebounded in late 2008, the treatment group posts returns twice that of the control group attributable to similar cross-sectional shocks (1.3% vs. 0.6%), and during the first quarter of 2009, differences remain between overlapped portfolios of the two groups. Results obtained from non-overlapping portfolio regressions in Panel B confirm those found previously; the control group generates negative excess returns from unique investments while the treatment group breaks even. Taken together, these results buttress evidence that overlapped portfolios among connected funds convey information to asset prices as a result of an active network. Furthermore, connected fund managers do not deviate from their networked investment ideas during adverse market conditions.

[**Table 8 here**]

5.4 Price Discovery Among Connected Funds

Lastly, I compare the timeline of returns to overlapping positions and how networks of fund managers respond. Following Coval and Stafford (2007), I calculate cumulative average abnormal returns (CAARs) to overlapping stocks near an event quarter. CAARs are monthly compounded differences between returns to an event-stock and the equal-weighted average return to stocks held by each group of funds in the cross-section. I define an event quarter as when five or more funds in the treatment and control groups, respectively, hold the same stock-quarter pairing, where the event at time t is measured at the end of the quarter where the overlapping threshold is crossed. I calculate CAARs for each event-stock starting 15 months leading up to the event and for the following 18 months.[17] This allows us to see if there are differential returns to overlapped positions between the treatment and control group. Further, it provides enough horizon to determine whether cumulative returns fully reverse, which would support a price impact story–aking to a pump-and-dump strategy–or whether they remain above the prior price point, which would provide evidence that privately generated information is impounding into asset prices.

[Figure 2 here]

As shown in Figure 2, CAARs for both the treatment (solid blue line) and control groups (solid red line) increase toward the event at time t, continue to increase through the next quarter, and begin to reverse in the following quarter. Notably, both groups see a similar trend, however, difference in CAARs between the two groups is stark; event-stocks in the treatment group achieve CAARs above 25.3%, whereas those of the control group only reach 9.5%. Furthermore, whereas CAARs fully reverse and become negative for the control group, they reverse only partially for the treatment group. This indicates that information propagating through the treatment group network is not fully transitory, and thus, conveys valuable information to asset prices.

Interestingly, the difference in CAARs between the two groups (solid gray line) highlights the differential impact the treatment group has on asset prices over the control group. From this, it is clear that differences in shared portfolio choices between connected and unconnected funds generates significant abnormal return, which persists long after the event period has passed. If information communicated through connected funds contained no differential information from that of unconnected managers, the difference in CAARs should fully reverse to zero.

Figure 2 also shows the average number of funds in each group holding the stock-quarter surrounding the event. The average number of funds are represented by dashed lines; blue for the treatment group and red for the control. For both groups of funds, the average number of funds holding the event stock-quarter increases monotonically towards the event, remains high during the following quarter, and reverts dramatically in the next quarter. This implies that funds are buy-herding into event stocks and are able to capture the peak returns to these positions before subsequently exiting. However, the treatment group appears to do so in a larger grouping leading up to the event, with a delta of roughly 10% more funds participating in the event-stock for the

[17]Results hold when the event quarter is tested for various number of funds holding the same stock-quarter and for differing pre- and post-measurement horizons.

treatment group. In sum, these results provide corroborating evidence that overlapped holdings are driven by an exchange of privately generated information as suggested in Pool et al. (2015) and Stein (2008).

6 Conclusion

Despite the expansive literature on mutual fund families, there has been little attention paid to the linkages between connected hedge funds traced by a common employment history. Using holdings information combined with a proprietary dataset of hedge fund manager employment histories, I show how connected hedge fund managers express correlated investment behavior. Firstly, learned skills from a formative mentor such as portfolios construction and trading techniques appear with commonality across portfolios of connected members. Second, connected funds leverage their social network by holding overlapped portfolios 48.5% more of the time than unconnected fund managers, suggesting an active network.

I then show how connected managers differ their trading and allocation decisions between uniquely and commonly held investments. My stock-level analyses on conditional herding find that connected managers buy- and sell-herd more than unconnected managers. Furthermore, they over-weight overlapped positions 12.4% more. This additional evidence supports the premise that an active network among connected managers influences investment decisions.

As modeled in the "quid pro quo" framework of Stein (2008), I find that overlapped portfolios drive risk-adjusted alpha, accounting for 70% of the excess return to the overall portfolio compared to 56% for the control group; a 25% improvement. This finding buttresses those of Pool et al. (2015) who find that information transfers from networked hedge fund managers to asset prices. Further, during periods of extreme market duress when managers have an incentive to deviate from the herd, I find connected managers maintain conviction in overlapped positions shared among the network.

Importantly, I show that cumulative average abnormal returns to connected-overlapped stocks of 15% is permanently conveyed to asset prices (Coval and Stafford (2007)). This suggests that incentives within networks of hedge fund managers are aligned to generate private information, which leads to price discovery for commonly held assets in connected portfolios. I calculate that a long/short portfolio of overlapped-connected/unique-unconnected stocks generates alpha of 4.5%

annually.

My findings are also relevant to the fund family literature. By identifying commonality in investment behavior among hedge fund managers connected through a shared employment ancestry, an informal hedge fund "family" can be identified, thereby reducing costly search for institutional investors. Furthermore, this identification can assist the institutional investor in diversifying risk not heretofore noted among a portfolio of connected hedge fund managers; that of commonly held positions.

References

Agarwal, Vikas, Vyacheslav Fos, and Wei Jiang, 2013, Inferring reporting-related biases in hedge fund databases from hedge fund equity holdings, *Management Science* 59, 1271–1289.

Anson, Mark JP, 2006, Handbook of alternative investments, *New York: Willey* .

Aragon, George O, 2007, Share restrictions and asset pricing: Evidence from the hedge fund industry, *Journal of Financial Economics* 83, 33–58.

Aragon, George O, and J Spencer Martin, 2012, A unique view of hedge fund derivatives usage: Safeguard or speculation?, *Journal of Financial Economics* 105, 436–456.

Bertrand, Marianne, and Antoinette Schoar, 2003, Managing with style: The effect of managers on firm policies, *The Quarterly Journal of Economics* 118, 1169–1208.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of political Economy* 992–1026.

Brown, Stephen J, Thomas L Fraser, and Bing Liang, 2008, Hedge fund due diligence: A source of alpha in a hedge fund portfolio strategy .

Brown, Stephen J, and Christopher Schwarz, 2013, Do market participants care about portfolio disclosure? evidence from hedge funds' 13f filings, *Evidence from Hedge Funds' 13F Filings (May 8, 2013)* .

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of finance* 52, 57–82.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The small world of investing: Board connections and mutual fund returns, *Journal of Political Economy* 116, 951–979.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Elton, Edwin J, Martin J Gruber, and T Clifton Green, 2007, The impact of mutual fund family membership on investor risk, *Journal of Financial and Quantitative Analysis* 42, 257–277.

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.

Froot, Kenneth A, David S Scharfstein, and Jeremy C Stein, 1992, Herd on the street: Informational inefficiencies in a market with short-term speculation, *The Journal of Finance* 47, 1461–1484.

Fung, William, and David A Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 65–80.

Gerritzen, Marc, Jens Carsten Jackwerth, and Alberto Plazzi, 2016, Birds of a feather–do hedge fund managers flock together?, *Available at SSRN 2729993* .

Gervais, Simon, Anthony W Lynch, and David K Musto, 2005, Fund families as delegated monitors of money managers, *Review of Financial Studies* 18, 1139–1169.

Griffin, John M, and Jin Xu, 2009, How smart are the smart guys? a unique view from hedge fund stock holdings, *Review of Financial Studies* 22, 2531–2570.

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *The Journal of Finance* 62, 251–301.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *The Journal of Finance* 60, 2801–2824.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *The Journal of Finance* 60, 1983–2011.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Mallaby, Sebastian, 2011, *More Money Than God: Hedge Funds and the Making of the New Elite* (Bloomsbury Publishing).

Markowitz, Henry, 1959, Portfolio selection: efficient diversification of investments, *New York* .

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.

Papageorgiou, Nicolas A, Jerry T Parwada, and Kian ME Tan, 2011, Where do hedge fund managers come from? past employment experience and managerial performance, *UNSW Australian School of Business Research Paper* .

Pástor, Ľuboš, and Robert F Stambaugh, 2003, Liquidity risk and expected stock retums, *The Journal of Political Economy* 111, 642–685.

Pastor, Lubos, Robert F Stambaugh, and Lucian A Taylor, Forthcoming, Do funds make more when they trade more?, *Journal of Finance* .

Pool, Veronika K, Noah Stoffman, and Scott E Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *The Journal of Finance* 70, 2679–2732.

Scharfstein, David S, and Jeremy C Stein, 1990, Herd behavior and investment, *The American Economic Review* 465–479.

Sharpe, William F, 1970, *Portfolio theory and capital markets*, volume 217 (McGraw-Hill New York).

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

Stein, Jeremy, 2008, Conversations among competitors, *American Economic Review* 98, 2150–62.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *The Journal of Finance* 54, 581–622.

Figure 1: Distribution of Quarterly Holdings

This figure shows the distribution of average portfolio holdings for each fund found in the treatment group and the unmatched sample for the period 2000-2013.

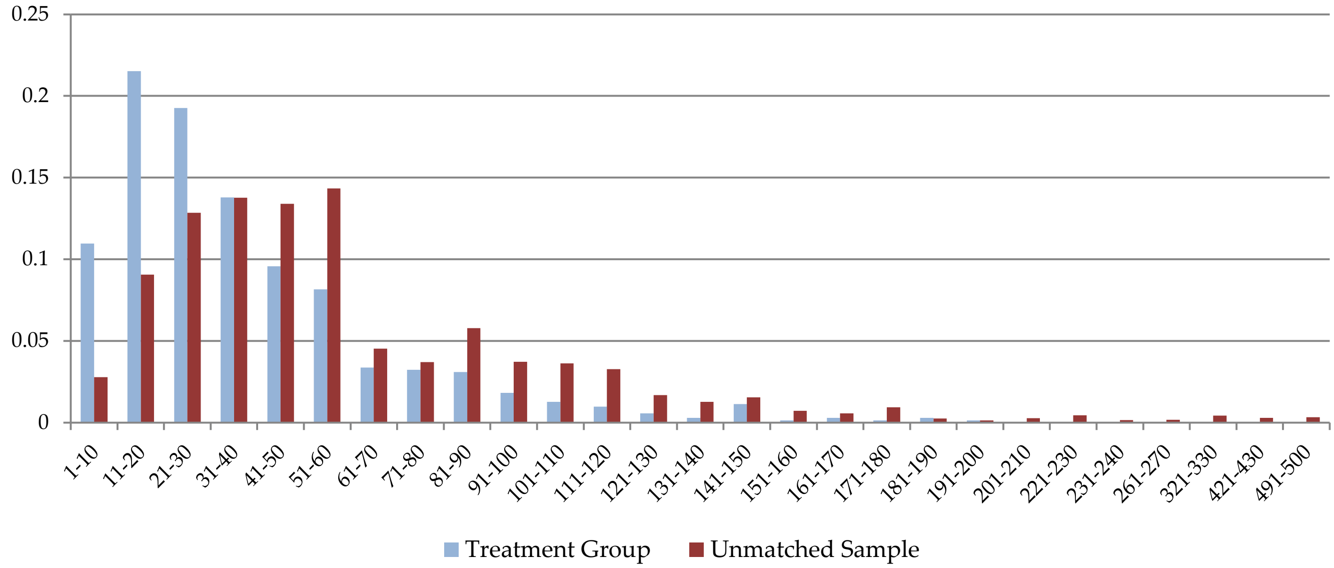

Figure 2: Monthly Cumulative Average Abnormal Returns (CAARs) near overlapping events

This figure shows CAARs for stocks held across all portfolios near overlapping events. CAAR is measured as monthly returns to stock i in excess of the equal-weighted average return of all stocks held by mutual funds at the start of the month, as used in Coval and Stafford (2007). An overlapping event is defined as five or more funds holding the same stock-quarter, and is highlighted in the figure as the vertical line at time t . A similar figure obtains when the overlapping event is defined when as few as two funds hold the same stock-quarter, or more than 10 funds, and when the time interval is shortened or lengthened. The dashed vertical line indicates when 13F holdings information becomes publicly available on average (t +45 days), as most fund managers wait until the mandated disclosure date to file 13F forms. Accordingly, informaion disclosed on date t +45 days contains holdings data for each fund held at the end of the prior quarter t .

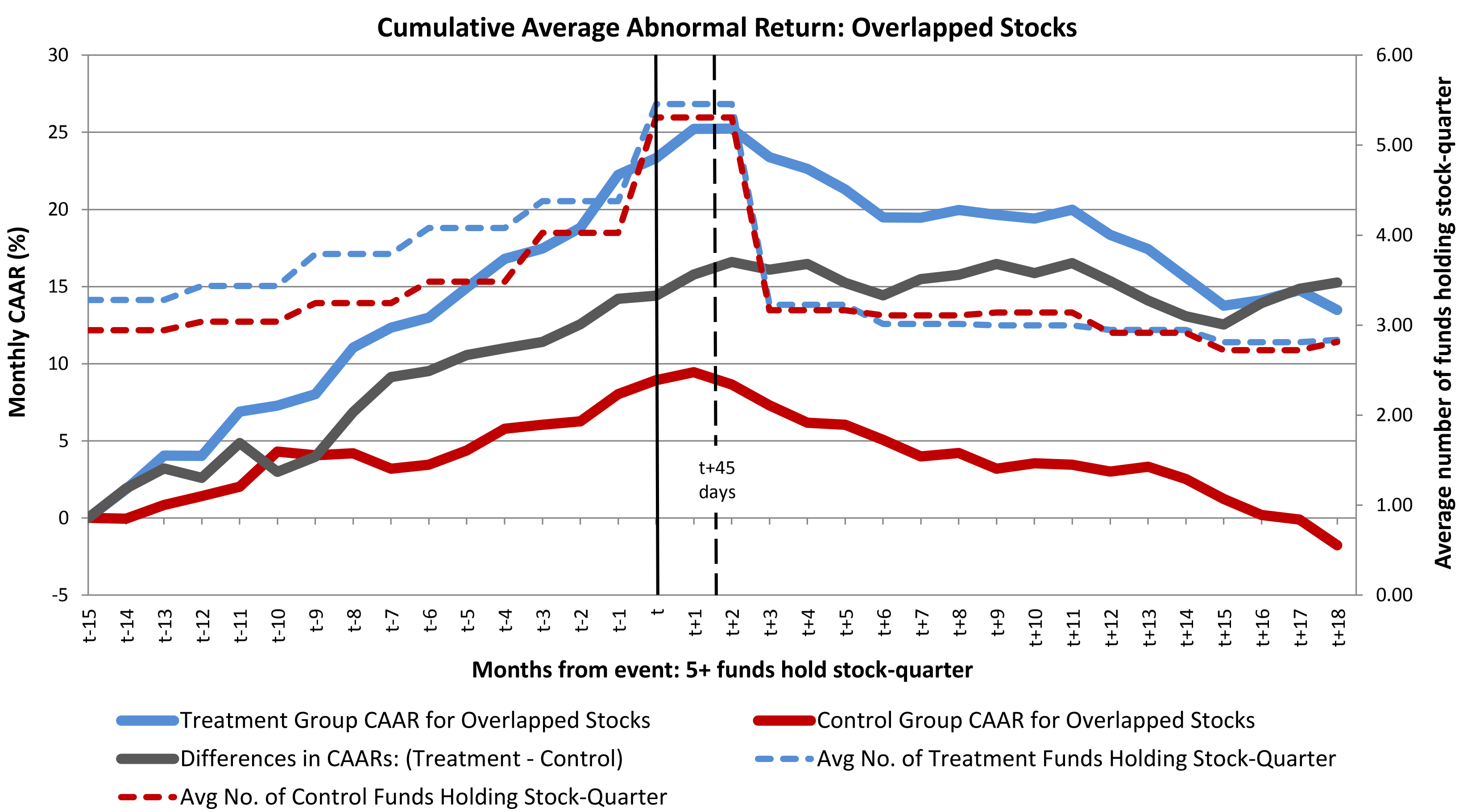

Table 1: Summary Statistics

This table presents summary statistics by sub-group for the quarterly sample period of 2000 to 2013. The treatment group is comprised of U.S. headquarterd long-short equity hedge funds ran by managers whom were previously employed by Tiger Management. The unmatched sample includes all U.S. headquartered long-short equity hedge funds in my sample from which I use to construct the matched control group. The matched control group is composed of the two nearest-neighbors match to the treatment group after matching on (1) fund start date, (2) fund longevity, (3) average number of positions, (4) average position weight, and (5) fund headquarter location.

	Unmatched Sample			Treatment Group			Control Group		
Variable	Mean	Median	Std	Mean	Median	Std	Mean	Median	Std
Number of Funds	232	232	0	42	42	0	84	84	0
Number of Positions	95	42	165	39	34	23	39	34	24
Position Weight (%)	1.05	0.21	2.45	2.58	1.59	3.39	2.57	1.60	3.37
Position Horizon (year)	0.93	0.69	0.99	0.73	0.66	0.38	0.81	0.66	0.64
Top Ten Positions (% of AUM)	58.3	56.3	22.9	64.1	61.8	18.7	63.8	61.1	17.4
Invested Long Assets ($MM)	1,239	349	2,741	1,560	729	2,484	779	318	1,222
Fund Longevity (Qtrs)	30.6	29.0	16.2	32.6	29.0	14.2	32.5	29.0	14.2
Raw Quarterly Return (%)	3.21	3.89	11.52	4.20	4.57	11.55	3.72	4.28	11.62

Table 2: Portfolio Overlap

This table presents the results of the quarterly panel regression

$$PortOverlap_{j,k,t} = \alpha + \beta SameNetwork_{j,k,t} + \delta SameCity_{j,k,t} + \Gamma' Controls_{j,k,t} + \varepsilon_{j,k,t},$$

where Port*Overlap* $_{j,k,t}$ measures the minimum portfolio overlap between funds j and k during quarter t. *SameNetwork* $_{j,k,t}$ is a dummy variable equal to one if funds j and k belong to the same hedge fund network. *SameCity* $_{j,k,t}$ is a dummy variable that is one if funds j and k are headquartered in the same city. *Controls* $_{j,k,t}$ is a vector of control variables that include (a) dummy variables equal to one if funds j and k pursue the same hedge fund strategy; (b) the average assets under management (AUM) based quintiles of funds j and k (*AUMQuintAvg*); and (c) the absolute value in the differences between AUM-based quintiles of funds j and k (*AUMQuintDiff*). Standard errors are two-way clustered at the fund level for each fund in the fund-pair and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	All Strategies			Equity Long/Short		
Dependent Variable	*PortOverlap* $_{j,k,t}$			*PortOverlap* $_{j,k,t}$		
SameNetwork		1.75***	1.37***		2.06***	1.82***
		(4.49)	(3.38)		(4.84)	(4.26)
SameCity	-0.19	-0.11	0.03	-0.10	-0.17	-0.02
	(-0.84)	(-0.50)	(0.17)	(-0.38)	(-0.66)	(-0.09)
Activist		-2.68***	-4.89***			
		(-4.38)	(-6.05)			
Equity Long/Short		-0.63	-1.30***			
		(-1.38)	(-3.08)			
Equity Market Neutral		-1.18**	-1.66***			
		(-2.21)	(-4.33)			
Event Driven		-0.86	-1.02**			
		(-1.36)	(-2.04)			
Fixed Income		-0.68	-0.86*			
		(-1.16)	(-1.69)			
Fund of Funds		-0.21	0.52			
		(-0.25)	(0.70)			
Global Macro		0.30	-0.04			
		(0.31)	(-0.06)			
Multi-Strategy		1.05*	0.38			
		(1.93)	(0.82)			
AUMQuintAvg			1.87***			1.34***
			(7.77)			(5.12)
AUMQuintDiff			-0.74***			-0.53***
			(-5.77)			(-4.89)
Constant	5.39***	5.54***	3.48***	4.99***	4.91***	2.94***
	(22.85)	(12.62)	(10.18)	(18.25)	(17.49)	(12.42)
R^2	0.000	0.009	0.077	0.000	0.004	0.053
Observations	1,393,952	1,393,952	1,393,952	708,873	708,873	708,873

Table 3: Herding Measure

This table presents the results of the herding measure for each of the treatment (Panel B), matched control (Panel C), and unmatched groups (Panel A). The herding measure comes from Lakonishok et al. (1992), and is expressed in Wermers (1999) as

$$HM_{i,t} = |p_{i,t} - E[p_{i,t}]| - E|p_{i,t} - E[p_{i,t}]| ,$$

where, $p_{i,t}$ is the proportion of funds buying stock i in quarter t relative to all funds trading the same stock-quarter. $E[p_{i,t}]$ is subtracted from $p_{i,t}$ to expose herding variation relative to unperterbed coincident trading that would be expected in a given quarter. The second term, $E|p_{i,t} - E[p_{i,t}]|$, is the adjustment factor to allow for random variation around the proportion of buyers over the full sample period where, in expectation, no herding arises. Measures reported are average values. HM represents the average of the herding measure. BHM represents the average of the conditional buy-side herding measure. SHM represents the average of the conditional sell-side herding measure. BHM-SHM represents the average spread to a portfolio of conditional herding measures; buy-side minus sell-side . * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Funds Trading Stock-Quarter Pairs				
Mean Measure	≥3	≥5	≥10	≥15	≥20
Panel A: Full Sample					
HM	0.00%	0.00%	0.00%	0.00%	0.00%
Observations	74,916	41,880	13,179	4,739	1,876
BHM	12.69%	11.42%	9.56%	8.64%	8.38%
Observations	51,204	28,870	9,063	3,241	1,263
SHM	11.24%	9.96%	8.02%	7.07%	6.48%
Observations	23,712	13,010	4,116	1,498	613
BHM - SHM	*1.45%*	*1.46%*	*1.54%*	*1.57%*	*1.90%*
Panel B: Treatment Group					
HM	0.00%	0.00%	0.00%	0.00%	
Observations	3,335	825	127	27	
BHM	15.11%	14.27%	11.02%	9.58%	
Observations	2,449	597	87	15	
SHM	15.60%	12.96%	9.28%	8.77%	
Observations	886	228	40	12	
BHM - SHM	*-0.48%*	*1.31%*	*1.75%*	*0.81%*	
Panel C: Control Group					
HM	0.00%	0.00%	0.00%		
Observations	2,443	346	8		
BHM	14.56%	11.94%	7.12%		
Observations	1,640	229	5		
SHM	11.69%	10.76%	7.95%		
Observations	803	117	4		
BHM - SHM	*2.87%*	*1.18%*	*-0.84%*		
Panel D: Differences in Conditional Means					
BHM	0.55***	2.33***	3.91***	9.58***	
tstat	(2.71)	(4.35)	(2.20)	(6.19)	
SHM	3.91***	2.20***	1.32	8.77***	
tstat	(11.22)	(3.32)	(0.31)	(4.75)	

Table 4: Adjusted Weight Panel Regressions

This table presents the results of the quarterly panel regression

$$AdjustedWeight_{i,j,t} = \alpha + \beta Overlap_{i,j,t} + \delta Treatment_j + \gamma(Overlap_{i,j,t} * Treatment_j) + \tau_t + \varepsilon_{i.j,t},$$

where $AdjustedWeight_{i,j,t}$ is the position weight of stock i in fund j in quarter t, adjusted by relative portfolio market capitalizations. The adjustment factor is stock i's market capitalization at the end of quarter t held by fund j, scaled by the sum of the market capitalizations for each stock held in the fund at the end of the quarter. $Overlap_{i,j,t}$ is equal to one if two or more funds hold the same stock-quarter pair. Note that $Overlap_{i,j,t}$ is measured for each group separately. $Treatment_j$ is equal to one for all hedge funds in the treatment group, and zero for all control funds. Newey-West autocorrelation and heteroskedasticity-consistent standard errors are clustered at the fund-stock pair level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	$AdjustedWeight_{i,j,t}$			
Constant	2.15***	2.23***	2.17***	2.24***
	(19.55)	(16.78)	(19.95)	(16.99)
$Overlap_{i,j,t}$	0.83***	0.69***	0.78***	0.66***
	(11.09)	(8.87)	(10.37)	(8.27)
$Treatment_j$		-0.26		-0.21
		(-1.10)		(-0.90)
$Overlap_{i,t}$ * $Treatment_j$		0.42**		0.36**
		(2.44)		(2.14)
R^2	0.015	0.016	0.023	0.024
Observations	156,304	156,304	156,304	156,304
Date FE	No	No	Yes	Yes

Table 5: Risk-Adjusted Performance

This table presents the results of the fund-month panel regression

$$(R_{j,t} - R_{f,t}) = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \beta_5 LIQ_t + \varepsilon_{j,t},$$

where $(R_{j,t} - R_{f,t})$ represents the excess return to portfolio j during month t. Monthly portfolio returns are constructed from all stocks held by fund j at the end of quarter t, multiplied by weighted returns to each stock over the ensuing three months. Risk adjustment factors include the Fama and French (1993) three-factor model (MKT, SMB, and HML), plus the momentum factor of Carhart (1997; MOM), and the liquidity factor of Pástor and Stambaugh (2003; LIQ). MKT represents excess return to the market factor. SMB is the spread return to a portfolio of small market capitalization stocks minus large capitalization stocks. HML is the spread return to a portfolio of high book-to-market stocks minus low book-to-market stocks. MOM is the spread return to a portfolio of stocks with prior positive returns minus stocks with prior negative returns. LIQ is the factor for cross-sectional permanent liquidity innovations. Differences in *Alphas* between the treatment and control group are significantly different from one another at the quarterly (5% level) and annual (1% level) frequency, but marginally insignificantly different at the monthly frequency shown here. Standard errors are clustered two-ways by fund and date, with t-statistics reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Treatment Group			Control Group		
Dependent Variable	R_j - R_f			R_j - R_f		
Alpha	0.44***	0.45***	0.43***	0.36***	0.38***	0.32***
	(6.54)	(6.89)	(6.14)	(6.98)	(7.50)	(6.38)
MKT	0.86***	0.84***	0.84***	0.85***	0.82***	0.82***
	(28.22)	(30.32)	(30.52)	(34.71)	(36.03)	(35.95)
SMB	0.26***	0.26***	0.25***	0.23***	0.24***	0.22***
	(4.40)	(4.55)	(4.20)	(6.86)	(7.11)	(6.66)
HML	-0.16**	-0.18**	-0.16*	0.04	0.02	0.05
	(-2.13)	(-2.39)	(-2.00)	(0.78)	(0.35)	(1.02)
MOM		-0.05***	-0.05***		-0.05***	-0.06***
		(-2.95)	(-3.14)		(-5.02)	(-6.02)
LIQ			0.05**			0.10***
			(2.02)			(5.10)
R^2	0.557	0.559	0.561	0.590	0.592	0.598
Observations	4,801	4,801	4,801	8,581	8,581	8,581

Table 6: Overlap vs. Non-Overlap Risk-Adjusted Performance

This table presents the results of the fund-month panel regression

$$(R_{j,t} - R_{f,t}) = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \beta_5 LIQ_t + \varepsilon_{j,t},$$

where $(R_{j,t} - R_{f,t})$ represents the excess return to portfolio j during month t. Monthly portfolio returns are constructed from all stocks held by fund j at the end of quarter t, multiplied by weighted returns to each stock over the ensuing three months. Risk adjustment factors include the Fama and French (1993) three-factor model (MKT, SMB, and HML), plus the momentum factor of Carhart (1997; MOM), and the liquidity factor of Pástor and Stambaugh (2003; LIQ). MKT represents excess return to the market factor. SMB is the spread return to a portfolio of small market capitalization stocks minus large capitalization stocks. HML is the spread return to a portfolio of high book-to-market stocks minus low book-to-market stocks. MOM is the spread return to a portfolio of stocks with prior positive returns minus stocks with prior negative returns. LIQ is the factor for cross-sectional permanent liquidity innovations. Differences in *Alphas* between the treatment and control group are significantly different from one another at the quarterly (5% level) and annual (1% level) frequency, but marginally insignificantly different at the monthly frequency shown here. Standard errors are clustered two-ways by fund and date, with t-statistics reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel A: Overlap Risk-Adjusted Performance

	Treatment Group			Control Group		
Dependent Variable	$R_j - R_f$			$R_j - R_f$		
Alpha	0.31***	0.32***	0.30***	0.20***	0.21***	0.18***
	(5.92)	(6.14)	(5.76)	(7.70)	(7.74)	(6.72)
MKT	0.54***	0.53***	0.53***	0.39***	0.39***	0.39***
	(16.99)	(16.38)	(16.38)	(24.34)	(24.41)	(24.47)
SMB	0.10***	0.10***	0.09***	0.05***	0.06***	0.05***
	(5.16)	(5.35)	(5.17)	(5.38)	(5.43)	(4.87)
HML	-0.11***	-0.12***	-0.11***	0.04*	0.04*	0.05***
	(-3.68)	(-4.02)	(-3.65)	(1.96)	(1.78)	(2.88)
MOM		-0.02*	-0.02**		-0.01	-0.01**
		(-1.96)	(-2.19)		(-1.64)	(-2.54)
LIQ			0.03**			0.05***
			(2.44)			(4.16)
R^2	0.513	0.514	0.515	0.450	0.450	0.455
Observations	4,730	4,730	4,730	8,364	8,364	8,364

Panel B: Non-Overlap Risk-Adjusted Performance

	Treatment Group			Control Group		
Dependent Variable	$R_j - R_f$			$R_j - R_f$		
Alpha	0.02	0.03	0.02	-0.05***	-0.05***	-0.05***
	(0.54)	(0.94)	(0.63)	(-4.23)	(-3.76)	(-3.97)
MKT	0.34***	0.33***	0.33***	0.07***	0.06***	0.06***
	(11.25)	(10.77)	(10.79)	(19.91)	(19.47)	(19.46)
SMB	0.16***	0.17***	0.16***	0.02***	0.02***	0.02***
	(3.19)	(3.32)	(3.15)	(3.22)	(3.88)	(3.85)
HML	-0.08	-0.09	-0.09	0.00	-0.00	-0.00
	(-1.24)	(-1.50)	(-1.27)	(0.26)	(-0.60)	(-0.45)
MOM		-0.04***	-0.04***		-0.02***	-0.02***
		(-3.37)	(-3.45)		(-3.96)	(-4.01)
LIQ			0.02			0.00
			(0.81)			(1.22)
R^2	0.303	0.307	0.307	0.156	0.166	0.166
Observations	4,706	4,706	4,706	5,373	5,373	5,373

Table 7: Liquidity Management During the Global Financial Crisis

This table presents the results of the fund-month panel regression

$(R_{j,t}-R_{f,t}) = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \beta_5 LIQ_t + \beta_6 Q3\_2008_t + \beta_7 Q4\_2008_t + \beta_8 Q1\_2009_t + \varepsilon_{j,t},$

during the global financial crisis of 2008-2009 where (Rj,t- Rf,t) represents the excess return to portfolio j during month t. Monthly portfolio returns are constructed from all stocks held by fund j at the end of quarter t, multiplied by weighted returns to each stock over the ensuing three months. Risk adjustment factors include the Fama and French (1993) three-factor model (*MKT* , *SMB* , and *HML*), plus the momentum factor of Carhart (1997; *MOM*), and the liquidity factor of Pástor and Stambaugh (2003; LIQ). *MKT* represents excess return to the market factor. *SMB* is the spread return to a portfolio of small market capitalization stocks minus large capitalization stocks. *HML* is the spread return to a portfolio of high book-to-market stocks minus low book-to-market stocks. *MOM* is the spread return to a portfolio of stocks with prior positive returns minus stocks with prior negative returns. *LIQ* is the factor for cross-sectional permanent liquidity innovations. Indicators variables *Q3_2008-Q2_2009* are equal to one for each of the respective quarters during the crisis. Standard errors are clustered two-ways by fund and date, with t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Treatment Group			Control Group		
Dependent Variable	R_j - R_f			R_j - R_f		
Alpha	0.45***	0.38***	0.39***	0.39***	0.28***	0.30***
	(6.47)	(4.96)	(5.33)	(8.15)	(5.24)	(5.79)
MKT	0.83***	0.85***	0.84***	0.80***	0.83***	0.83***
	(29.93)	(32.55)	(30.90)	(36.11)	(37.36)	(35.90)
SMB	0.26***	0.25***	0.25***	0.26***	0.22***	0.22***
	(4.55)	(4.04)	(4.23)	(7.79)	(6.48)	(6.72)
HML	-0.16*	-0.15*	-0.14*	0.07	0.07	0.07
	(-1.93)	(-1.74)	(-1.68)	(1.47)	(1.33)	(1.31)
MOM	-0.06***	-0.05***	-0.04**	-0.07***	-0.06***	-0.06***
	(-3.20)	(-3.10)	(-2.46)	(-6.92)	(-5.95)	(-5.05)
LIQ	0.05*	0.07**	0.04	0.09***	0.11***	0.09***
	(2.00)	(2.32)	(1.59)	(4.76)	(5.53)	(4.57)
Q3_2008	-0.84			-2.86***		
	(-1.16)			(-6.16)		
Q4_2008		1.71***			1.41***	
		(2.71)			(3.71)	
Q1_2009			1.52***			1.25***
			(3.13)			(3.57)
R^2	0.561	0.562	0.562	0.603	0.599	0.599
Observations	4,801	4,801	4,801	8,581	8,581	8,581

Table 8: Overlap vs. Non-Overlap Liquidity Management During the Global Financial Crisis

This table presents the results of the fund-month panel regression

$$(R_{j,t}-R_{f,t}) = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \beta_5 LIQ_t + \beta_6 Q3\_2008_t + \beta_7 Q4\_2008_t + \beta_8 Q1\_2009_t + \varepsilon_{j,t},$$

during the global financial crisis of 2008-2009 where (Rj,t- Rf,t) represents the excess return to portfolio j during month t. Monthly portfolio returns are constructed from all stocks held by fund j at the end of quarter t, multiplied by weighted returns to each stock over the ensuing three months. Risk adjustment factors include the Fama and French (1993) three-factor model (MKT, SMB, and HML), plus the momentum factor of Carhart (1997; MOM), and the liquidity factor of Pástor and Stambaugh (2003; LIQ). MKT represents excess return to the market factor. SMB is the spread return to a portfolio of small market capitalization stocks minus large capitalization stocks. HML is the spread return to a portfolio of high book-to-market stocks minus low book-to-market stocks. MOM is the spread return to a portfolio of stocks with prior positive returns minus stocks with prior negative returns. LIQ is the factor for cross-sectional permanent liquidity innovations. Indicators variables *Q3_2008-Q2_2009* are equal to one for each of the respective quarters during the crisis. Standard errors are clustered two-ways by fund and date, with t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel A: Overlap Risk-Adjusted Performance

	Treatment Group			Control Group		
Dependent Variable	R_j - R_f			R_j - R_f		
Alpha	0.32***	0.27***	0.27***	0.22***	0.16***	0.16***
	(6.10)	(4.98)	(5.04)	(8.19)	(5.80)	(5.69)
MKT	0.52***	0.54***	0.54***	0.37***	0.39***	0.39***
	(16.87)	(16.80)	(16.46)	(25.91)	(24.25)	(24.59)
SMB	0.10***	0.09***	0.10***	0.07***	0.05***	0.05***
	(5.72)	(5.00)	(5.28)	(6.44)	(4.68)	(5.02)
HML	-0.10***	-0.10***	-0.09***	0.07***	0.06***	0.07***
	(-3.41)	(-3.34)	(-2.94)	(3.65)	(3.28)	(3.67)
MOM	-0.03**	-0.02**	-0.01	-0.02***	-0.01**	-0.01
	(-2.35)	(-2.13)	(-1.21)	(-3.50)	(-2.50)	(-1.04)
LIQ	0.03**	0.04***	0.02	0.04***	0.06***	0.04***
	(2.36)	(3.61)	(1.42)	(3.73)	(4.40)	(3.23)
Q3_2008	-0.64			-1.74***		
	(-1.58)			(-5.87)		
Q4_2008		1.29***			0.61**	
		(3.51)			(2.49)	
Q1_2009			1.42***			1.23***
			(5.45)			(6.65)
R^2	0.516	0.518	0.518	0.463	0.456	0.459
Observations	4,730	4,730	4,730	8,364	8,364	8,364

Panel B: Non-Overlap Risk-Adjusted Performance

	Treatment Group			Control Group		
Dependent Variable	R_j - R_f			R_j - R_f		
Alpha	0.03	0.01	0.02	-0.05***	-0.06***	-0.06***
	(0.74)	(0.16)	(0.51)	(-3.75)	(-4.21)	(-4.28)
MKT	0.32***	0.33***	0.33***	0.06***	0.06***	0.06***
	(10.28)	(11.02)	(10.63)	(19.44)	(19.18)	(20.23)
SMB	0.17***	0.16***	0.16***	0.02***	0.02***	0.02***
	(3.25)	(3.06)	(3.16)	(3.83)	(3.82)	(3.92)
HML	-0.08	-0.08	-0.08	-0.00	-0.00	0.00
	(-1.24)	(-1.14)	(-1.20)	(-0.36)	(-0.21)	(0.02)
MOM	-0.04***	-0.04***	-0.04***	-0.02***	-0.02***	-0.02***
	(-3.41)	(-3.43)	(-3.29)	(-4.03)	(-3.99)	(-3.64)
LIQ	0.02	0.02	0.02	0.00	0.00	0.00
	(0.82)	(0.91)	(0.81)	(1.14)	(1.56)	(0.58)
Q3_2008	-0.19			-0.08		
	(-0.35)			(-1.29)		
Q4_2008		0.56			0.15	
		(1.25)			(1.58)	
Q1_2009			0.12			0.26**
			(0.34)			(2.57)
R^2	0.307	0.308	0.307	0.167	0.167	0.168
Observations	4,706	4,706	4,706	5,373	5,373	5,373

ETFs in Mutual Fund Families – Cannibalization, Subsidization, and Flows

Caitlin Dannhauser
Villanova University

Harold Spilker
Boston College

March 15, 2017

Abstract

Meeting heterogeneous investor needs, mutual fund families now offer ETF versions of their index funds. Using a twin based study, we find mutual fund twins and their families benefit from the relationship, while the effect is ambiguous for ETF twins. Compared to the average index mutual fund, twins have a 23% lower tax burden driven by a 69% lower long-term capital gains yield. Unrealized capital gains also decrease by 7%. ETF twin investors face higher long-term capital gains yields and unrealized capital gains, but are compensated with lower total expense ratios. Overall, the family benefits from higher flows to twins.

Dannhauser is at the Villanova School of Business, Villanova University, 800 Lancaster Avenue Villanova, PA 19085, and can be reached at caitlin.dannhauser@villanova.edu or (610) 519-4348. Spilker is at the Carroll School of Management, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, and can be reached at tray.spilker@bc.edu or (617) 710-1289. We thank Alan Marcus, Jeffrey Pontiff, Ronnie Sadka, Jerome Taillard, and Hassan Tehranian for helpful comments.

1. Introduction

On the supply side of the mutual fund industry two stylized facts exist: (1) the majority of funds are organized as Open-End Funds (OEFs) and (2) they operate within a fund family. The prevalence of these industry features has emerged despite well-documented limitations to both. For instance, Stein (2006) theorizes that the dominance of OEFs, may be socially excessive due to the known externalities of trading, brokerage, and operating expenses, and unexpected capital gains imposed by short-term traders on other fund investors (Chordia (1996)). Traditional mutual funds organize as families to earn rents from economies of scale (Baumol et al. (1989)) and to cater to heterogeneous investor needs (Nanda, Narayanan and Warther (2000)). However, by organizing in a family structure, manager incentives may be distorted (Gaspar, Massa, and Matos (2006)) as evidenced by cross-fund subsidization (Bhattacharya, Lee, and Pool (2013)). Using a twin-based study to overcome endogeneity concerns, this paper contributes to the literature by examining an emerging trend in the organization of mutual funds: the incorporation of Exchange Traded Funds (ETFs) into a fund family.

The features that distinguish ETFs from their OEF peers helps mitigate the impact of externalities on fund investors. ETFs are exchange traded and rely on in-kind creation and redemption, such that investors bear their own transaction-induced costs and the fund itself rarely has to transact in the underlying. Thus, ETFs generally have higher levels of transparency, greater tax efficiencies, and lower management fees (Poterba and Shoven (2002)). These features have attracted broad array of investors to the new product, resulting in ETFs underpinning the recent trend toward index investing from active management. In fact, ETFs now represent 48% of the $4.23 trillion indexed funds market that has benefited from net inflows of $1.2 trillion since 2007. Of this, net new cash flows to U.S. domestic equity ETFs are almost double that of index mutual funds.[1] In light of such demand for ETFs, traditional mutual fund families face intense pressure to offer ETFs to retain existing shareholders and to attract new ones.

Fund families face a conflict. In effort to increase flows to the family, they can introduce ETFs to meet investor demand, but they do so at the risk of cannibalizing flows to higher fee OEFs within the family.[2] Yet, the operational efficiencies of ETFs allow mutual fund families to potentially offset tax externalities imposed on traditional fund investors by short-term traders. Using a twin-based identification strategy similar to Nohel, Wang, and Zheng (2010), Cici, Gibson, and Moussawi (2010), and Evans and Fahlenbrach (2012) we obtain a clean setting to test the implications of the proliferation of this alternative fund structure on the traditional mutual fund industry and mainstream investors.[3] Here, twins are defined as an index mutual fund and ETF in the same fund family that follow the same benchmark; essentially, twins are identical funds differentiated only by vehicle structure.[4] Our goal is to understand why fund families may introduce this share-class structure given the conflict it poses, and then consider the implications for mutual fund and ETF investors who exhibit differing preferences, with a focus on the tax, expense, and flow implications.

The tax efficiencies of ETFs stem from their use of a technique known as "in-kind creation and redemption," which are non-taxable events involving the exchange of a unit of ETF shares for a pre-specified basket of the underlying.[5] These exchanges occur only between the ETF sponsor and its Authorized Participants (APs), and enable the ETF to minimize both realized and unrealized capital

[1] ICI Factobook 2016
[2] Fidelity infamously resisted offering ETFs until recently. Krouse, Sarah, January 3, 2017, Wall Street Journal. "Fidelity Embraces What It Once Avoided: The ETF"
[3] http://www.google.com/patents/US20110258089
[4] The correlation of annual returns for the mutual fund and ETF twins is over 99%
[5] Specifically, Section 852(b)(6) of the US Tax Code relieves registered investment companies (RICs) from Section 311(b), which requires a corporation to pay taxes on distributed property, when redemptions are made in shares upon the demand of its shareholder.

gains distributions. Because mutual fund managers rarely use the in-kind feature (Poterba and Shoven (2002)), this leads to a tax externality because the tax burden of mutual fund investors is dependent on the behavior of others (Dickson, Shoven, and Sialm (2000)). When a mutual fund manager sells shares due to reallocation or to meet redemption requests, the remaining shareholders bear the tax burden on the fund's realized capital gains. The unique structure of twin funds allows fund families to capitalize on differences in investor preferences. Fund families that incorporate ETFs can exploit the tax efficiencies of ETFs to the benefit of their higher-fee paying mutual funds in two ways. First, is through a structure that Vanguard patented making the ETF a separate share class of the larger mutual fund.[6] As noted by Senior Investment Advisor at Vanguard, Joel Dickson, the structure uses the ETF's in-kind mechanism to get "rid of gains."[7] Second, ETFs and mutual fund twins exist as distinct funds within the fund family. In this setting, mutual funds may respond to a redemption request by delivering the basket of shares to the ETF in exchange for a creation unit of the ETF, typically 50,000 to 100,000 shares. The mutual fund can then sell the ETF shares into the market, using the funds from the sale to meet the cash redemption. This strategy is possible because the majority of fund sponsors are also APs for their own ETFs.[8]

We find that mutual fund families who introduce twin funds are able to exploit the tax efficiency of ETFs to lower their current and potential tax distributions. In fact, the use of ETFs by mutual fund twins results in a 22.8% reduction in the tax burden, or over $300 billion dollars in tax savings for mutual fund investors. Driving this lower tax burden is a 69.4% savings in the long-term capital gains distributions relative to the average mutual fund. Future taxable distributions – measured as

[6] US Patent Number 2002/0128947 A1, filed March 7, 2001 and published September 12, 2002
[7] https://advisors.vanguard.com/iam/pdf/Efficiency_transcript2.pdf
[8] The cost basis for acquirers of the underlying and ETF shares exchanged is the net asset value (NAV) at market close (Forstenhausler (2010), Gastineau (2005)). Therefore, rather than paying taxes on the accumulated share gains of the underlying, the mutual fund instead pays taxes on the difference between the NAV and the price at which the ETF is sold in the market.

reported 2014 unrealized capital gains scaled by total assets – are 7% lower for mutual fund twins. For ETFs, the overall tax burden is unaffected by this arrangement, but investors in ETF twins face higher long-term and unrealized capital gains.

Furthermore, we find that fund families do not fully pass on economies of scale benefits from managing two identical pools of assets. While ETF twins have 20% lower expense ratios, the expense ratios for mutual funds twins are not significantly different than their non-twin peers. Finally, we find that mutual fund and ETF twins generate 50% and 25% greater flows, respectively, than their non-twin peers. This suggests that fund families benefit by appealing to investor heterogeneity through this form of product diversification.

Since the introduction of a new investment is not random, we use two identification strategies to examine the implications of managing mutual fund and ETF twins. In our first test, we use a fixed effects model on the full panel and use Morningstar category fixed effects to account for unobserved style based heterogeneity (Bergstresser and Pontiff (2013)). We include year fixed effects to control for general market trends and the cyclicality of fund distributions (Sialm and Zhang (2015)). In our second specification, we take advantage of the recency of ETFs. Since index mutual funds date back to the 1970s and ETFs were only introduced in the early 1990s, there are several mutual funds that were in operation prior to the introduction of its ETF twin. We use these mutual funds as our treatment group in a difference-in-difference specification. For this test, we use non-twin mutual funds matched on Morningstar category and fund size as our control group.

This paper is contributes to three broad strands of the literature: fund family, tax burden, and ETF administration. First, the fund family literature has examined the impact of multiple funds operating within a family structure. It is well known that mutual funds benefit from economies of scale from organizing as a family (Baumol et al. (1989), Khorana and Servaes (2012), Khorana and Servaes (1999)) and by offering product differentiation families can appeal to investor heterogeneity (Massa (2003)), but limited evidence that families pass the savings on to investors (Freeman and Brown (2001)). Despite potential benefits, family relationships may lead managers to act in the best interest of the overall family, rather than the individual fund. To date the literature suggests that families engage in cross-fund subsidization by favoring high-fee and high performing mutual funds (Gaspar, Massa, and Matos (2006)) or providing liquidity to family funds under stress to benefit the family as a whole (Bhattacharya, Lee, and Pool (2013)). We add to the fund family literature by studying the incorporation of ETFs, a distinct investment vehicle, into mutual fund families. Our results identify a new cross-fund subsidization feature found in the OEF-ETF twin structure, while showing that offering ETFs does not lead to cannibalization of flows for higher-fee index funds.

The second strand of related literature is on the influence of taxes. Generally, papers find that taxes are an important consideration for mutual fund managers and investors alike. Dickson and Shoven (1993) document that capital gains distributions significantly impact after-tax returns, while Sialm and Zhang (2015) show that tax-efficient asset management generates superior before- and after-tax performance. Bergstresser and Poterba (2002) find that money flows to funds that are able to deliver lower tax burdens relative to funds with similar pre-tax returns. Sialm and Starks (2012) find that mutual fund managers consider the tax status of their investors when determining distributions and holdings. Beyond, realized capital gain, Barclay, Pearson, and Weisbach (1998) document empirically and theoretically that unrealized capital gains are an important consideration for fund managers. By offering ETFs, we find that traditional mutual fund families are able to mitigate a major externality imposed on tax sensitive investors.[9]

Finally, as the assets under management by ETFs have increased, so too has the academic community's interest in the investment vehicle. To date, the majority of studies have examined the impact of ETF membership on various characteristics of the underlying. ETFs are shown to decrease liquidity (Hamm (2011)), increase volatility (Ben-David, Franzoni, and Moussawi (2014)), and lead to greater co-movement (Da and Shive (2013)) for constituent stocks. Dannhauser (2016) shows that ETF constituency lower corporate bond yields due to a migration of liquidity traders from the underlying market to the basket security. Two recent papers have focused on ETF administration. Specifically, Blocher and Whaley (2014) document that ETFs generate significant revenues from security lending. In perhaps the most closely related paper to ours, Cheng, Massa, and Zhang (2014) investigate the consequences of ETFs affiliated with banks. They find that ETFs leverage information from banks' lending activities, thereby helping the banks' own mutual funds through cross-trading, while supporting the banks' stock price. Importantly, their study focuses on European ETFs, which are synthetically replicated using swaps.

2. Data and Summary Statistics

This section summarizes the sources for our data, describes the methodology used to identify mutual fund and ETF twins, defines the tax burden, and presents summary statistics.

[9] Mutual fund family cross-subsidizations (Gaspar, Massa, and Matos (2006) and Bhattacharya, Lee, and Pool (2013)); side-by-side management arrangements (Nohel, Wang, and Zheng (2010), Cici, Gibson, and Moussawi (2010), and Evans and Fahlenbrach (2012)).

2.1. Data

Our annual data covers U.S. equity index mutual funds and ETFs over the period 1997 to 2014 from the CRSP Survivor-Bias-Free U.S. Mutual Fund database. The CRSP database includes mutual fund characteristics such as fund returns, assets under management, fund dividends, long-term and short-term capital gains distributions, fees, and investment objectives. Since, nearly all ETFs are passive investment vehicles, we restrict the sample to include only index funds. In particular, we use the *index_fund_flag* variable of CRSP to retain pure index funds or index-based funds. Since many funds contain multiple share classes, we aggregate variables by asset-weighting the individual share classes. The only variables that are summed over individual share classes are total assets and total 2014 unrealized capital gains, which we obtain from Morningstar Direct. We account for the share class features of Vanguard that includes an ETF under the same fund as the mutual fund classes. To do so, we aggregate at the fund and structure level, resulting in a mutual fund and an ETF observation for a single fund. We also exclude fund liquidation years and require that a fund has a positive tax burden for at least one year of our sample.[10]

We merge the CRSP data with data from Morningstar Direct on fund CUSIP. From the sample of CRSP index funds we further refine the data by eliminating all inverse and enhanced funds using the Morningstar category and fund name. We find the index followed by the remaining funds from three sources. First, we use Morningstar Direct to obtain the benchmarks by matching on CUSIP, then by ticker. For the funds with no Morningstar entry, we develop an algorithm to identify commonly followed indices using the mutual fund names. For instance, we search fund names for keywords, such as, S&P 500, MidCap 400 Value, or Russell 2000 Growth, to identify the benchmarks

[10] Results are robust to the inclusion of liquidation years and funds without distributions.

S&P 500, S&P MidCap 400 Value Index, and Russell 2000 Growth Index, respectively. Finally, we hand collect index benchmark data for approximately 500 funds using prospectuses. We delete any funds without an identifiable benchmark. For funds with hand-collected benchmarks, we use the Morningstar data to assign a Morningstar category on benchmark.

Finally, the time series of tax rates on dividend, short-term, and long-term capital gains are sourced from the National Bureau of Economic Research (NBER).[11] The marginal rates are only available up to 2013 as of February 26, 2015. As a consequence, our tax burden tests are run for 1997 to 2013 despite having information on the distribution yields for 2014.

2.2. Identification of Mutual Fund and ETF Twins

A mutual fund and ETF are identified as twins if they are in the same fund family and follow the same index. Essentially, twins are the exact same fund with different structures. For the majority of management company and index combinations this selection process results in a single mutual fund and a single ETF. However, in select cases a fund family may run many mutual funds following the same benchmark. For instance, Vanguard has an ETF, and three open-ended funds (an index fund, an institutional fund, and a tax-managed fund) that all follow the S&P 500. In these instances, we consider all three mutual funds as broad matches. However, in each case of multiples we identify the true twin as either the oldest mutual fund, or in the case of Vanguard, the mutual fund that contains the ETF as a share class. In total we identify seventy-four twin combinations composed of 152 distinct funds. The correlation of mutual fund and ETF twin returns pre and post expenses are 99.90% and 99.86%, respectively. Correlations that are essentially one confirm our claims that twins are identical funds distinguished only by their structure. Of the seventy-four

[11] http://www.nber.org/~taxism

twins, forty-eight are administered by Vanguard. While Vanguard is a dominant manager in this study, we argue that it is important to understand how this family operates since their assets at the end of 2014 were in excess of the entire hedge fund industry.[12] Furthermore, we categorize the twins based on the timeline of introduction. Thirty-one of the twin fund combinations began as an open-end mutual fund, with the ETF added to the family at least two years after the mutual fund introduction. Eight began as ETFs with the mutual fund following at least two years after. The remaining thirty-five twins introduced the ETF and mutual fund in the same year.

We further restrict the sample to include only funds in the same Morningstar category as the twins. We also delete State Street's S&P 500 funds, so that the SPDR S&P 500 ETF, commonly referred to by its ticker, SPY, does not dominate our ETF results. We are left with 8,503 fund-years representing 617 ETFs and 437 mutual funds. Over 13% of the sample fund-months are associated with a twin arrangement.

2.3. Tax Burden

Mutual funds and ETFs are both registered under the Investment Company Act of 1940, making them pass through entities. Thus on an annual basis, funds distribute capital gains and dividend income to shareholders. If an investment company distributes all of its investment income to its shareholders, the company itself will have no tax liability. The distributions are taxable for investors who hold the funds in a taxable account. Dickson, Shoven, and Sialm (2000) describe this as an externality, since the tax burden of mutual fund investors depends on the behavior of others. Due to their liberal use of in-kind rather than cash distributions, ETFs are considered a more tax efficient alternative to traditional mutual funds. Traditional mutual funds have the ability to use in-kind

[12] http://www.wsj.com/articles/vanguard-sets-record-funds-inflow-1420430643

redemption, but rarely if ever utilize the feature. At the fund level, the U.S. Tax Code relieves the ETF from any tax consequences related to the distribution of appreciated assets when made in response to an investor demand. Since redemptions are met with a tax-free exchange of the basket of securities in return for the ETF shares, the fund does not incur capital gains which would need to be distributed to investors. In fact, ETF sponsors rarely need to transact in the underlying, thereby reducing the potential of incurring taxable capital gains.

We follow Bergstresser and Poterba (2002) and Sialm (2009) in computing the tax burden ($TB_{f,t}$) for fund, f, at time, t. The tax burden is the sum of the marginal investor's tax liabilities on the fund's distribution yields. Specifically, the tax burden is defined as

$$\begin{aligned} TB_{f,t} &= \left(\frac{DIV_{f,t}}{NAV_{f,t-1}} * Tax_t^{DIV}\right) + \left(\frac{LCG_{f,t}}{NAV_{f,t-1}} * Tax_t^{LCG}\right) + \left(\frac{SCG_{f,t}}{NAV_{f,t-1}} * Tax_t^{SCG}\right) \\ &= \left(Y_{f,t}^{DIV} * Tax_t^{DIV}\right) + \left(Y_{f,t}^{LCG} * Tax_t^{LCG}\right) + \left(Y_{f,t}^{SCG} * Tax_t^{SCG}\right), \end{aligned} \tag{1}$$

where $Y_{f,t}^{DIV}$, $Y_{f,t}^{LCG}$, and $Y_{f,t}^{SCG}$ are the fund's dividend yield, long-term capital gains yield, and short-term capital gains yield, respectively. Tax_t^{DIV}, Tax_t^{LCG}, and Tax_t^{SCG} are the tax rates on dividends, and long-term and short-term capital gains. Tax rates are computed as the weighted averages of the marginal tax rates of investors in different tax brackets, where the weights correspond to the amount of declared dividends and capital gains (Feenberg and Coutts (1993)).

The CRSP mutual fund database provides the level of distributions per share by type. Dividend distributions by mutual funds are made net of fund expenses. By definition, short-term gains are for investments held by the fund for less than one year. Long-term capital gains are generally on positions held for over a year. If the term of the capital gain is not specified, we follow Sialm and Starks (2012) and assume it is a long-term distribution. Only taxable dividends are considered, as taxes are not charged for untaxed or tax-exempt dividends. The tax burden can be interpreted as the portion of the fund's previous value that an investor pays in taxes.

We are also interested in the unrealized capital gains disclosed in annual reports. Unrealized capital gain, referred to as the tax overhang, is equal to the cumulative price appreciation of the fund net of distributions. And in expectation, lower capital gains overhang will reduce future taxable fund distributions. We collect 2014 reported unrealized capital gains at the share class level from Morningstar Direct. Unfortunately, Morningstar Direct does not retain historical values of this variable. Although a time-series would be preferred, since unrealized capital gains is cumulative we believe that the results using only 2014 data are valid. Therefore, we compute, $UNR_{f,2014}$, as the total reported unrealized capital gains in fund f as a percentage of total fund assets. We also considered computation of unrealized capital gains burdens computed in the literature by Bergstresser and Pontiff (2013), Odean (1998) and Sialm and Starks (2012). However, each of these measures relies on an assumption about the accounting method employed by a fund. Generally, they assume "smart" tax realization strategies, with the highest-basis shares sold first. This assumption would bias us against finding results, as the use of ETFs by mutual funds would raise the possibility of a change in strategies. Therefore, we use reported unrealized capital gains.

2.4. Summary Statistics

Table 1 presents the summary statistics for our data. Panel A documents the number of mutual funds, ETFs, and twins of both types in each year of the study. The first twin arrangement was created by the first ETF, SPY, in 1993, which is excluded from our data. In our setting, the first twin was introduced in 2000. The number of ETF twins has grown from one in 2000 to 68 in 2014.

[Insert Table 1]

Panel B of table 1 contains the mean value of different characteristics of the funds in our study. The panel shows that twins are generally created by larger families and have lower overall expense ratios. Mutual fund and ETF twins also have more assets under management and lower tax burdens than their non-twin peers. Interestingly, the average unrealized capital gains for mutual fund twins is lower than non-twin mutual funds, but for ETF twins it is significantly greater. These summary results on the capital gains overhang support the intuition provided by Barclay, Pearson, and Weisbach (1998). The lower unrealized capital gains of mutual funds attract new investors, while ETF managers bear the higher capital gains overhang as the capital gains are unlikely to be realized.

3. Empirical Analysis of Mutual Fund and ETF Twins

In this section we study the implications of a twin relationship for investors. In particular we examine the consequences of side-by-side management on the total tax burden, the long-term and short-term capital gains yields, the capital gains overhang, the overall expense ratio, the 12b-1 distribution fees, the management fees, and annual flows. To address endogeneity concerns related to the introduction of new investment vehicles, we execute two identification strategies: a fixed effects model and a difference-in-difference specification. The following subsections discuss each of the methodologies and results in greater detail.

3.1. Fixed Effects Model

Our fixed effects model uses both Morningstar category and year fixed effects to account for unobserved style related heterogeneity (Bergstresser and Pontiff (2013)), and secular trends over our sample period. In this model, we regress one of our dependent variables on three binary variables of interest and covariates specified in the literature. The independent variables that we consider are the overall tax burden, $TB_{f,t}$, the long-term capital gains yield, $Y_{f,t}^{LCG}$, and the short-term capital gains yield, $Y_{f,t}^{SCG}$. We are also interested in the unrealized capital gains of the fund measured as a percentage of total assets, $UNR_{f,2014}$.

Additionally, in our study of fees paid by investors, we consider the overall expense ratio, $Expense_{f,t}$, and its major components, the distribution fee, $12b1_{f,t}$, and the management fee, $Mgmt_{f,t}$. In particular, we execute the following specification

$$Y_{f,s,t} = \alpha + \beta_1 MF_f + \beta_2 MF\ Twin_{f,t} + \beta_3 ETF\ Twin_{f,t} + \beta_4 X_{f,t} + \eta_s + \tau_t + \varepsilon_{f,s,t}, \qquad (2)$$

where , $Y_{f,s,t}$, is the value of one of our dependent variables for fund f in style s in year t. MF_f is equal to one for a mutual fund and zero for an ETF. This dummy variable accounts for differences between the tax distributions, fees, and flows of mutual funds relative to ETFs. $MF\ Twin_{f,t}$ is set to one for all years that a mutual fund operates as a twin, and $ETF\ Twin_{f,t}$ is set to one for all years that an ETF operates as a twin. Both values are zero otherwise. For contemporaneous twins, $MF\ Twin_{f,t}$ and $ETF\ Twin_{f,t}$ are always equal to one. For twins where the mutual fund or ETF existed on its own for at least two years, the covariate is equal to zero until the twin is introduced. These two dummy variables measure the difference between the outcome variable for twins relative to that of their peers of the same investment vehicle. For instance, $MF\ Twin_{f,t}$ shows a differential effect of mutual funds operating with a twin relative to that of all stand-alone mutual funds. $X_{f,t}$ is the vector of controls including the log of fund size, the log of family index assets, and fund age (in years). η_s is the Morningstar category fixed effect. Since many of our twin funds are contemporaneous starts, we do not use fund fixed effects. Doing so would exclude the contemporaneous twins from our study because there is no variation in their treatment status. τ_t is the year fixed effect.

3.1.1. Tax Effect Results

Table 2 presents the results of the tax fixed effects panel regressions. The first column reports results without the two twin dummies to first identify whether ETFs are truly more tax efficient as frequently claimed by practitioners. For the remaining columns, the dependent variable is overall tax burden for columns 2 and 3, the long-term capital gains yield for columns 4 and 5, the short-term capital gains yield for columns 6 and 7, and the unrealized capital gains for columns 8 and 9. For each of the regression pairs, the latter column includes the covariates discussed above.

[Insert Table 2]

In column 1, the coefficient on the mutual fund dummy, MF, is positive and significant, empirically confirming that mutual funds have higher tax burdens than ETFs. The coefficient on $MF\ Twin$ in both columns 2 and 3 is negative and significant, suggesting that mutual funds are able to use their ETF twin to lower the tax burdens. After controlling for observables, the coefficient on $ETF\ Twin$ is insignificant, suggesting that the overall tax burden is unaffected by this family relationship.

We examine two of the components of tax burden to determine the driver of the relationship; long- and short-term capital gains yield. We do not consider the dividend yield because index funds do not have control over the constituents and their dividend policy. Furthermore, dividend distributions are made net of expenses. As expected, we find that the lower overall tax burden is being driven entirely by a reduction in the long-term capital gains yield for mutual fund twins. Also, the results indicate that the arrangement leads to higher long-term capital gains for ETF investors.

The short-term capital gains yield of both mutual fund and ETF twins is unaffected by the relationship.

3.1.2. The Expense Effect

Having shown above that mutual funds benefit from having an ETF twin, we now examine if any economies of scale from managing two identical pools of assets are passed on to investors in the form of lower fees. Table 3 shows the findings of our fixed effects panel regressions. As above, column 1 excludes the twin dummies to document a general difference in expenses for mutual funds and ETFs. Columns 2 and 3 have the total expense ratio as the dependent variable. Columns 4 and 5 use 12b-1 fees as the left-hand variable. 12b-1 fees are expenses charged directly to the assets of the fund for marketing and distribution. Finally, columns 6 and 7 use management fees as the dependent variable. The management fee is paid directly to the fund's advisor, and as claimed by Wahal and Wang (2011), is a clean measure of the price of the services provided by the advisor.

[Insert Table 3]

Column 1 confirms that ETFs generally have lower expense ratios than their mutual fund peers, which supports general market commentary on ETFs relative to mutual funds. The lower expense ratio is often attributed to the relative ease of ETF management since the managers themselves rarely trade in the underlying markets. Column 2 suggests that investors in mutual fund twins may have lower total expense ratios. However, the result is only robust for ETF twins after the inclusion of the controls of fund size, family size, and age. In column 3, we find that ETF twins have lower expense ratios of 8.6 basis points, which is a 20% decrease for the average ETF. Examining the main components of the overall expense ratio in columns 5 and 7, we see that mutual fund twins' 12b-1 fees are higher and management fees are lower than their non-twin mutual fund peers. Conversely, ETF twins have statistically higher 12b-1 fees and lower management fees.

3.1.3. *Flows Results*

Next we consider if two identical asset pools in two different vehicles has an impact on the flows to the funds. We compute flows following Sirri and Tufano (1998) as,

$$Flow_{f,t} = \frac{TNA_{f,t} - TNA_{f,t-1} * (1 + R_{f,t})}{TNA_{f,t-1}} \tag{3}$$

where $TNA_{f,t}$ is the total net assets and $R_{f,t}$ is the total return of fund f in period t. We compute flow in two ways. First, we compute the average monthly flow in a year. Second, we compute the total annual flow. We winsorize both measures at the 1% level to mitigate the impact of outliers. Table 4 presents the results of the fixed effects regression with our flow measures the dependent variables. In Columns 1 through 3, we use the average monthly measure, and in columns 4 through 6, the total annual measure. Columns 2 and 5 include the covariates specified in equation (2), while columns 3 and 6 add the annual lagged flow measure.

[Insert Table 4]

The mutual fund and ETF twin dummies are positive and statistically significant regardless of the model. These results suggest that fund families are able to generate significant inflows from having index twins. Of note, the average annual flows to mutual fund twins are 50% greater than that of their non-twin peers. For ETF twins, annual flows are 25% greater. These results suggest that fund families are able to drive flows to their more profitable mutual funds using the twin structure. These results support the findings from Massa (2003) that product differentiation, which appeals to different investor needs, drives fund proliferation and flows.

3.2. *Difference-in-Difference: Mutual Fund First*

To provide additional evidence of the impact of side-by-side management on the tax distributions to and expenses paid by investors, we exploit the historical differences between mutual funds and ETFs. The first index mutual fund was introduced in 1971 by Wells Fargo Bank, while the first ETF was launched more than twenty years later by State Street Global Advisors. The relative maturity of mutual funds allows us to conduct a difference-in-difference test as many mutual funds have an established record prior to the introduction of their ETF twin.

In this setting, we use mutual funds that operated for at least two years prior to the introduction of their ETF twin as the treatment group. We denote the year that the ETF is introduced as year zero. The time frame for these tests is two years before and two years after a twin introduction, excluding the introductory year. We exclude year zero to allow for gradual adoption of the ETF by investors and learning by the twin mutual fund advisor. We limit observations to only two years on either side of treatment for two reasons. First, doing so allows us to increase our sample size. Second, limiting the number of periods mitigates concerns related to serial correlation bias from difference-in-difference studies as discussed in Bertrand, Duflo, and Mullainathan (2004). We develop the control sample of mutual funds without a twin by matching on Morningstar category and fund size in the year before treatment occurs. Our matching strategy requires the control sample to be in the same Morningstar category, and then we use both the two and three nearest neighbors matched on fund size. We use two and three nearest matches to balance the trade-off between bias and variance associated with nearest neighbor matching. We match with replacement contingent on a control having multiple matches not having overlapping pre- and post-periods. Furthermore, treatment funds are eligible controls for the years prior to the two pre-periods. Our regressions are as follows:

$$Y_{f,t} = \alpha_f + \tau_t + \beta_1(MF\ Twin_f * Post_{f,t}) + \beta_2 Post_{f,t} + \beta_3 X_{f,t} + \varepsilon_{f,t} \tag{4}$$

$Y_{f,t}$ is one of the dependent variables detailed above. α_f is a fund fixed effect. In this setting we are able to use fund fixed effects since contemporaneous twins are not included. This fixed effect allows us to control for a fund's different characteristics, such as index or family, which would lead to different propensities to generate taxable distributions or to adjust expenses. τ_t is the year fixed effect. The interaction, $MF\ Twin_f * Post_{f,t}$, is the covariate of interest. The coefficient on this variable, β_1, identifies the effect the ETF introduction had on the twin relative to its matched controls. $X_{f,t}$ is the vector of controls that vary at the fund-year level: the log of fund size, the log of family size, and fund age (in years).

3.2.1. Tax Effect Results

In Table 5 we present the results for the difference-in-difference regressions with the tax variables as the outcome variables. Panel A presents the results using two nearest neighbors matching and Panel B uses the three nearest neighbors. The first and second columns of both panels show the results for the overall tax burden, the third and fourth for the long-term capital gains yield, and the fifth and sixth for the short-term capital gains yield.

[Insert Table 5]

The coefficient on the interaction term, $MF\ Twin_f * Post_{f,t}$, is negative and significant for the overall tax burden regressions. Figure 1 provides a graphical representation of the effect of the twin initiation on the overall tax burden. In this figure we plot the average tax burden for the treatment and three nearest neighbors control groups in the four years before and after the introduction of the ETF twin, as well as the year of introduction; year zero. The figure confirms the critical assumption of common trends prior to treatment and documents the divergence following the introduction. Furthermore, it appears that managers did require time to learn how to maximize the new family dynamics.

[Insert Figure 1]

The last four columns of Table 5 demonstrate that a reduction in long-term capital gains yields is the source of the lower tax burden for mutual funds. In particular, the long-term capital gains yield for the treatment funds is 1.95% lower in the two years following the introduction of the ETF twin. The short-term capital gains yield is unchanged following treatment. Figure 2 presents a graphical representation of the long-term capital gains effect. The plot shows that mutual fund twins drive their long-term capital gains distributions to zero following the introduction of their ETF twins.

[Insert Figure 2]

Overall, these results confirm the findings of the fixed effects regressions; mutual funds are able to use the co-existence of a twin ETF to lower the long-term capital gains distributions to investors, thus lowering the total fund's tax burden.

3.2.2. The Expense Effect

We are able to use the difference-in-difference specification to examine whether mutual funds adjust their fees following the introduction of the ETF twin. Since, our specification uses two years after the treatment event, this allows for the typical mutual fund board, which meets annually, at least three opportunities to adjust fees. Table 6 shows the expense results of the difference-in-difference specification where the dependent variables are expenses.

[Insert Table 6]

In every column the coefficient on the interaction variable of interest is insignificant and nearly equal to zero. These findings confirm those of the fixed effects regressions that mutual fund investors do not receive any reduction in fees from economies of scale. Specifically, this table shows that mutual funds do not have significantly different expense ratios, 12b-1 fees, or management fees in the two years following their ETF twin introduction relative to the matched control funds.

4. Robustness

In this section we conduct robustness tests related to the tax effects results. We first conduct a falsification test of our difference-in-difference specification that rolls treatment back two years. This test confirms that the assumption of common trends is satisfied and that the ETF introduction drives the lower tax burden. Finally, we run our regressions using capital gains distribution levels rather than yields to confirm that returns are not driving our results.

4.1. Difference-in-Difference Falsification Test

We perform a falsification test to address concerns of differential pre-treatment trends driving the tax results. This test confirms the essential assumption of difference-in-difference identification of common pre-treatment trend. We estimate the same model with fund and year fixed effects as in Tables 5 and 6, but move the treatment period two years earlier. This test focuses on the differential effect of twin versus non-twin mutual funds two years prior to the introductory event. In this test, the pre-period is now three and four years prior to the true introduction. The post-period is now the year before and the year of introduction. Table 7 reports the results of this falsification test.

[Insert Table 7]

We find that none of the coefficients reported are statistically significant. These results confirm that the two groups have common trends in the pre-period examined in the initial difference-in-difference test. Furthermore, the results validate our intuition that the twin introduction is the main driver behind the identified change in taxable distributions.

4.2. Capital Gain Distribution Levels

The previous tests relied on yield measures, raising concerns that differences in returns to the index funds are contributing to the tests. In the same way that dividend yield is preferred to dividend levels, the conversion to yields accounts for the size of a distribution relative to the investment. Nevertheless, we address these concerns using the level of capital gains distributions on the left-hand side. As discussed in Sialm and Zhang (2015), dividend distributions are made after expenses are paid to the fund. This netting not only highlights the importance of standardizing by NAV, which is also a net value, but also makes interpretation of the total tax burden and dividend levels difficult. Table 8 presents the results of these tests using long-term and short-term distributions levels as the dependent variables.

[Insert Table 8]

The coefficients on the twin variables confirm the results of those in Table 2. In particular, the level of long-term capital gains is statistically higher for mutual fund twins than for non-twin mutual funds (16 basis points). For ETF twins there is no difference in the level of distributions.

5. Conclusion

As ETFs become an increasingly important investment alternative for institutional and retail investors, fund families have been forced to adapt. To retain old investors and attract new investors, many mutual fund families have begun to offer mutual fund and ETF twins. Twins are comprised of an ETF and mutual fund in the same family that follow the same index. Essentially, these funds are identical pools of assets that differ only in their investment structure.

We show that inclusion of an ETF as a twin in a mutual fund family has significant implications for investors and the family. In particular, mutual fund investors benefit from lower tax burdens, long-term capital gains distributions, and unrealized capital gains overhang. For ETF twin investors, the overall tax burden is unaffected, but there is some evidence that long-term capital gains distributions are greater. We also show that fund families do not fully pass along economies of scale from this arrangement, as only ETF twins have lower expense ratios. Finally, we show that twin funds experience greater inflows than their non-twin peers, suggesting that this form of product differentiation is beneficial to the family as a whole, and is a determinant for fund families to introduce products that cater to heterogeneous investor preferences.

References

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of financial Economics*, *17*(2), 223-249.

Barclay, Michael J, Neil D Pearson, and Michael S Weisbach, 1998, Open-end mutual funds and capital-gains taxes, *Journal of Financial Economics* 49, 3–43.

Baumol, W., S.M. Goldfeld, L.A. Gordon, and F.M. Köhn, 1989, *TheEconomics Mutual Fund Markets: Competition Versus Regulation* (Kluwer Academic).

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2014, Do ETFs Increase Volatility?, National Bureau of Economic Research.

Bergstresser, Daniel, and Jeffrey Pontiff, 2013, Investment taxation and portfolio performance, *Journal of Public Economics* 97, 245–257.

Bergstresser, Daniel, and James Poterba, 2002, Do after-tax returns affect mutual fund inflows?, *Journal of Financial Economics* 63, 381–414.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How Much Should We Trust Differences-in-Differences Estimates?, *Quarterly Journal of Economics* 119.

Bhattacharya, Utpal, Jung H Lee, and Veronika K Pool, 2013, Conflicting family values in mutual fund families, *Journal of Finance* 68, 173–200.

Blocher, Jesse, and Robert E Whaley, 2014, Passive Investing, *Available at SSRN 2474904*.

Cheng, Si, Massimo Massa, and Hong Zhang, 2014, The Dark Side of ETF Investing: A World-Wide Analysis, .

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

Cici, Gjergji, Scott Gibson, and Rabih Moussawi, 2010, Mutual fund performance when parent firms simultaneously manage hedge funds, *Journal of Financial Intermediation* 19, 169–187.

Da, Zhi, and Sophie Shive, 2013, When the bellwether dances to noise: Evidence from exchange-traded funds, .

Dannhauser, Caitlin D, 2016, The Impact of Innovation: Evidence from Corporate Bond ETFs, *Unpublished Working Paper*.

Dickson, Joel M, and John B Shoven, 1993, Ranking Mutual Funds on an After-Tax Basis, *NBER Working Paper*.

Dickson, Joel M, John B Shoven, and Clemens Sialm, 2000, Tax Externalities of Equity Mutual Funds, *National Tax Journal* 53, 607–28.

Evans, Richard B, and Rüdiger Fahlenbrach, 2012, Institutional investors and mutual fund governance: Evidence from retail–institutional fund twins, *Review of Financial Studies* 25, 3530–3571.

Feenberg, Daniel, and Elisabeth Coutts, 1993, An introduction to the TAXSIM model, *Journal of Policy Analysis and Management* 12, 189–194.

Freeman, John P, and Stewart L Brown, 2001, Mutual Fund Advisory Fees: The Cost of Conflicts of Interest, *Journal of Corporation Law* 26, 609.

Gaspar, JOSÉ-MIGUEL, Massimo Massa, and Pedro Matos, 2006, Favoritism in Mutual Fund Families? Evidence on Strategic Cross-Fund Subsidization, *Journal of Finance* 61, 73–104.

Hamm, Sophia, 2011, The effect of ETFs on stock liquidity, *Available at SSRN 1687914*.

Khorana, Ajay, and Henri Servaes, 1999, The determinants of mutual fund starts, *Review of Financial Studies* 12, 1043–1074.

Khorana, Ajay, and Henri Servaes, 2012, What drives market share in the mutual fund industry?, *Review of Finance* 16, 81–113.

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.

Nanda, V., Narayanan, M. P., & Warther, V. A. (2000). Liquidity, investment ability, and mutual fund structure. *Journal of Financial Economics*, *57*(3), 417-443.

Nohel, Tom, Z Jay Wang, and Lu Zheng, 2010, Side-by-Side Management of Hedge Funds and Mutual Funds, *Review of Financial Studies* 23, 2342–2373.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *Journal of Finance* 53, 1775–1798.

Poterba, James M, and John B Shoven, 2002, Exchange-Traded Funds: A New Investment Option for Taxable Investors, *American Economic Review* 92, 422–427.

Sialm, Clemens, 2009, Tax Changes and Asset Pricing, *American Economic Review* 99, 1356–83.

Sialm, Clemens, and Laura Starks, 2012, Mutual fund tax clienteles, *Journal of Finance* 67, 1397–1422.

Sialm, Clemens, and Hanjiang Zhang, 2015, Tax Efficient Asset Management: Evidence from Equity Mutual Funds, .

Sirri, Erik R, and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589–1622.

Wahal, S., & Wang, A. Y. (2011). Competition among mutual funds. *Journal of Financial Economics*, *99*(1), 40-59.

Figure 1: Overall Tax Burden of a Mutual Fund around the Introduction of an ETF Twin

Plotted below is the average of the overall tax burden computed following Bergstresser and Poterba (2002), Sialm (2009). MF Twins are index mutual funds whose family introduced an ETF fund that follows the same benchmark. The control funds are stand-along mutual funds matched on Morningstar category and fund size the year prior to the twin introduction. Year 0 is the year of introduction.

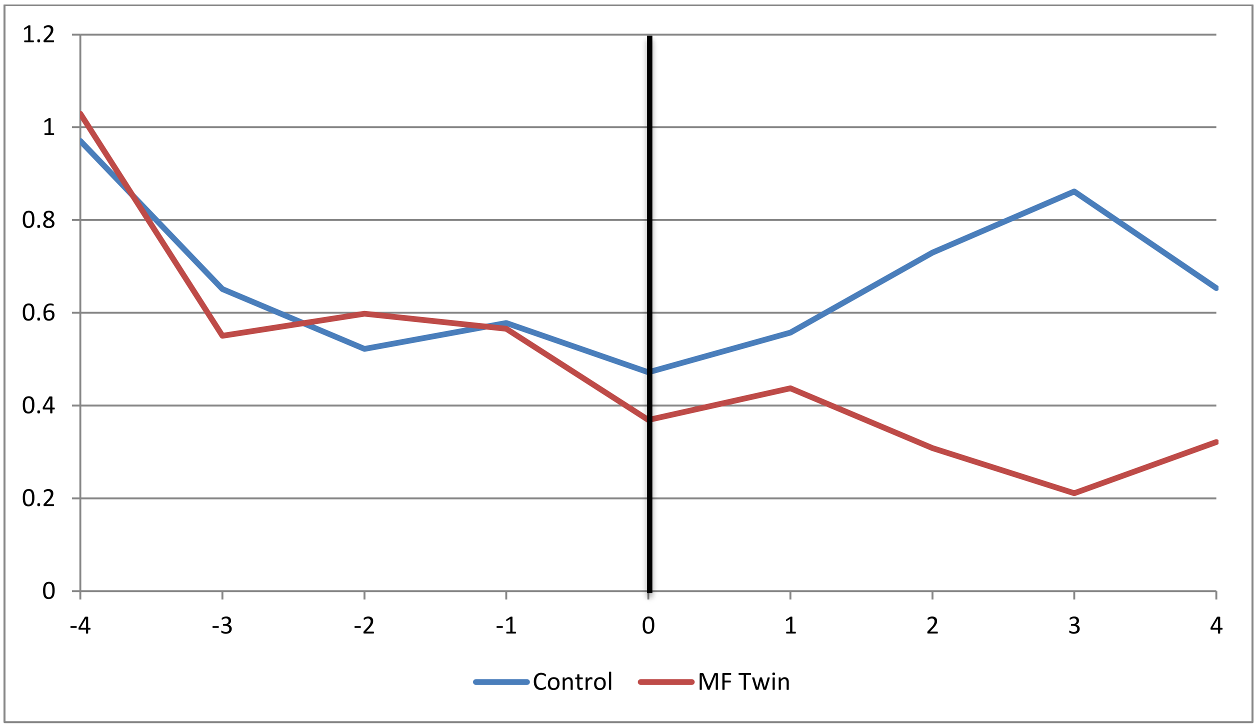

Figure 2: Long-Term Capital Gains of a Mutual Fund around the Introduction of an ETF Twin

Plotted below is the average of the long-term capital gains yield of a mutual fund, computed as the annual long-term capital gains distributions scaled by lagged net asset value. MF Twins are index mutual funds whose family introduced an ETF fund that follows the same benchmark. The control funds are stand-alone mutual funds matched on Morningstar category and fund size the year prior to the twin introduction. Year 0 is the year of introduction.

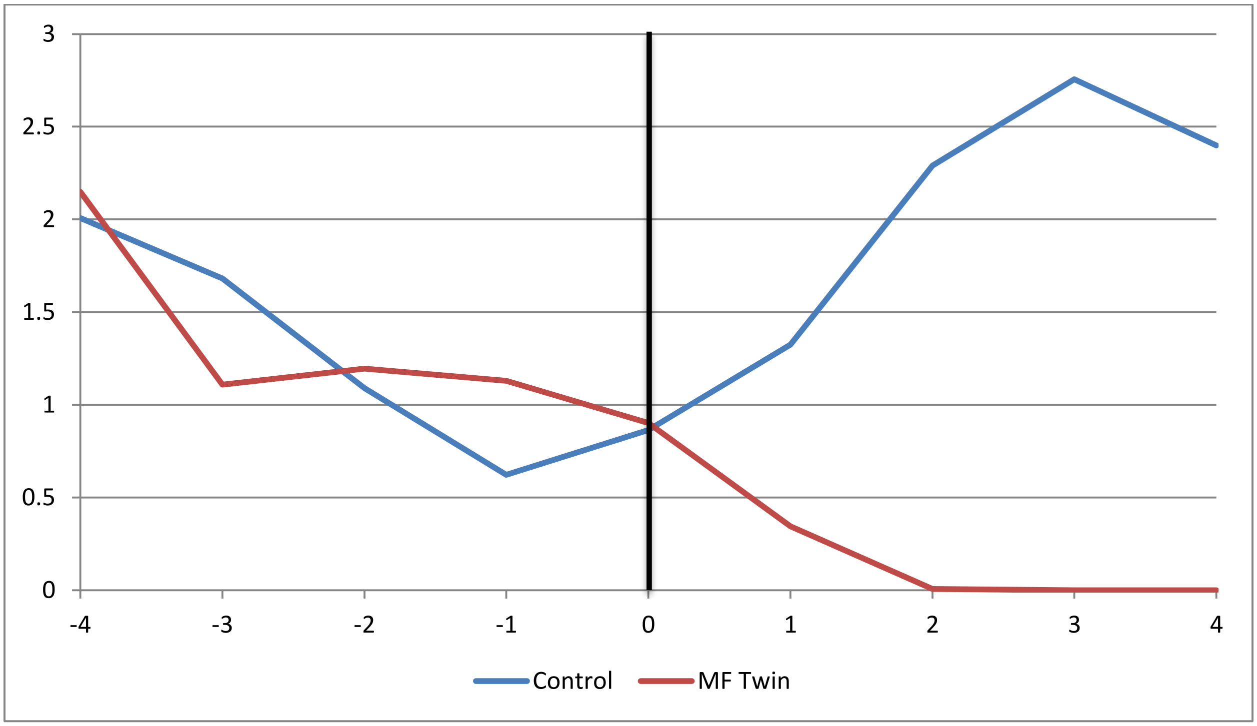

Table 1: Summary Statistics

Summary statistics by investment vehicle type and by twin status for index funds for the annual sample period of 1997 to 2014. Twins are mutual funds and ETFs that operate in the same fund family and follow the same index. Panel A presents the number of observations of mutual funds, ETFs, and twin funds by year. Twin funds are further broken down into mutual fund and ETF twins. Panel B presents the mean of observable summary startistics for the different categories. *Fund Size* is the dollar of assets in millions for all share classes of the fund. *Family Size* is the total index assets under management by the fund family. *Age, Expense Ratio*, *12b-1 Fee*, *Management Fee*, *Tax Burden*, *Dividend Yield*, *Long-Term Capital Gains (LCG) Yield*, and *Short-Term Capital Gains (SCG) Yield* are all the asset-weighted values for all share classes of a fund. *Unrealized Gains* is the total 2014 reported unrealized capital gains of all share classes of a fund scaled by total fund assets.

Panel A: Number of Fund Types Per Year

Year	# MF	# ETF	# Twins	#MF Twins	# ETF Twins
1997	86	1	0	0	0
1998	101	7	0	0	0
1999	138	8	0	0	0
2000	181	47	3	2	1
2001	201	66	8	5	3
2002	219	72	8	5	3
2003	215	80	10	6	4
2004	218	109	39	20	19
2005	207	145	46	24	22
2006	211	222	66	34	32
2007	227	292	77	40	37
2008	285	359	89	48	41
2009	280	386	90	48	42
2010	297	443	116	60	56
2011	319	508	141	75	66
2012	314	546	142	76	66
2013	309	523	147	81	66
2014	319	562	151	83	68
Total	4,127	4,376	1,133	607	526

Panel B: Characteristics of the Average Fund by Type

	Overall	MF	MF Twin	ETF	ETF Twin
Fund Size	1929	2534	7873	1359	3221
Family Size	199717	137128	688198	258529	734772
Age	6.89	9.06	9.31	4.85	4.95
Expense Ratio (%)	0.47	0.54	0.34	0.42	0.18
12b-1 Fees (%)	0.04	0.08	0.06	0.00	0.02
Management Fee (%)	0.23	0.16	0.17	0.29	0.13
Tax Burden (%)	0.52	0.60	0.48	0.45	0.39
Dividend Yield	1.48	1.24	1.54	1.70	1.57
LCG Yield (%)	0.59	1.18	0.18	0.03	0.01
SCG Yield (%)	0.18	0.33	0.26	0.04	0.00
Unrealized Gains (%)	11.53	22.69	16.18	6.71	13.05
Avg Monthly Flows (%)	3.40	2.86	4.14	3.91	4.42
Annual Flows (%)	32.91	26.24	37.92	39.46	53.36

Table 2: Tax Fixed Effects Panel Regressions

This table presents the results of the annual fixed effects panel regression

$$Tax_{f,s,t} = \alpha + \beta_1 MF_f + \beta_2 MF\ Twin_{f,t} + \beta_3 ETF\ Twin_{f,t} + \beta_4 X_{f,t} + \eta_s + \tau_t + \varepsilon_{f,s,t}$$

where $Tax_{f,s,t}$ is one of four tax measures for fund f in Morningstar Category s in year t. The measures include total *Tax Burden*, long-term capital gains yield, *LCG*, short-term capital gains yield, *SCG*, and the unrealized capital gains as a percentage of total assets, *UNR*. Tests with *Tax Burden* use data from 1997-2013, for *LCG* and *SCG* from 1997-2014, and *UNR* is only available for 2014. *MF* is equal to one if a fund is a mutual fund. Twins are mutual funds and ETFs in the same fund family that follow the same index. *MF Twin* is equal to one for the years in which a mutual fund has an ETF twin and zero otherwise. *ETF Twin* is equal to one when an ETF has a mutual fund twin and zero otherwise. $X_{i,t}$ includes covariates that change at the style and year level and include *Fund Size* $_{f,t}$ the log of total portfolio assets, *Family Size* $_{f,t}$ the log of total fund family assets, and the age in years of a fund. *Age* $_{f,t}$. η_s is a Morningstar Category fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable:	**Tax Burden**	**Tax Burden**	**Tax Burden**	**LCG**	**LCG**	**SCG**	**SCG**	**UNR**	**UNR**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
MF	0.131***	0.244***	0.174***	1.380***	0.991***	0.403***	0.460***	14.748***	12.150***
	4.44	7.74	4.73	12.92	7.97	5.83	5.36	8.95	8.36
MF Twin		-0.186***	-0.136**	-1.152***	-0.821***	-0.176	-0.144	-6.365***	-6.996***
		-4.29	-2.57	-9.58	-5.35	-1.46	-1.03	-3.61	-3.74
ETF Twin		-0.042*	-0.023	-0.001	0.189**	-0.068*	-0.036	5.136***	4.286***
		-1.75	-0.84	-0.01	2.16	-1.69	-0.85	4.20	3.37
Fund Size	0.011**		0.011*		-0.030		-0.015		-0.452
	2.01		1.92		-1.35		-1.60		-1.41
Family Size	-0.020***		-0.014**		-0.072***		-0.008		0.449*
	-4.11		-2.44		-3.18		-0.81		1.79
Age (Years)	0.009***		0.009***		0.052***		-0.015***		0.766***
	2.82		2.67		3.30		-2.98		3.93
Constant	1.021***	0.859***	0.949***	1.577***	2.460***	0.197**	0.320**	7.280***	-1.017
	7.11	6.59	6.35	2.59	3.48	2.03	2.50	11.97	-0.39
Year FE	Y	Y	Y	Y	Y	Y	Y	N	N
Morningstar Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-sqr	0.174	0.169	0.176	0.169	0.182	0.059	0.066	0.396	0.440
Obs	7,622	7,622	7,622	8,503	8,503	8,503	8,503	805	805

Table 3: Expense Fixed Effects Panel Regressions

This table presents the results of the annual fixed effects panel regression

$$Expense_{f,s,t} = \alpha + \beta_1 MF_f + \beta_2 MF\ Twin_{f,t} + \beta_3 ETF\ Twin_{f,t} + \beta_4 X_{f,t} + \eta_s + \tau_t + \varepsilon_{f,s,t}$$

where $Expense_{f,s,t}$ is one of three tax measures for fund f in Morningstar Category s in year t. The measures include *Exp. Ratio,* the year-end expense ratio, *12b1 Fees*, the fund's 12b1 fee, and *MgmtExp*, a fund's year-end management expense. *MF* is equal to one if a fund is a mutual fund. Twins are mutual funds and ETFs in the same fund family that follow the same index. *MF Twin* is equal to one for the years in which a mutual fund has an ETF twin and zero otherwise. *ETF Twin* is equal to one when an ETF has a mutual fund twin and zero otherwise. $X_{f,t}$ includes covariates that change at the style and year level and include $Fund\ Size_{f,t}$ the log of total portfolio assets, and $Family\ Size_{f,t}$ the log of total fund family assets. η_s is a Morningstar Category fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable:	**Exp. Ratio**	**Exp. Ratio**	**Exp. Ratio**	**12b1 Fees**	**12b1 Fees**	**Mgmt Exp**	**Mgmt Exp**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
MF	0.088***	0.273***	0.085**	0.084***	0.055***	-0.014	-0.006
	2.75	6.12	2.15	9.18	5.82	-0.38	-0.13
MF Twin		-0.317***	-0.017	-0.023*	0.025*	-0.084**	-0.131***
		-5.46	-0.31	-1.70	1.75	-2.19	-2.70
ETF Twin		-0.230***	-0.086***	0.017***	0.038***	-0.157***	-0.192***
		-10.73	-3.65	4.31	6.02	-8.11	-7.36
Fund Size	-0.037***		-0.036***		-0.004*		0.026**
	-6.91		-6.63		-1.95		2.09
Family Size	-0.060***		-0.057***		-0.010***		0.004
	-14.88		-12.73		-5.61		0.68
Age (Years)	0.004*		0.003		0.000		0.002
	1.74		1.47		0.08		0.55
Constant	1.097***	0.298***	1.079***	-0.039***	0.079***	0.174***	-0.019
	20.57	7.03	17.42	-3.16	3.39	3.18	-0.15
Year FE	Y	Y	Y	Y	Y	Y	Y
Morningstar Category FE	Y	Y	Y	Y	Y	Y	Y
R-sqr	0.411	0.230	0.414	0.167	0.214	0.375	0.381
Obs	7,211	7,211	7,211	7,212	7,212	7,103	7,103

Table 4: Flows Effect of Mutual Fund and ETF Twins

This table presents the results of the annual fixed effects panel regression

$$Flow_{f,s,t} = \alpha + \beta_1 MF_f + \beta_2 MF\ Twin_{f,t} + \beta_3 ETF\ Twin_{f,t} + \beta_4 X + \eta_s + \tau_t + \varepsilon_{f,s,t}$$

where $Flow_{f,s,t}$ is one of two flow measures for fund f in Morningstar Category s in year t. The measures includes the average asset weighted monthly flows of all classes of fund f, *Avg. Monthly* ,and the value weighted total annual flows of all classes of fund f, *Annual* . *MF* is equal to one if a fund is a mutual fund. Twins are mutual funds and ETFs in the same fund family that follow the same index. *MF Twin* is equal to one for the years in which a mutual fund has an ETF twin and zero otherwise. *ETF Twin* is equal to one when an ETF has a mutual fund twin and zero otherwise. X includes covariates that change at the style and year level. Contemporaneous controls include $Fund\ Size_{f,t}$ the log of total portfolio assets, $Family\ Size_{f,t}$ the log of total fund family assets, and $Age_{f,t}$ in years. *Flows (Lagged)* is the one year lag of the flow measure. η_s is a Morningstar Category fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable:	**Avg. Monthly**	**Avg. Monthly**	**Avg. Monthly**	**Annual**	**Annual**	**Annual**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
MF	-1.369***	-0.212	-0.011	-18.992***	-5.519	-3.896
	-4.63	-0.62	-0.05	-8.26	-1.52	-1.48
MF Twin	1.875***	2.285***	1.483***	20.075***	20.034***	13.402***
	4.15	4.98	4.44	5.54	4.40	4.33
ETF Twin	0.634**	0.844***	0.581***	17.058***	14.254***	10.748***
	2.42	4.08	3.28	5.00	4.60	4.75
Fund Size		0.077	0.036		3.376***	2.845***
		1.33	0.91		5.14	5.40
Family Size		-0.191***	-0.115***		-1.358***	-1.194***
		-4.86	-3.84		-3.43	-3.40
Age (Years)		-0.340***	-0.157***		-3.631***	-1.944***
		-7.06	-5.68		-6.23	-4.99
Lagged Flows			0.279***			0.186***
			15.89			11.01
Constant	7.911***	8.892***	4.442***	84.103***	78.297***	44.113***
	12.17	11.83	8.19	10.61	8.53	5.59
Year FE	Y	Y	Y	Y	Y	Y
Morningstar Category FE	Y	Y	Y	Y	Y	Y
R-sqr	0.093	0.183	0.240	0.109	0.180	0.195
Obs.	8,410	8,410	7,405	7,551	7,551	6,577

Table 5: Mutual Fund Difference-in-Difference Around the Introduction of an ETF Twin

This table presents the results of the annual fixed effects panel regression

$$Tax_{f,t} = \alpha + \lambda(MF\ Twin_f * Post_{f,t}) + \beta_1 Post_{f,t} + \beta_2 X_{f,t} + \vartheta_f + \tau_t + \varepsilon_{i,s,t}$$

where $Tax_{f,t}$ is one of three tax measures for fund f in year t. The measures include total *Tax Burden*, long-term capital gains yield, *LCG,* and short-term capital gains yield, *SCG*. *MF Twin* is equal to one for a mutual fund that existed prior to the introduction of a twin ETF. $Post_{f,t}$ is set to one for the two years following the ETF addition for the mutual fund twin and its controls and to zero for the two years prior. The year of the ETF launch is excluded. The controls are matched on Morningstar Category and fund size using the new nearest neighbors in Panel A and the three nearest neighbors in Panel B. $X_{f,t}$ includes covariates that change at the fund and year level and include $Fund\ Size_{f,t}$ the log of total portfolio assets, and $Family\ Size_{f,t}$ the log of total fund family assets. ϑ_f is a fund fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Panel A: Two Nearest Neighbors Match						**Panel B: Three Nearest Neighbors Match**					
Dependent Variable:	**Tax Burden**	**Tax Burden**	**LCG**	**LCG**	**SCG**	**SCG**	**Tax Burden**	**Tax Burden**	**LCG**	**LCG**	**SCG**	**SCG**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
MF Twin * Post	-0.498***	-0.467**	-2.855***	-2.823***	0.046	0.141	-0.320*	-0.299*	-1.966***	-1.948***	0.055	0.145
	-2.80	-2.58	-3.81	-3.68	0.16	0.48	-1.84	-1.68	-2.68	-2.61	0.23	0.60
Post	0.274*	0.286*	1.737***	1.823***	-0.183	-0.177	0.080	0.078	0.852	0.833	-0.224	-0.209
	1.84	1.89	2.75	2.85	-0.74	-0.72	0.62	0.60	1.57	1.53	-1.26	-1.18
Fund Size		-0.081		-0.710		0.014		-0.192		-1.090*		-0.106
		-0.44		-0.91		0.05		-1.30		-1.76		-0.52
Family Size		-0.141		-0.392		-0.393		-0.054		-0.031		-0.327*
		-0.95		-0.63		-1.63		-0.38		-0.05		-1.70
Age Years		-0.095		-0.026		-0.267		-0.015		0.110		-0.083
		-0.69		-0.04		-1.18		-0.19		0.35		-0.80
Constant	0.722**	3.053*	1.670	10.662	0.076	4.753*	0.778**	2.680*	1.888	9.516	0.178	4.089**
	2.02	1.79	1.11	1.48	0.13	1.71	2.32	1.82	1.34	1.54	0.39	2.04
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-sqr	0.558	0.567	0.615	0.622	0.422	0.446	0.555	0.562	0.575	0.586	0.456	0.473
Obs.	156	156	156	156	156	156	208	208	208	208	208	208

Table 6: Mutual Fund Expense Difference-in-Difference Around the Introduction of an ETF Twin

This table presents the results of the annual fixed effects panel regression

$$Expense_{f,t} = \alpha + \lambda(MF\ Twin_f * Post_{f,t}) + \beta_1 Post_{f,t} + \beta_2 X_{f,t} + \vartheta_f + \tau_t + \varepsilon_{i,s,t}$$

where $Expense_{f,t}$ is one of three tax measures for fund f in year t. The measures include *Exp. Ratio*, the year-end expense ratio, *12b-1 Fees*, the fund's 12b1 fee, and *MgmtExp*, a fund's year-end management expense. *MF Twin* is equal to one for a mutual fund that existed prior to the introduction of a twin ETF. $Post_{f,t}$ is set to one for the two years following the ETF addition for the mutual fund twin and its controls and to zero for the two years prior. The year of the ETF launch is excluded. The controls are matched on Morningstar Category and fund size using the new nearest neighbors in Panel A and the three nearest neighbors in Panel B. $X_{f,t}$ includes covariates that change at the fund and year level and include $Fund\ Size_{f,t}$ the log of total portfolio assets, and $Family\ Size_{f,t}$ the log of total fund family assets. ϑ_f is a fund fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Panel A: Two Nearest Neighbors Match						**Panel B: Three Nearest Neighbors Match**					
Dependent Variable:	**Exp Ratio**	**Exp Ratio**	**12b-1 Fees**	**12b-1 Fees**	**Mgmt Fees**	**Mgmt Fees**	**Exp Ratio**	**Exp Ratio**	**12b-1 Fees**	**12b-1 Fees**	**Mgmt Fees**	**Mgmt Fees**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
MF Twin * Post	-0.019	-0.010	-0.002	-0.003	0.010	0.010	-0.009	0.004	0.004	0.000	-0.001	0.004
	-0.62	-0.33	-0.28	-0.45	0.48	0.49	-0.33	0.14	0.57	0.04	-0.03	0.14
Post	0.017	0.018	0.007	0.007	-0.011	-0.008	0.030	0.024	-0.000	0.001	0.021	0.018
	0.65	0.75	1.21	1.16	-0.65	-0.47	1.43	1.17	-0.07	0.16	1.07	0.92
Fund Size		-0.069**		0.004		-0.043*		-0.077***		0.010		-0.036
		-1.99		0.50		-1.79		-3.26		1.59		-1.62
Family Size		-0.039		0.001		-0.006		-0.026		0.005		-0.010
		-1.18		0.18		-0.27		-0.85		0.65		-0.36
Age Years		0.068***		0.004		0.034**		0.024**		0.002		0.015
		3.38		0.73		2.38		2.25		0.53		1.46
Constant	0.356***	0.833**	0.051***	-0.017	0.129***	0.322	0.362***	1.053***	0.053***	-0.080	0.122***	0.413
	6.40	2.02	4.11	-0.17	3.46	1.11	7.32	3.08	4.13	-0.86	2.74	1.28
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-sqr	0.950	0.958	0.977	0.977	0.908	0.917	0.942	0.949	0.985	0.985	0.888	0.892
Obs.	132	132	132	132	132	132	176	176	176	176	176	176

Table 7: Robustness - Mutual Fund Difference-in-Difference with Treatment Two Years After the Introduction of an ETF Twin

This table presents the results of the annual fixed effects panel regression

$$Tax_{f,t} = \alpha + \lambda(MF\ Twin_f * Post_{f,t}) + \beta_1 Post_{f,t} + \beta_2 X_{f,t} + \vartheta_f + \tau_t + \varepsilon_{i,s,t}$$

where $Tax_{f,t}$ is one of three tax measures for fund f in year t. The measures include total *Tax Burden*, long-term capital gains yield, *LCG,* and short-term capital gains yield, *SCG*. *MF Twin* is equal to one for a mutual fund that existed prior to the introduction of a twin ETF. $Post_{f,t}$ is set to one for the two years following treatment. Treatment occurs two years prior to the ETF addition for the mutual fund twin and its controls and to zero for the two years prior. The year of the hypothetical treatment is excluded. The controls are matched on Morningstar Category and fund size using the new nearest neighbors in Panel A and the three nearest neighbors in Panel B. $X_{f,t}$ includes covariates that change at the fund and year level and include $Fund\ Size_{f,t}$ the log of total portfolio assets, and $Family\ Size_{f,t}$ the log of total fund family assets. ϑ_f is a fund fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Panel A: Two Nearest Neighbors Match						**Panel B: Three Nearest Neighbors Match**					
Dependent Variable:	**Tax Burden**	**Tax Burden**	**LCG**	**LCG**	**SCG**	**SCG**	**Tax Burden**	**Tax Burden**	**LCG**	**LCG**	**SCG**	**SCG**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
MF Twin * Post	0.395	0.353	1.565	1.615	0.026	-0.137	0.164	0.116	0.729	0.779	-0.026	-0.144
	1.58	1.43	1.34	1.37	0.07	-0.33	0.77	0.53	0.86	0.89	-0.08	-0.45
Post	-0.693***	-0.497**	-2.440**	-2.698**	-0.478	-0.340	-0.364**	-0.269*	-1.293**	-1.231**	-0.312	-0.224
	-3.49	-2.10	-2.64	-2.38	-1.49	-0.87	-2.58	-1.77	-2.31	-2.02	-1.53	-1.00
Fund Size		-0.541		0.637		-0.426		-0.287		0.020		-0.354
		-1.48		0.36		-0.71		-1.51		0.03		-1.27
Family Size		0.534		-3.180		1.692		0.145		-1.494		0.661
		0.84		-1.04		1.60		0.35		-0.89		1.08
Age (Years)		0.393**		1.756*		0.263		0.355*		1.374*		0.332
		2.11		1.97		0.85		1.97		1.90		1.26
Constant	1.201***	-2.337	2.207	16.251	0.958	-13.389*	0.896***	0.045	1.823	8.819	0.658	-4.069
	3.19	-0.49	1.26	0.71	1.58	-1.69	2.88	0.02	1.48	0.74	1.46	-0.94
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-sqr	0.715	0.751	0.593	0.624	0.604	0.636	0.592	0.619	0.556	0.576	0.506	0.525
Obs.	84	84	84	84	84	84	136	136	136	136	136	136

Table 8: Robustness - Tax Distribution Levels

This table presents the results of the annual fixed effects panel regression

$$Tax\ Level_{f,s,t} = \alpha + \beta_1 MF_f + \beta_2 MF\ Twin_{f,t} + \beta_3 ETF\ Twin_{f,t} + \beta_4 X_{f,s,t} + \varepsilon_{f,s,t}$$

where *Tax Level* $_{f,s,t}$ is either the total annual level of long-term or short-term capital gains by fund f in Morningstar Category s in year t. MF is equal to one if a fund is a mutual fund. Twins are mutual funds and ETFs in the same fund family that follow the same index. *MF Twin* is equal to one for the years in which a mutual fund has an ETF twin and zero otherwise. *ETF Twin* is equal to one when an ETF has a mutual fund twin and zero otherwise. $X_{f,t}$ includes covariates that change at the style and year level and include *Fund Size* $_{f,t}$, the log of total portfolio assets, and *Family Size* $_{f,t}$, the log of total fund family assets. η_s is a Morningstar Category fixed effect and τ_t is a year fixed effect. Standard errors are clustered at the fund level and t-statistics are reported below the coefficients. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable:	**LCG Level**	**LCG Level**	**SCG Level**	**SCG Level**
MF	0.259***	0.163***	0.057***	0.059***
	11.73	5.88	3.99	3.69
MF Twin	-0.207***	-0.160***	-0.002	0.005
	-6.92	-3.80	-0.08	0.14
ETF Twin	-0.013	0.019	-0.031**	-0.021
	-1.04	1.18	-2.56	-1.56
Fund Size		-0.015**		-0.008***
		-2.33		-3.16
Family Size		-0.005		-0.000
		-1.25		-0.06
Age (Years)		0.019***		-0.000
		3.98		-0.31
Constant	0.289**	0.437***	0.064***	0.104***
	2.46	3.51	3.23	3.69
Year FE	Y	Y	Y	Y
Morningstar Category FE	Y	Y	Y	Y
R-sqr	0.116	0.136	0.054	0.058
Obs.	8,503	8,503	8,503	8,503

Hidden Style: The Latent Factors of Hedge Fund Performance

Harold D. Spilker III*

April 1, 2017

Abstract

Latent factors have been used successfully to characterize hedge fund returns beyond the observable asset based factors. Using principal component analysis, Fung and Hsieh (1997a) extract five latent components of hedge fund returns that explain up to 43% of the common return variation of hedge funds. Employing the alternative estimator of Connor and Korajczyk (1986, 1987), I show that asymptotic principal components (APC) can explain 42% more of the common variation of hedge fund returns on average and over a larger sample period. Further, the hedge fund factor model of Fung and Hsieh (2004, 2006) achieve larger $\bar{R}^2$s when employing APC extracted factors as regressors than those obtained from the traditional approach. I also identify an increase in the common variation across hedge fund excess return in the time-series via the extracted latent factors. This increase corresponds with a rise in flows to hedge fund strategies and an attendant crowding effect noted in the literature. My results suggest an impetus for future researchers to employ APC factors when characterizing hedge fund performance.

*Spilker is from Boston College, Carroll School of Management, Chestnut Hill, Massachusetts, 02467; e-mail: tray.spilker@bc.edu

Introduction

The attractiveness of using asset class factor models is the ability to attribute performance of an investment vehicle (e.g., a mutual or hedge fund) to major asset classes.[1] From this, one can discern an investment "style" of a particular fund manager, as shown in Sharpe (1992), by regressing fund excess return on many styles or asset classes (styles). In fact, extracting from the arbitrage pricing theory model (APT) of Ross (1976), even style can be employed as factors.[2] Harlow and Brown (2006) show that some mutual funds exhibit performance persistence when adjusted for these style exposures, particularly when augmented by asset class factors in the form of sector exposures.[3]

The literature has followed the rise of hedge funds in a similar fashion. In their pioneering paper, Fung and Hsieh (1997a) show that mutual fund styles explain very little of hedge fund performance on average, which they attribute to the dynamic investment behavior of hedge funds. Accordingly, the authors augment the Sharpe (1992) style model by identifying dynamic factors. Using principal components analysis (PCA), they extract five mutually orthogonal principal components explaining 43% of the cross sectional return variance. Only three of these components qualify as dynamic strategies, which the authors identify in their 2001 paper and link to observable prices.[4] Combining factors from equity and fixed income mutual funds, as well as their dynamic strategy factors, Fung and Hsieh (2004) construct a seven-factor asset based model that explains up to 80% of return variation across funds of hedge funds (FoFs).[5]

Using hedge fund factor exposures as identified in Fung and Hsieh (2004), Fung, Hsieh, Naik, and Ramadorai (2008) analyzed FoFs over the period of 1995-2000 and found a subset of funds that delivered alpha over longer periods. Jagannathan, Malakhov, and Novikov (2010) extend the model with a correction for autocorrelation[6] and find that only superior

[1] as noted by Jones (2001)

[2] see Fama and French (1993) and Carhart (1997)

[3] see Pstor and Stambaugh (2002)

[4] Fung and Hsieh (2001) develop factors for three common trend-following strategies

[5] see Jones and Wermers (2011) for a summary of the mutual fund and hedge fund literatures

[6] according to Getmansky, Lo, and Makarov (2004)

funds generated persistent returns over the period of 1996-2005. Correcting for non-normality of hedge fund returns, Kosowski, Naik, and Teo (2007) find that top-decile hedge funds outperform bottom-decile funds by 5.8% in the ensuing year.[7]

The technique of principal components analysis, which Fung and Hsieh (1997a) employ to identify dynamic strategies, has limitations. In particular, employing principal components analysis to extract latent factors relies on a balanced panel such that over longer horizons, only those hedge funds with commensurate operating lives are analyzed, or conversely, a restriction on the time horizon is necessary. Accordingly, I employ the asymptotic principal components (APC) model of Connor and Korajczyk (1986), which when augmented by their missing observation technique,[8] allows for the inclusion of all return observations in the analysis (full N, full T), thus increasing the explanatory power of each factor (Connor and Korajczyk (1986, 1987) henceforth CK). While Fung and Hsieh (1997a) find the top five principal components explain 43% of the variation of funds, the replication of these results eludes researchers given the data source nor the time period employed in the paper are disclosed. Using traditional PCA on my available dataset, I replicate the technique of Fung and Hsieh (1997a; henceforth FH) and show that the top five principal components explain up to 40% of the common variation between hedge funds using average $\bar{R}^2$ for individual funds, or 87.5% with an equal-weighted portfolio. Comparatively, I show that employing APC significantly improves adjusted r-sqaures ($\bar{R}^2$) over the replicated results of Fung and Hsieh, achieving a fit of 50% and 99%, respectively, when components are extracted according to APC theory. Further, applying components extracted using FH to out of sample returns results in a decrease in average $\bar{R}^2$ of as little as 18% for individual funds, or 55% for equal-weighted portfolios.

The remainder of the paper is organized as follows. The first section describes the data employed in the sample, followed in section 2 by a discussion of the theoretical methodology. Section 3 provides results of the empirical findings of the paper, and I conclude the paper in section 4.

[7] their method is a hedge fund extension of Kosowski, Timmermann, Wermers, and White (2006)

[8] see Connor and Korajczyk (1987)

1 Data

The data comes from Lipper-TASS for the period 1994:01-2012:12 on all hedge fund strategies as identified in the dataset. The strategies include Convertible Arbitrage, Dedicated Short Bias, Emerging Markets, Equity Market Neutral, Event Driven, Fixed Income Arbitrage, Fund of Funds, Global Macro, Long/Short Equity Hedge, Managed Futures, Multi-Strategy, and Options Strategy.

While TASS data extends before 1994, the capturing of data from delisted hedge funds ("dead" funds) began in 1994. Thus, the inclusion of dead funds reduces much of the survivorship bias that may be present in the "live" funds database.[9] Recently, Edelman, Fung and Hsieh (2013) find that the various motivations for delisting from commercial databases have offset one another over the recent decade; however, the inclusion of dead funds data is commensurate with the bulk of the literature. The dataset consists of just over 18,400 unique hedge fund vehicles that have at least 24 monthly observations who report net-of-fee returns (see Table 1). The largest category of hedge fund in the sample are FoFs, with about 6,200 funds, followed by long/short equity hedge (3,885) and multi-strategy funds (2,321).

Incubation bias in hedge fund returns arises when firms "incubate" funds for several years and decided to report returns to commercial databases such as TASS if the early return stream is "good enough." Fung and Hsieh (2002a) show that incubation bias in the TASS dataset is resolved at the 12 month mark.[10] Thus, I remove the first 12 months of each fund's return history. Lastly, I remove replicated funds from the data set that are pursuing the same strategies within a particular management company, and thus, have almost identical returns. These "replicated" funds are developed by hedge fund management companies to cater to specific needs of the investor interested in the same strategy (e.g., offshore vehicles for tax exempt entities, or non-US dollar based funds) in effort to garnering flows, and would

[9]see Brown, Goetzmann, and Park (1997); Ackermann, McEnally, and Ravenscraft (1999); and Malkiel and Saha (2005)

[10]Park (1995) showed incubation period of 27 months for MAR CTA database (now Morningstar CISDM); Brown, Goetzmann, and Park (1997) similarly showed incubation periods of 27 months for CTAs and 15 months for hedge funds in the TASS data set

lead to erroneous conclusions about cross-sectional correlations if left in the data set. After removing funds that fit these criteria, we are left with just over 2,500 unique hedge funds (bottom of table 1).

I also divide the sample between fund of funds (FoFs) and all remaining strategies in the sample except for FoFs (henceforth, NoFoFs); a delineation shown in the literature. Fung and Hsieh (1997a) perform PCA analysis on all funds except FoFs (NoFoFs here) and emerging market funds. In a later paper, Fung and Hsieh (2000a, 2002a) highlight the affects of the various biases found in the commercially available hedge fund databases and show that employing FoFs closely mimics returns to actual hedge fund portfolios, which are free from said biases. Thus, I include both for the purpose of this analysis.

2 Methodology

As noted in Jones (2001), using principal components analysis (PCA, as in FH) is particularly attractive when analyzing portfolio returns as it requires no specification of the factors ex ante. Thus, it is an ideal method to test asset pricing models. Following the literature, we assume hedge fund returns follow a factor structure as in an APT model,

$$R^n = B^n F + e^n, \tag{1}$$

where, R^n is an $n \times 1$ vector of excess returns at time t, B^n is an $n \times k$ matrix of factor loadings (sensitivities), F is a $k \times T$ vector of systemic factors, and e^n is an $n \times T$ matrix of asset specific residuals. From this factor structure, PCA extracts the k-largest eigenvalues and corresponding eigenvectors from the covariance matrix of returns, Σ. Here, the covariance matrix will be $n \times n$, requiring extraordinary computing resources for the decomposition as n gets large. Further, the PCA method relies on a balanced panel for the extraction, reducing the number of observations in panels with missing data.

Based on the work of Chamberlin and Rothschild (1983), Connor and Korajczyk (1986) define Asymptotic Principal Components as the eigenvectors obtained from the k-largest eigenvalues of the cross-product matrix of returns

$$\Omega^n = \left(\frac{1}{n}\right) R^{n\prime} R^n. \tag{2}$$

This $T \times T$ matrix reduces the computational burden of factor reduction, while achieving consistent estimates of the latent factors. Thus, we are able to estimate latent factors from large cross-sections whereas under traditional PCA this would be impossible. The convenience of the CK approach is that it does not rely upon a normal distribution of returns nor a diagonal covariance matrix, and allows for time-varying factor risk premia. Note that the full decomposition of the cross-product matrix, Ω^n, which includes unobservable and error terms, takes the form:

$$\begin{aligned} \Omega^n &= \left(\frac{1}{n}\right) R^{n\prime} R^n = A^n + B^n + B^{n\prime} + Z^n \\ &\equiv \left(\frac{1}{n}\right) F' B^{n\prime} B^n F + \left(\frac{1}{n}\right) F' B^{n\prime} e^n + e^{n\prime} B^n F + \left(\frac{1}{n}\right) e^{n\prime} e^n. \end{aligned} \tag{3}$$

However, under the assumption of independence between factors and residuals (CK assumption 7), the B^n and Z^n terms are equal to zero, leaving the observable $T \times T$ matrix, Ω^n.

The CK method by itself, however, does not avoid the loss of firm observations resulting from missing data in the panel. To accommodate for missing data, CK propose an alternative estimator for Ω^n. In their specification, Connor and Korajczyk (1987) define R^m be the $m \times T$ matrix of excess returns where missing data are replaced with zeroes. Let I^m be an indicator matrix for which I^m is equal to one if R^m is observed, and zero otherwise. If returns in R^m follow the process in equation (1), then the estimates of the latent factors extracted from the alternative estimator,

$$\Omega^m = (R^{m\prime} R^m)/(I^{m\prime} I^m), \tag{4}$$

are identical to those from equation (2); i.e., $\Omega^n = \Omega^m$. Hence, we are able to extract as much information contained in the returns data, thereby enhancing efficiency of our estimator.

3 Empirical Tests and Findings

3.1 Latent Factor Regressions

Applying the PCA technique over the whole sample period necessarily drops funds with missing return data in order to form a balanced panel. Therefore, I create sub-periods according to Fung and Hsieh (2004) at least through the end of 2002, when the paper's sample period end. Beyond 2002, I exploit shocks in the liquidity factor of Sadka (2006)[11] to identify additional regimes for analysis.[12] By reducing the horizon for the analysis, I hope to extract the maximal common variance across hedge fund returns using the FH method by reducing the number of excluded funds resulting from the balanced panel requirement.

Using the eigenvectors extracted from the covariance matrix of excess returns as our vector of factors (via PCA), for both in- and out-of-sample funds, I regress excess returns of individual funds and an equal-weighted portfolio on the factors as outlined in equation (1). Specifically, I regress excess returns comprised of the funds from which the factors were extracted onto the latent factors (in sample), $R^n_{in} = B^n F + e^n$. Then the same regression is run using funds that have some returns in the dataset for the given time period and strategy, but were dropped from the PCA technique due to the balanced panel restriction (out-of-sample), $R^n_{out} = B^n F + e^n$.

Results from these regressions can be found in the left hand pane of Tables 2 and 3 for FoFs and all other funds (NoFoFs), respectively. Notice that on average, the ability for the extracted factors to explain out-of-sample excess returns is less than for in-sample returns (Panels A and B, respectively). For FoFs, the first five principal components explain 42% (43.2%) of the common variation among funds, whilst the out-of-sample adjusted-R^2 falls to 26.9% (29.6%) for individual (equal-weighted) regressions over the entire sample period (1994:01-2012:12). For NoFoFs, $\bar{R}^2$s predictably decline to 36.6% (63.1%) and 20.3% (58.8%) for in- and out-out-sample regressions, respectively.

By construction, FoFs exhibit lower volatility than individual funds, thus we should expect less common variation from the NoFoF funds. The tables also include the number of funds included in the PCA factor extraction and subsequent regressions by time period. On average we are leaving a large percentage of funds out of the decomposition using traditional PCA, thereby reducing efficiency of our estimators. For example, over the entire sample period of 1994-2012, PCA extracts components from only nine funds from a sample of 518 FoFs, and 36 of the remaining 2,061 hedge funds. Thus, when regressing excess returns from

[11] Thanks to Ronnie Sadka for making the liquidity factors from Sadka (2006) available at his website: https://www2.bc.edu/~sadka/

[12] These regimes are supported by Edelman, Fung, Hsieh, and Naik (2012) through the end of 2010

the remaining 509 out-of-sample FoFs onto the components extracted from these nine funds, $\bar{R}^2$s fall to 26.9%; the components have less common variance from which to explain excess returns.

To further this point, the in- and out-of-sample $\bar{R}^2$s for FoFs during the 2007:07-2009:04 period of 50.2% and 49.7% for the first five components, respectively, shows evidence of minimal decay in explanatory power for out-of-sample regressions when the in-sample size is large relative to the out-of-sample size (N=241 versus N=116).

Following the same approach, I run the regressions of equation (1) of excess return on the factors extracted with the CK alternative asymptotic estimator. This ensures we extract as much common variance across the excess returns of hedge funds in the sample as possible, the results of which are displayed in the right hand panes of Tables 2 and 3. Highlighting the efficiency of the APC technique, these regressions bear $\bar{R}^2$s from the first five principal components of 61.3% (94.2%) for individual (equal-weighted) FoFs over the entire sample period, and 35.2% (90.2%) for NoFoFs. For FoFs, a gain of almost 30% is achieved (61.3% from 42% in-sample) for individual fund regressions and a doubling is achieved for equal-weighted portfolios (94.2% from 43.2% in-sample) when employing the CK methodology. NoFoF APCs (principal components from all other funds other than FoFs) modestly lose power at the individual fund level (35.2% from 36.6% in-sample), but gain ground on equal-weighted regressions (90.2% from 63.1% in-sample) from the first five latent components.

Given the CK approach is inclusive of all funds, a more appropriate comparison would have CK results next to the PCA out-of-sample results. Here, we see even more dramatic improvements in explanatory power between the two techniques, with a roughly 15-point gain between NoFoF $\bar{R}^2$s using the first five components as evidence over the entire sample period.

Relative to PCA, the efficiency of the CK technique is also on display here. Over the full sample period, the CK method gains little from the inclusion of principal components beyond 5 (7.5% for FoFs), while the FH method has a larger percentage gain from the inclusion of more components into the regression (33% for FoFs). Thus, the CK method achieves a greater explanation of the common variation among hedge funds with fewer latent factors.

The results also provide some time series implications. By analyzing the various horizon regimes, we see a notable increase in $\bar{R}^2$s from the first period (1994:01-1998:09) to the most recent periods on average across fund types (FoF or NoFoFs) and across factor extraction methods (FH or CK). For example, the $\bar{R}^2$s increase from 18% to 41.8% for the first five principal components of FH out-of-sample regressions among FoFs between the first and last period, and similarly, the CK method for the sample period and funds improves from 57% to 67.5%. These results coincide with a rise in the number of hedge funds and asset flows into these strategies and suggests that the proliferation of hedge fund vehicles has resulted in a crowding effect with attendant increases in fund return correlations.[13]

3.2 Identifying Latent Factors

Clearly, the method of CK is superior to that of PCA in explaining common variation across hedge fund returns. However, the abstract nature of these components begs for identification. Fung and Hsieh satisfy this need for linkages to observable prices by identifying market factors and trading strategies that closely price, and thus convert, returns based factors into asset based factors. With the inclusion of an emerging market factor in Fung and Hsieh (2006), the Fung and Hsieh model culminates in eight factors, which is comprised of two equity factors similar to the first two factors of Fama and French (1993), the S&P 500 minus the risk free rate (SP-Rf) and the S&P 500 minus the Russell 2000 index (SP-RL); two fixed income factors consisting of the excess return to ten-year treasuries (TY-Rf) and the return of Moody's BAA corporate bonds over ten-year treasuries (BAA-TY); three trend following factors consisting of excess returns on bonds (PTFSBD-Rf), foreign currencies (PTFSFX-Rf), and commodities (PTFSCOM-Rf);[14] and lastly, a factor for emerging markets, taken from the International Finance Corporation index (IFC-Rf).[15]

Before identifying latent factors, I begin with those that can be witnessed; the reported returns. Table 4 shows $\bar{R}^2$s from regressions of fund excess returns, R_i, on the observable

[13]Fung and Hsieh (2007) and Asness, Krail, and Liew (2001) find increasing correlations between hedge funds and standard market indices.

[14]as added in Fung and Hsieh (2001)

[15]Thanks to David Hsieh for providing these factors on his website: http://people.duke.edu/~dah7/

hedge fund factors of Fung and Hsieh, as follows:

$$\begin{aligned} R_i = \alpha_i &+ \beta_{1,i}(SP - Rf) + \beta_{2,i}(SP - RL) + \beta_{3,i}(TY - Rf) \\ &+ \beta_{4,i}(BAA - TY) + \beta_{5,i}(PTFSBD - Rf) + \beta_{6,i}(PTFSFX - Rf) \\ &+ \beta_{7,i}(PTFSCOM - Rf) + \beta_{8,i}(IFC - Rf) + e_i. \end{aligned} \tag{5}$$

The top line of each panel shows the $\bar{R}^2$ from a multivariate regression across all 8-factors, and below, univariate R^2s for each factor individually. Relative to Fung and Hsieh, I achieve an equal-weighted $\bar{R}^2$ of 68.8% during the same horizon in their 2004 paper of 2000:04-2002:12 when projecting FoF excess returns onto the 8-factor model (left side of Panel B),[16] reaching a maximum of 83.1% during the most recent period (2009:05-2012:12), and an overall fit of 64% for the entire period.[17] For individual funds, I obtain a maximum $\bar{R}^2$ during the most recent period of 58.9% and an overall fit of 46.9% for the whole period. On the right hand side of the table, I show results for regressing NoFoF returns onto the factors. Notably, while the individual fund results fall almost predictably across the board relative to FoFs, the equal-weighted NoFoF (EW-NoFoF) results achieve higher $\bar{R}^2$s on average. This may suggest that FoFs provide some diversification benefit beyond a naive 1/N investment policy.[18] Dissecting this further, we see that over the full sample period, NoFoFs load more heavily on the equity factors (SP-Rf and SP-RL), including the emerging market equity strategies (IFC-Rf), while loadings on the fixed income factors decline (TY-Rf and BAA-TY). This holds for both equal-weighted and individual fund regressions, suggesting our FoFs sample have a more balanced exposure towards fixed income risk factors and away from equity factors.

Identification of the principal components (PCs) takes the same approach. Tables 5 and 6 show R^2s and associated t-stats from regressing PCs (latent factors) onto the 8-factors of Fung and Hsieh. Specifically, each PC is projected onto each factor one by one as follows:

[16] Fung and Hsieh (2004) only used the first 7-factors

[17] Fung and Hsieh (2004) obtain an R^2 of 80% over the period 2000:04-2002:12 using the returns of the Hedge Fund Research FoF index (HFRFOF) on 7-factors

[18] see DeMiguel, Garlappi, and Uppal (2007)

$$\begin{aligned}
PC_i &= \alpha_i + \beta_{1,i}(SP - Rf) + e_i \\
PC_i &= \alpha_i + \beta_{2,i}(SP - RL) + e_i \\
&\vdots \\
PC_i &= \alpha_i + \beta_{8,i}(IFC - Rf) + e_i.
\end{aligned} \tag{6}$$

Then, the same PC is regressed across all factors (multivariate) to achieve the reported $\bar{R}^2$ in the last column of each panel according to:

$$\begin{aligned}
PC_i = \alpha_i &+ \beta_{1,i}(SP - Rf) + \beta_{2,i}(SP - RL) + \beta_{3,i}(TY - Rf) \\
&+ \beta_{4,i}(BAA - TY) + \beta_{5,i}(PTFSBD - Rf) + \beta_{6,i}(PTFSFX - Rf) \\
&+ \beta_{7,i}(PTFSCOM - Rf) + \beta_{8,i}(IFC - Rf) + e_i.
\end{aligned} \tag{7}$$

The left side of the table shows results from the components obtained from the FH method, and the right side shows those from the CK method. It is quickly evident that the 8-factors explain a great deal more of the variation in extracted components from the CK method relative to those of the FH method. For example, over the entire sample period, the $\bar{R}^2$ for the first principal component (PC1) from the CK method reaches 76.4% while PC1 from the FH method only obtains 26.8%. That is, the 8-factors of FH explain three times more of the first principal component extracted via the CK method than they do for the first component of FH. This result bolsters the case for the ability of the 8-factor model to explain hedge fund returns. Further, the t-stats indicate the direction of exposure to each observed factor. For PC1, we see the component has positive and significant loadings on SP-Rf, TY-Rf, and IFC-Rf, while loading negatively and significantly on the remaining factors. This suggests that PC1 - from both extraction methods - could be viewed as a portfolio with long exposure to the US equity market, US treasuries, and emerging markets, combined with short exposure to all remaining factors in the model. We see the same outcome on R^2s for the NoFoFs, with an even greater improvement between the two methods.

The orthogonality of the principal components, however, limits our ability to strictly identify the components with observable factors. Particularly, as the components as extracted are only mutually orthogonal to one another and not to the observable factors. The 8-factors of FH themselves are not orthogonal to each other, thus complicating the identification exercise. Yet, we can be sure the 8-factor model says much more about the components extracted via the CK method on average, than the traditional PCA of FH.

4 Conclusion

Fung and Hsieh (1997a) set forth a novel approach to link return based factors to asset based factors for hedge funds. Using traditional principal component extraction to explain hedge fund returns has proven beneficial in understanding the return dynamics of hedge fund strategies. Employing an alternative estimator proposed by Connor and Korajczyk (1986, 1987) from which to extract latent factors, I have shown leads to efficiency gains and greater power when explaining the common variation across hedge fund excess returns. Further, the latent factors extracted from the asymptotic estimator exhibit a better fit with the observable hedge fund factors of Fung and Hsieh. Lastly, the latent factors have corroborated the evidence of increased correlations among observed excess returns in the time series, suggesting a crowding effect among respective hedge fund strategies and the resultant decline in returns witnessed in the literature. There remains, however, some explanatory power of the principal components over the observable factor model, thereby forcing us to consider their use to augment models restricted solely to observable factors. Accordingly, future research of comparative hedge fund return characteristics should consider latent factor extraction via the asymptotic estimator method of CK .

Table 1: Summary Statistics

	N	Month Avg Excess Return	Month Avg Stdev	Annual Avg Excess Return	Annual Avg Stdev	Annual Avg Sharpe
Full Dataset*						
All Funds	18412	0.27	3.40	3.27	11.77	0.28
Convertible Arbitrage	308	0.28	2.42	3.45	8.40	0.41
Dedicated Short Bias	56	-0.07	6.07	-0.88	21.02	-0.04
Emerging Markets	1033	0.33	5.47	3.98	18.96	0.21
Equity Market Neutral	693	0.34	3.72	4.10	12.90	0.32
Event Driven	802	0.52	3.06	6.43	10.60	0.61
Fixed Income Arbitrage	473	0.36	2.28	4.44	7.91	0.56
Fund of Funds	6261	-0.03	2.48	-0.30	8.58	-0.04
Global Macro	921	0.40	3.52	4.89	12.20	0.40
Long/Short Equity Hedge	3885	0.45	4.62	5.48	15.99	0.34
Managed Futures	1112	0.21	5.04	2.53	17.45	0.15
Multi-Strategy	2321	0.52	2.50	6.44	8.65	0.74
Options Strategy	51	0.83	4.09	10.49	14.16	0.74
Other	495	0.56	2.85	6.88	9.88	0.70
Cleaned Sample**						
All Funds	2579	0.51	4.08	6.23	14.14	0.44
Convertible Arbitrage	67	0.28	2.47	3.37	8.56	0.39
Dedicated Short Bias	20	-0.09	6.55	-1.02	22.68	-0.04
Emerging Markets	83	0.64	5.50	7.93	19.05	0.42
Equity Market Neutral	136	0.35	2.78	4.34	9.63	0.45
Event Driven	242	0.68	3.34	8.51	11.56	0.74
Fixed Income Arbitrage	69	0.50	2.28	6.10	7.90	0.77
Fund of Funds	518	0.26	2.57	3.20	8.89	0.36
Global Macro	81	0.52	4.63	6.45	16.03	0.40
Long/Short Equity Hedge	900	0.61	5.10	7.53	17.67	0.43
Managed Futures	232	0.43	5.94	5.30	20.58	0.26
Multi-Strategy	133	0.56	3.05	6.90	10.55	0.65
Options Strategy	7	0.53	3.51	6.61	12.17	0.54
Other	91	0.91	3.60	11.44	12.46	0.92

* Base dataset has been screened for funds with at least 24 months of reported return on a net -of-fee basis

** The cleaned dataset restricts replicative share classes that are typically domiciled off-shore from their respective management company and non-US dollar based

Table 2: Funds of Funds

Average adjusted R^2s from regressing excess monthly returns on extracted principal components. Funds have been corrected for incubation bias (Fung and Hsieh 2002a; 12 months), and consist of onshore, US Dollar denominated hedge fund vehicles, with at least 24 months of observations.

	Principal Components Analysis (PCA) - Fung and Hsieh (1997a)							Asymptotic Principal Components (APC) - Connor and Korajzyck (1986, 1987)						
	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12
Panel A: Individual Fund Regressions														
In Sample Means														
N=	9	51	125	142	159	241	158	518	144	164	225	410	357	288
1	0.2623	0.1695	0.1244	0.1861	0.1317	0.1500	0.3898	0.4841	0.1805	0.2920	0.1865	0.4902	0.5883	0.5925
2	0.2967	0.2280	0.1759	0.2867	0.4088	0.2824	0.5319	0.5462	0.5084	0.3827	0.3934	0.5599	0.6748	0.6033
3	0.3580	0.2973	0.2698	0.2940	0.4242	0.4777	0.5384	0.5736	0.5269	0.4490	0.4106	0.5667	0.7327	0.6123
4	0.3970	0.3493	0.2918	0.2941	0.4515	0.4691	0.5591	0.5959	0.5362	0.4974	0.4668	0.6135	0.7499	0.6470
5	**0.4199**	**0.3615**	**0.4098**	**0.3094**	**0.5076**	**0.5016**	**0.5667**	**0.6126**	**0.5696**	**0.5213**	**0.4731**	**0.6234**	**0.7788**	**0.6752**
6	0.4620	0.3793	0.4339	0.3710	0.5190	0.6072	0.5898	0.6232	0.5926	0.5811	0.5308	0.6592	0.8057	0.6862
7	0.4773	0.3984	0.4363	0.3946	0.5292	0.6254	0.5961	0.6323	0.6027	0.6417	0.5652	0.6727	0.8190	0.7085
8	0.4915	0.4088	0.4450	0.3962	0.5299	0.6479	0.6121	0.6460	0.6212	0.6643	0.5831	0.6816	0.8331	0.7220
9	0.5251	0.4581	0.4243	0.4077	0.5435	0.6442	0.6258	0.6521	0.6470	0.6828	0.5936	0.6879	0.8452	0.7305
10	**0.5596**	**0.4676**	**0.4757**	**0.4365**	**0.5737**	**0.6449**	**0.6483**	**0.6594**	**0.6636**	**0.7157**	**0.6201**	**0.6965**	**0.8458**	**0.7513**
Out of Sample Means														
N=	509	93	39	83	251	116	130							
1	0.2158	0.1139	0.0855	0.1100	0.1054	0.1013	0.2943							
2	0.2363	0.1453	0.1931	0.1994	0.3389	0.3590	0.4006							
3	0.2561	0.1483	0.2326	0.2649	0.3385	0.4899	0.4073							
4	0.2593	0.2091	0.2700	0.2876	0.4366	0.5092	0.4155							
5	**0.2686**	**0.1796**	**0.3410**	**0.1974**	**0.4871**	**0.4972**	**0.4177**							
6	0.2736	0.1389	0.3479	0.2256	0.4801	0.5780	0.3960							
7	0.2787	0.1989	0.2891	0.2917	0.5004	0.6139	0.4761							
8	0.2860	0.1992	0.2748	0.2501	0.5058	0.6249	0.5137							
9	0.3211	0.2473	0.0449	0.3089	0.5032	0.6124	0.5118							
10	**0.3604**	**0.2487**	**0.2869**	**0.2543**	**0.5288**	**0.6029**	**0.5144**							
Panel B: Equal-Weighted Portfolio Regressions														
In Sample Means														
N=	9	51	125	142	159	241	158	518	144	164	225	410	357	288
1	0.3589	0.3095	0.2240	0.0797	0.3364	0.2418	0.4833	0.7267	0.3397	0.9020	0.0616	0.9753	0.9545	0.9641
2	0.3665	0.3826	0.4028	0.6374	0.7615	0.4410	0.7084	0.8421	0.9535	0.9492	0.9204	0.9867	0.9945	0.9653
3	0.4032	0.5209	0.6117	0.6394	0.7615	0.6708	0.7084	0.9062	0.9551	0.9634	0.9177	0.9878	0.9942	0.9644
4	0.4322	0.5924	0.6181	0.6491	0.7692	0.6714	0.7417	0.9087	0.9554	0.9609	0.9231	0.9882	0.9968	0.9813
5	**0.4323**	**0.5947**	**0.8041**	**0.6896**	**0.8057**	**0.7042**	**0.7492**	**0.9416**	**0.9571**	**0.9792**	**0.9296**	**0.9913**	**0.9991**	**0.9808**
6	0.4569	0.6175	0.8333	0.6935	0.8182	0.8154	0.7807	0.9447	0.9565	0.9917	0.9346	0.9940	0.9992	0.9804
7	0.4574	0.6399	0.8422	0.7033	0.8201	0.8396	0.7865	0.9471	0.9623	0.9921	0.9371	0.9946	0.9991	0.9807
8	0.4614	0.6421	0.8427	0.7034	0.8202	0.8662	0.7865	0.9709	0.9675	0.9912	0.9649	0.9946	0.9991	0.9801
9	0.5067	0.7102	0.8555	0.7136	0.8318	0.8683	0.7990	0.9708	0.9693	0.9915	0.9693	0.9959	0.9991	0.9880
10	**0.5514**	**0.7202**	**0.8748**	**0.7380**	**0.8478**	**0.8684**	**0.8350**	**0.9715**	**0.9721**	**0.9922**	**0.9778**	**0.9959**	**0.9990**	**0.9910**
Out of Sample Means														
N=	509	93	39	83	251	116	130							
1	0.2603	0.1591	0.3062	0.0058	0.3104	0.2337	0.4848							
2	0.2626	0.1948	0.4318	0.3653	0.7244	0.4791	0.7091							
3	0.2838	0.2775	0.5246	0.3728	0.7245	0.7040	0.7091							
4	0.2948	0.3545	0.5241	0.5561	0.7370	0.7042	0.7125							
5	**0.2955**	**0.3573**	**0.5439**	**0.5602**	**0.7690**	**0.7748**	**0.7138**							
6	0.3139	0.3680	0.6029	0.5608	0.7751	0.8236	0.7558							
7	0.3139	0.3697	0.6043	0.5610	0.7758	0.8496	0.7582							
8	0.3156	0.3696	0.6312	0.5626	0.7766	0.8884	0.7622							
9	0.3692	0.4938	0.6650	0.5863	0.7813	0.8903	0.7700							
10	**0.4177**	**0.5160**	**0.7940**	**0.6407**	**0.8012**	**0.8932**	**0.7904**							

Table 3: All Funds except for Funds of Funds

Average adjusted R^2s from regressing excess monthly returns on extracted principal components. Funds have been corrected for incubation bias (Fung and Hsieh 2002a; 12 months), and consist of onshore, US Dollar denominated hedge fund vehicles, with at least 24 months of observations.

	Principal Components Analysis (PCA) - Fung and Hsieh (1997a)							Asymptotic Principal Components (APC) - Connor and Korajzyck (1986, 1987)						
	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12
Panel A: Individual Fund Regressions														
In Sample Means														
N=	36	186	540	557	540	744	458	2061	599	705	998	1486	1182	1029
1	0.1317	0.0571	0.1442	0.2050	0.2175	0.2297	0.3059	0.2514	0.2405	0.1829	0.2196	0.2199	0.3025	0.3109
2	0.1602	0.1347	0.1823	0.2620	0.2671	0.3515	0.3709	0.2785	0.3400	0.2436	0.2823	0.2276	0.3898	0.3133
3	0.2489	0.2588	0.1871	0.2920	0.2969	0.3795	0.3789	0.3211	0.3667	0.2672	0.3349	0.2650	0.4300	0.3686
4	0.2935	0.2796	0.2110	0.3221	0.3022	0.3905	0.3949	0.3360	0.3854	0.2910	0.3607	0.2849	0.4749	0.4032
5	**0.3658**	**0.2991**	**0.2244**	**0.3560**	**0.3120**	**0.4042**	**0.4077**	**0.3516**	**0.4081**	**0.3206**	**0.3902**	**0.3111**	**0.5028**	**0.4393**
6	0.3976	0.3091	0.2691	0.3715	0.3128	0.4370	0.4428	0.3595	0.4490	0.3431	0.4235	0.3268	0.5268	0.4428
7	0.4228	0.3224	0.2830	0.3753	0.3291	0.4436	0.4523	0.3730	0.4627	0.4072	0.4474	0.3297	0.5602	0.4551
8	0.4458	0.3282	0.3135	0.3859	0.3379	0.4507	0.4563	0.3869	0.4736	0.4182	0.4543	0.3317	0.5775	0.4835
9	0.4688	0.3323	0.3243	0.3844	0.3418	0.4543	0.4643	0.3987	0.4915	0.4508	0.4625	0.3422	0.5917	0.4956
10	**0.4863**	**0.3674**	**0.3059**	**0.4041**	**0.3478**	**0.4451**	**0.4748**	**0.4121**	**0.4997**	**0.4662**	**0.4805**	**0.3519**	**0.6081**	**0.5099**
Out of Sample Means														
N=	2025	413	165	441	946	438	571							
1	0.0549	0.0654	0.1790	0.2106	0.1621	0.1425	0.2237							
2	0.0601	0.1409	0.1533	0.2350	0.2104	0.2614	0.2694							
3	0.1280	0.2251	0.1194	0.2687	0.2371	0.3139	0.2647							
4	0.1503	0.2252	0.2108	0.3048	0.2385	0.3362	0.2840							
5	**0.2028**	**0.2023**	**0.1214**	**0.3252**	**0.2577**	**0.3747**	**0.2958**							
6	0.2191	0.1766	0.1230	0.3288	0.2598	0.3828	0.3222							
7	0.2275	0.1880	0.2198	0.3410	0.2660	0.3902	0.3345							
8	0.2359	0.2245	0.2265	0.3535	0.2704	0.3671	0.3320							
9	0.2407	0.2299	0.2501	0.3851	0.2593	0.3625	0.3230							
10	**0.2471**	**0.2332**	**0.1799**	**0.4174**	**0.2756**	**0.4024**	**0.3351**							
Panel B: Equal-Weighted Portfolio Regressions														
In Sample Means														
N=	36	186	540	557	540	744	458	2061	599	705	998	1486	1182	1029
1	0.0452	0.0012	0.5832	0.5341	0.8683	0.4732	0.6926	0.8291	0.2484	0.7700	0.7597	0.9747	0.9035	0.9706
2	0.1244	0.0124	0.6281	0.6449	0.8750	0.8379	0.8313	0.8285	0.8470	0.8306	0.7665	0.9785	0.9618	0.9704
3	0.3686	0.6123	0.6803	0.7885	0.8804	0.8389	0.8326	0.8958	0.8792	0.8741	0.9478	0.9785	0.9626	0.9789
4	0.4184	0.6212	0.6803	0.8383	0.8913	0.8390	0.8365	0.9021	0.9176	0.9046	0.9793	0.9816	0.9939	0.9784
5	**0.6313**	**0.6368**	**0.6998**	**0.8607**	**0.9048**	**0.8446**	**0.8381**	**0.9018**	**0.9268**	**0.9077**	**0.9828**	**0.9819**	**0.9938**	**0.9812**
6	0.6383	0.6684	0.8543	0.8631	0.9059	0.8766	0.8706	0.9014	0.9421	0.8993	0.9865	0.9821	0.9934	0.9826
7	0.6459	0.6931	0.8955	0.8758	0.9098	0.8849	0.8708	0.9029	0.9437	0.9705	0.9870	0.9852	0.9936	0.9825
8	0.6557	0.7087	0.9022	0.8829	0.9106	0.8849	0.8724	0.9225	0.9642	0.9672	0.9864	0.9848	0.9943	0.9907
9	0.6574	0.7106	0.9060	0.8870	0.9107	0.8944	0.8749	0.9275	0.9684	0.9787	0.9866	0.9874	0.9941	0.9927
10	**0.6630**	**0.7493**	**0.9065**	**0.8879**	**0.9114**	**0.8992**	**0.8879**	**0.9461**	**0.9679**	**0.9768**	**0.9865**	**0.9872**	**0.9942**	**0.9926**
Out of Sample Means														
N=	2025	413	165	441	946	438	571							
1	0.0096	0.0426	0.3457	0.6727	0.7840	0.6415	0.7482							
2	0.0325	0.0474	0.3595	0.7365	0.8551	0.8368	0.8574							
3	0.3212	0.5239	0.3794	0.7784	0.8703	0.8493	0.8575							
4	0.3596	0.5348	0.3794	0.8574	0.8740	0.8515	0.8575							
5	**0.5878**	**0.5445**	**0.3803**	**0.8720**	**0.8917**	**0.8865**	**0.8585**							
6	0.6049	0.5536	0.5192	0.8900	0.8917	0.9170	0.8813							
7	0.6179	0.6059	0.6368	0.8996	0.8917	0.9203	0.8848							
8	0.6186	0.6074	0.7096	0.9005	0.8918	0.9225	0.8848							
9	0.6189	0.6158	0.7212	0.9011	0.8918	0.9235	0.8877							
10	**0.6210**	**0.6321**	**0.7221**	**0.9150**	**0.8920**	**0.9263**	**0.8911**							

Table 4: Hedge Fund R^2s

Average R^2s obtained from regressing fund excess returns on asset based risk factors (8-factor model of Fung and Hsieh 2001, 2002, 2004, 2006). All regressions are univariate, except for "8-Factor," which includes all factors in the regression. The R^2s from the 8-factor regression are adjusted R^2s.

	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12	94 - 12	jan94 - sep98	oct98 - mar00	apr00 - dec02	jan03 - jun07	jul07 - apr09	may09 - dec12
	Fund of Funds							All Funds Except Fund of Funds						
	Panel A: Individual Fund Regressions													
8-Factor	0.4685	0.4894	0.3546	0.3420	0.4685	0.5633	0.5894	0.3138	0.4105	0.2352	0.2995	0.2650	0.3818	0.3954
SP-Rf	0.2638	0.2976	0.1354	0.2135	0.3115	0.2766	0.4650	0.2076	0.2905	0.1435	0.2414	0.2078	0.3164	0.3093
SP-RL	0.0890	0.0947	0.2390	0.1629	0.2222	0.0904	0.1207	0.0862	0.1336	0.1932	0.1162	0.1596	0.1591	0.1264
TY-Rf	0.0733	0.0756	0.0759	0.0971	0.0636	0.1545	0.2663	0.0485	0.0757	0.0922	0.1139	0.0490	0.1217	0.1563
BAA-TY	0.1535	0.0827	0.0965	0.1726	0.0509	0.3059	0.2822	0.0791	0.0816	0.1012	0.1370	0.0455	0.1745	0.1579
PTFSBD-Rf	0.0799	0.1930	0.1185	0.0578	0.0493	0.1982	0.1804	0.0565	0.1607	0.1192	0.0735	0.0476	0.1395	0.1501
PTFSFX-Rf	0.0591	0.0916	0.1184	0.0936	0.0861	0.2028	0.0726	0.0440	0.0812	0.1015	0.0954	0.0640	0.1682	0.0789
PTFSCOM-Rf	0.0415	0.1246	0.1134	0.0657	0.0607	0.1361	0.0723	0.0382	0.1418	0.0954	0.0951	0.0528	0.1574	0.0760
IFC-Rf	0.3486	0.2635	0.1788	0.2473	0.3956	0.4578	0.4860	0.2074	0.2401	0.1555	0.2379	0.2067	0.3344	0.3065
	Panel B: Equal-Weighted Regressions													
8-Factor	0.6395	0.6221	0.8683	0.6879	0.7571	0.7450	0.8312	0.7905	0.7398	0.9537	0.8703	0.8586	0.8164	0.9031
SP-Rf	0.3608	0.3416	0.0166	0.2326	0.4635	0.3660	0.7021	0.5465	0.5013	0.0762	0.5253	0.5599	0.6291	0.7918
SP-RL	0.0975	0.0001	0.4290	0.3827	0.3574	0.0387	0.1308	0.1757	0.0735	0.3977	0.3279	0.4815	0.1606	0.1922
TY-Rf	0.0336	0.0467	0.0023	0.0031	0.0387	0.0706	0.3820	0.0333	0.0395	0.0235	0.0437	0.0229	0.0105	0.3387
BAA-TY	0.1481	0.0160	0.0220	0.0907	0.0028	0.3190	0.4007	0.1243	0.0098	0.0020	0.1705	0.0000	0.1629	0.3217
PTFSBD-Rf	0.0213	0.0005	0.0368	0.0245	0.0001	0.2256	0.1676	0.0369	0.0233	0.1066	0.0186	0.0002	0.1744	0.1745
PTFSFX-Rf	0.0002	0.1405	0.0459	0.0210	0.1252	0.2493	0.0174	0.0008	0.0591	0.0573	0.0003	0.1207	0.2317	0.0100
PTFSCOM-Rf	0.0001	0.1477	0.0836	0.0010	0.0330	0.0881	0.0424	0.0010	0.0350	0.0670	0.0028	0.0487	0.0710	0.0199
IFC-Rf	0.4950	0.2749	0.2723	0.4306	0.6781	0.6785	0.7335	0.6072	0.4020	0.2811	0.6365	0.6994	0.8369	0.8167

Table 5: Fund of Funds (FoFs)

R^2s from regressing each extracted principal components onto each hedge fund risk factor individually along with associated T-stats beneath. The adjusted R^2 column represents multivariate regressions of each principal component regressed onto all factors of the 8-factor model of Fung and Hsieh (2004, 2006).

	SP-Rf	SP-RL	TY-Rf	BAA-TY	PTFSBD-Rf	PTFSFX-Rf	PTFSCOM-Rf	IFC-Rf	Adjusted R^2	SP-Rf	SP-RL	TY-Rf	BAA-TY	PTFSBD-Rf	PTFSFX-Rf	PTFSCOM-Rf	IFC-Rf	Adjusted R^2
	Principal Components Analysis (PCA) - Fung and Hsieh (1997a)									Asymptotic Principal Components (APC) - Connor and Korajzyck (1986, 1987)								
94 - 12																		
PC1	0.1273	0.0443	0.0497	0.0882	0.0959	0.0983	0.1146	0.0937	0.2683	0.4785	0.1192	0.0660	0.2304	0.0710	0.0333	0.0239	0.6122	0.7640
	5.7418	**-3.2372**	**3.4363**	**-4.6770**	**-4.8956**	**-4.9640**	**-5.4091**	**4.8342**		**14.4005**	**-5.5313**	**3.9962**	**-8.2254**	**-4.1563**	**-2.7896**	**-2.3514**	**18.8878**	
PC2	0.2308	0.0701	0.0008	0.0074	0.0017	0.0066	0.0148	0.2048	0.3448	0.0413	0.0008	0.0202	0.0116	0.0812	0.1779	0.0975	0.0090	0.2468
	8.2351	**-4.1262**	-0.4306	-1.2981	-0.6288	1.2208	1.8399	**7.6288**		**-3.1214**	0.4150	**-2.1601**	1.6297	**4.4694**	**6.9934**	**4.9414**	-1.4338	
PC3	0.0459	0.0105	0.0171	0.0107	0.0007	0.0001	0.0079	0.0130	0.0576	0.0260	0.0759	0.0145	0.0411	0.0118	0.0405	0.0155	0.0346	0.2261
	3.2956	1.5462	1.9852	-1.5623	0.3937	-0.1455	-1.3417	1.7280		**2.4570**	**-4.3073**	-1.8265	**3.1130**	1.6462	**3.0894**	1.8862	**2.8465**	
PC4	0.0504	0.0045	0.0008	0.0001	0.0046	0.0004	0.0000	0.0005	0.1488	0.1082	0.0018	0.0061	0.0133	0.0030	0.0001	0.0092	0.0130	0.1572
	3.4621	-1.0140	0.4354	0.1626	1.0257	0.3016	-0.0409	-0.3211		**-5.2358**	0.6318	-1.1808	1.7446	-0.8229	-0.1518	1.4510	-1.7251	
PC5	0.0131	0.0065	0.0166	0.0291	0.0005	0.0107	0.0232	0.0059	0.0306	0.0004	0.0212	0.0090	0.0009	0.0022	0.0094	0.0039	0.0007	0.0661
	1.7323	-1.2198	1.9553	**-2.6043**	0.3414	-1.5660	**-2.3150**	1.1604		-0.3018	**-2.2121**	1.4289	0.4605	-0.7057	1.4613	0.9405	0.4038	
jan94 - sep98																		
PC1	0.0265	0.0896	0.0234	0.0226	0.2246	0.2114	0.2978	0.0253	0.4355	0.0001	0.0770	0.0749	0.0681	0.2465	0.3134	0.3798	0.0087	0.5786
	1.2234	**-2.3267**	1.1469	-1.1286	**-3.9910**	**-3.8400**	**-4.8291**	1.1943		-0.0873	**2.1413**	**-2.1103**	**2.0050**	**4.2415**	**5.0104**	**5.8038**	-0.6949	
PC2	0.2042	0.0047	0.0099	0.0035	0.0177	0.0071	0.0320	0.4223	0.4190	0.6580	0.0568	0.0013	0.0038	0.1362	0.0009	0.0008	0.6669	0.8657
	-3.7564	0.5114	0.7398	-0.4414	0.9956	0.6283	-1.3475	**-6.3402**		**10.2866**	-1.8194	-0.2717	-0.4610	**-2.9448**	0.2172	-0.2100	**10.4926**	
PC3	0.0551	0.0555	0.0051	0.0088	0.0466	0.0021	0.0000	0.0262	0.1011	0.0144	0.0255	0.0004	0.0121	0.0000	0.0375	0.0369	0.0031	0.1108
	-1.7917	-1.7976	0.5298	-0.6972	-1.6395	-0.3413	0.0406	-1.2176		-0.8973	-1.1992	0.1399	-0.8215	-0.0095	-1.4644	1.4508	0.4121	
PC4	0.0163	0.0067	0.0303	0.0180	0.0989	0.0670	0.0008	0.0233	0.1487	0.0428	0.0011	0.0271	0.0311	0.0306	0.0644	0.0010	0.0007	0.1267
	-0.9549	-0.6086	-1.3110	1.0033	**2.4565**	-1.9879	0.2132	-1.1462		1.5676	0.2455	-1.2370	1.3290	1.3173	-1.9457	-0.2366	0.2029	
PC5	0.0015	0.0001	0.0089	0.0046	0.0018	0.0741	0.0054	0.0212	-0.0281	0.0062	0.1410	0.0045	0.0150	0.0220	0.0399	0.0130	0.1054	0.4415
	-0.2913	0.0693	-0.7012	0.5043	-0.3143	**2.0974**	0.5489	-1.0914		-0.5871	**-3.0041**	-0.5011	0.9157	-1.1129	1.5120	-0.8500	**-2.5455**	
oct98 - mar00																		
PC1	0.2589	0.0419	0.0269	0.0250	0.1682	0.1489	0.1259	0.2364	0.2312	0.0396	0.3889	0.0119	0.0023	0.0723	0.0372	0.0865	0.2526	0.6690
	2.3643	-0.8364	-0.6656	0.6408	-1.7984	-1.6730	-1.5178	**2.2257**		0.8121	**-3.1908**	-0.4384	-0.1914	-1.1163	-0.7859	-1.2306	**2.3251**	
PC2	0.2680	0.1996	0.0520	0.0995	0.0587	0.0443	0.0120	0.0150	0.4095	0.5343	0.3858	0.0800	0.2583	0.0187	0.0488	0.0631	0.0085	0.7810
	-2.4204	-1.9972	0.9368	-1.3293	0.9985	0.8610	0.4414	0.4944		**4.2847**	**3.1700**	-1.1793	**2.3603**	-0.5525	-0.9057	-1.0378	0.3712	
PC3	0.3386	0.0004	0.0013	0.0040	0.0004	0.0151	0.2622	0.5147	0.3690	0.2041	0.0130	0.0070	0.0208	0.0085	0.0720	0.0030	0.2163	0.0488
	-2.8621	-0.0798	0.1443	-0.2532	0.0841	0.4957	**2.3845**	**-4.1194**		**-2.0257**	-0.4586	-0.3355	0.5831	0.3693	1.1139	0.2205	**-2.1014**	
PC4	0.1795	0.2275	0.0172	0.0592	0.0571	0.0066	0.0733	0.0076	-0.2892	0.0728	0.0864	0.0107	0.0249	0.1100	0.0260	0.1212	0.1897	0.1706
	-1.8709	**-2.1708**	0.5291	-1.0036	-0.9840	0.3249	1.1248	-0.3501		1.1209	1.2303	-0.4168	0.6386	1.4061	0.6540	-1.4854	1.9353	
PC5	0.0192	0.0099	0.0183	0.0512	0.0121	0.0491	0.1594	0.0257	0.2471	0.0148	0.0936	0.0933	0.1079	0.1342	0.0549	0.0228	0.0320	0.1222
	-0.5596	0.4008	-0.5456	0.9296	0.4436	0.9084	-1.7415	-0.6500		-0.4904	1.2857	1.2828	-1.3914	1.5748	-0.9641	0.6108	-0.7276	
apr00 - dec02																		
PC1	0.6742	0.0371	0.2505	0.2844	0.0161	0.2734	0.0589	0.5048	0.7906	0.6967	0.0380	0.2303	0.2891	0.0109	0.2364	0.0702	0.5692	0.7653
	-8.0093	1.0937	**-3.2190**	**3.5098**	0.7127	**3.4152**	1.3930	**-5.6216**		**-8.4377**	1.1072	**-3.0459**	**3.5510**	0.5840	**3.0977**	1.5297	**-6.3997**	
PC2	0.0770	0.1100	0.0174	0.0001	0.2772	0.0619	0.0261	0.1605	0.5808	0.0621	0.2793	0.0193	0.0054	0.1159	0.0872	0.0177	0.2020	0.5618
	-1.6083	1.9576	0.7402	0.0439	**-3.4482**	-1.4302	-0.9107	**-2.4347**		1.4326	**-3.4659**	-0.7802	-0.4116	**2.0155**	1.7213	0.7481	**2.8012**	
PC3	0.0048	0.0019	0.1998	0.1398	0.0219	0.1109	0.0037	0.0040	0.2360	0.0001	0.0019	0.0226	0.0214	0.0218	0.0006	0.0534	0.0002	-0.1664
	0.3850	-0.2443	**2.7825**	**-2.2445**	-0.8330	1.9663	-0.3378	0.3548		0.0519	0.2461	-0.8470	0.8232	0.8315	0.1382	1.3223	0.0687	
PC4	0.0096	0.0140	0.0169	0.0158	0.0009	0.0272	0.0122	0.0024	-0.1253	0.0908	0.0085	0.0984	0.0989	0.0345	0.0032	0.0543	0.0356	0.0424
	0.5477	0.6623	0.7289	-0.7062	0.1659	-0.9302	0.6179	-0.2723		1.7593	-0.5148	1.8395	-1.8442	-1.0527	0.3154	-1.3342	1.0693	
PC5	0.0005	0.0593	0.0231	0.0074	0.0349	0.0086	0.0254	0.0361	0.0613	0.0373	0.0298	0.0023	0.0087	0.1310	0.1411	0.0055	0.0015	0.1808
	-0.1182	-1.3985	-0.8561	0.4803	-1.0582	0.5181	0.8993	1.0775		1.0961	0.9755	-0.2663	0.5202	**2.1614**	**-2.2568**	-0.4148	0.2164	
jan03 - jun07																		
PC1	0.1330	0.1140	0.0002	0.0019	0.0431	0.1980	0.0632	0.1882	0.2607	0.4484	0.3695	0.0317	0.0021	0.0001	0.1405	0.0445	0.6715	0.7508
	-2.8238	**2.5865**	0.1042	0.3111	-1.5303	**-3.5831**	-1.8728	**-3.4720**		**6.5010**	**-5.5205**	-1.3056	0.3313	0.0586	**2.9158**	1.5569	**10.3092**	
PC2	0.3657	0.3659	0.0002	0.0201	0.0136	0.0023	0.0279	0.4909	0.5676	0.0001	0.0000	0.0106	0.0099	0.0421	0.0958	0.0063	0.0051	0.0123
	5.4754	**-5.4773**	0.0935	-1.0331	-0.8480	0.3495	1.2217	**7.0816**		0.0702	0.0137	0.7470	-0.7214	-1.5116	**-2.3471**	-0.5738	0.5160	
PC3	0.1907	0.0051	0.0293	0.0572	0.0043	0.0021	0.0417	0.0183	0.5442	0.1320	0.0351	0.0003	0.0126	0.0005	0.0003	0.0071	0.0001	0.2558
	-3.5008	0.5143	-1.2524	1.7765	0.4736	0.3279	1.5036	0.9852		**2.8121**	-1.3761	-0.1272	-0.8152	-0.1632	0.1329	-0.6078	-0.0616	
PC4	0.0222	0.0121	0.0109	0.0000	0.0001	0.0145	0.0141	0.0044	0.0407	0.2039	0.0287	0.0508	0.0706	0.0005	0.0005	0.0187	0.0162	0.2479
	1.0858	-0.7973	0.7556	0.0237	0.0743	-0.8744	-0.8635	0.4778		**-3.6493**	1.2387	-1.6674	1.9872	-0.1649	0.1584	0.9952	-0.9266	
PC5	0.0031	0.0368	0.0250	0.0122	0.0198	0.0006	0.0008	0.0487	0.0823	0.0625	0.1059	0.0227	0.0016	0.0019	0.0010	0.0039	0.0798	0.1841
	-0.4020	-1.4099	1.1537	-0.8014	1.0253	0.1718	0.2003	-1.6322		-1.8620	**2.4820**	-1.1001	0.2882	0.3162	0.2336	-0.4532	**-2.1235**	
jul07 - apr09																		
PC1	0.7835	0.2545	0.0054	0.0122	0.1161	0.1838	0.1310	0.6551	0.7620	0.4995	0.0730	0.0362	0.3067	0.2168	0.3044	0.1431	0.7742	0.8577
	-8.5067	**2.6128**	0.3280	0.4972	1.6209	**2.1222**	1.7363	**-6.1629**		**-4.4677**	1.2553	-0.8666	**2.9745**	**2.3530**	**2.9583**	1.8276	**-8.2815**	
PC2	0.0021	0.0178	0.0029	0.0057	0.0121	0.0072	0.1238	0.1508	0.6382	0.1576	0.2112	0.0750	0.0012	0.0008	0.0380	0.1455	0.0030	0.5364
	0.2070	0.6012	-0.2404	0.3372	0.4954	-0.3817	1.6812	1.8849		-1.9346	**2.3142**	1.2736	-0.1519	0.1279	0.8894	1.8451	-0.2463	
PC3	0.0492	0.1465	0.1635	0.2657	0.2476	0.0961	0.1252	0.0043	0.5187	0.1746	0.2215	0.0810	0.1848	0.0437	0.0128	0.0700	0.1041	0.6247
	1.0174	-1.8532	-1.9769	**2.6901**	**2.5656**	1.4583	1.6917	-0.2929		**-2.0569**	**2.3856**	1.3273	**-2.1296**	-0.9556	-0.5096	-1.2270	-1.5247	
PC4	0.0332	0.0045	0.0410	0.0809	0.0705	0.0436	0.0766	0.0595	-0.3233	0.1699	0.0257	0.0508	0.0060	0.0239	0.0057	0.0137	0.1335	-0.1506
	0.8289	0.2994	0.9243	-1.3266	-1.2320	-0.9550	-1.2883	1.1253		**2.0236**	-0.7260	-1.0343	0.3461	-0.7001	-0.3373	-0.5278	1.7554	
PC5	0.0716	0.0054	0.2841	0.2250	0.0067	0.0128	0.0004	0.0596	0.1950	0.0017	0.0299	0.0814	0.0342	0.0711	0.0080	0.0172	0.0575	0.0525
	1.2416	-0.3310	**-2.8171**	**2.4098**	0.3679	0.5088	0.0910	1.1256		0.1862	0.7855	-1.3311	0.8416	1.2376	-0.4021	-0.5922	1.1051	
may09 - dec12																		
PC1	0.6031	0.2858	0.3045	0.3346	0.4566	0.1088	0.0900	0.5478	0.7456	0.7596	0.1765	0.3581	0.3837	0.2175	0.0200	0.0462	0.8156	0.9000
	7.9896	**-4.1001**	**4.2878**	**-4.5961**	**-5.9406**	**-2.2639**	**-2.0386**	**7.1325**		**11.5187**	**-3.0005**	**4.8406**	**-5.1132**	**-3.4165**	-0.9256	-1.4258	**13.6289**	
PC2	0.1769	0.0007	0.0006	0.0015	0.0243	0.1564	0.1086	0.1681	0.4944	0.0070	0.0113	0.0006	0.0040	0.0234	0.0094	0.0237	0.0000	0.0415
	-3.0047	0.1736	-0.1611	-0.2483	-1.0228	**-2.7904**	**-2.2620**	**-2.9128**		-0.5446	-0.6926	0.1556	0.4119	-1.0042	0.6306	1.0092	-0.0446	
PC3	0.0000	0.0003	0.0096	0.0117	0.0506	0.0286	0.0421	0.0743	0.3215	0.0069	0.0093	0.0005	0.0007	0.0051	0.0023	0.0525	0.0034	-0.0677
	-0.0345	0.1071	0.6381	-0.7043	1.4961	-1.1113	-1.3593	-1.8365		0.5420	0.6277	-0.1450	0.1726	-0.4657	-0.3127	1.5255	-0.3766	
PC4	0.0000	0.0203	0.0011	0.0276	0.0038	0.0000	0.1845	0.0192	0.2874	0.0364	0.1310	0.0082	0.0036	0.3432	0.0881	0.0267	0.0398	0.3095
	0.0109	0.9324	-0.2197	1.0925	-0.3986	0.0032	**3.0827**	-0.9069		-1.2590	**2.5163**	-0.5881	0.3907	**4.6850**	**2.0144**	1.0729	-1.3199	
PC5	0.0000	0.0337	0.0003	0.0053	0.0063	0.0086	0.1204	0.0136	0.1954	0.0009	0.0185	0.0144	0.0164	0.0166	0.2679	0.2484	0.0000	0.2476

Table 6: All Funds Except Fund of Funds (NoFoFs)

R^2s from regressing each extracted principal components onto each hedge fund risk factor individually along with associated T-stats beneath. The adjusted R^2 column represents multivariate regressions of each principal component regressed onto all factors of the 8-factor model of Fung and Hsieh (2004, 2006).

	SP-Rf	SP-RL	TY-Rf	BAA-TY	PTFSBD-Rf	PTFSFX-Rf	PTFSCOM-Rf	IFC-Rf	Adjusted R^2	SP-Rf	SP-RL	TY-Rf	BAA-TY	PTFSBD-Rf	PTFSFX-Rf	PTFSCOM-Rf	IFC-Rf	Adjusted R^2
	Principal Components Analysis (PCA) - Fung and Hsieh (1997a)									Asymptotic Principal Components (APC) - Connor and Korajzyck (1986, 1987)								
94 - 12																		
PC1	0.0020	0.0039	0.0049	0.0039	0.0280	0.0644	0.1436	0.0108	0.2085	0.6228	0.2153	0.0432	0.1421	0.0663	0.0286	0.0220	0.6839	0.8643
	-0.6785	0.9417	-1.0580	0.9419	**2.5512**	**3.9457**	**6.1549**	1.5733		**19.3161**	**-7.8756**	**3.1955**	**-6.1188**	**-4.0074**	**-2.5793**	**-2.2551**	**22.1103**	
PC2	0.2670	0.0770	0.0842	0.1056	0.2000	0.1041	0.0359	0.2731	0.4556	0.0022	0.0152	0.0110	0.0529	0.0076	0.0060	0.0008	0.0112	0.1064
	9.0721	**-4.3418**	**4.5590**	**-5.1657**	**-7.5173**	**-5.1258**	**-2.9026**	**9.2136**		0.7086	1.8669	1.5857	**-3.5546**	-1.3193	-1.1633	-0.4171	1.6012	
PC3	0.0070	0.0025	0.0166	0.0225	0.0020	0.0118	0.0220	0.0004	0.0567	0.0462	0.0096	0.0205	0.0238	0.1383	0.1518	0.2099	0.0054	0.3438
	1.2630	-0.7502	-1.9513	**2.2818**	0.6770	-1.6445	**-2.2537**	0.3029		**-3.3095**	-1.4777	**-2.1759**	**2.3453**	**6.0214**	**6.3596**	**7.7479**	-1.1042	
PC4	0.2193	0.0958	0.0019	0.0041	0.0345	0.0228	0.0085	0.1616	0.4390	0.0191	0.0227	0.0003	0.0069	0.0019	0.0239	0.0062	0.0306	0.1070
	7.9677	**-4.8941**	0.6473	-0.9643	**2.8403**	**2.2967**	1.3933	**6.5999**		**2.0967**	**2.2895**	-0.2716	1.2531	-0.6578	**2.3516**	1.1884	**2.6713**	
PC5	0.0025	0.0046	0.0076	0.0232	0.0000	0.0014	0.0099	0.0003	0.0310	0.0386	0.0016	0.0011	0.0022	0.0149	0.0046	0.0066	0.0002	0.0804
	-0.7505	1.0244	1.3164	**-2.3192**	-0.0807	-0.5644	-1.5044	0.2615		**3.0131**	0.5930	0.4943	-0.7118	1.8471	-1.0196	-1.2252	0.2302	
jan94 - sep98																		
PC1	0.2316	0.0573	0.0000	0.0004	0.1088	0.0140	0.0733	0.0894	0.2874	0.5392	0.1236	0.0021	0.0078	0.2736	0.0070	0.1188	0.4307	0.7732
	4.0710	-1.8281	0.0335	0.1522	**-2.5917**	-0.8834	**-2.0860**	**2.3237**		**-8.0219**	**2.7847**	-0.3426	0.6579	**4.5511**	0.6237	**2.7234**	**-6.4503**	
PC2	0.0029	0.1393	0.0005	0.0000	0.0297	0.0893	0.2460	0.0133	0.2307	0.1857	0.0032	0.0871	0.0541	0.0720	0.0558	0.2457	0.1179	0.5459
	0.4010	**-2.9839**	0.1682	0.0284	-1.2982	**-2.3220**	**-4.2356**	0.8598		**3.5411**	-0.4188	**-2.2909**	1.7738	**2.0653**	1.8029	**4.2322**	**2.7109**	
PC3	0.3168	0.0375	0.1389	0.0901	0.0314	0.0006	0.0636	0.1690	0.5578	0.0037	0.1957	0.0069	0.0045	0.0057	0.0206	0.1526	0.0005	0.1647
	5.0500	-1.4634	**-2.9791**	**2.3342**	1.3355	0.1860	1.9329	**3.3442**		-0.4513	**-3.6586**	0.6195	-0.4969	-0.5612	-1.0748	**-3.1473**	0.1657	
PC4	0.0981	0.1799	0.0394	0.0044	0.1197	0.0011	0.0028	0.0291	0.3077	0.0293	0.0193	0.0112	0.0076	0.0123	0.0418	0.0238	0.0276	-0.0148
	-2.4464	**-3.4734**	1.5024	-0.4951	**2.7346**	-0.2512	-0.3923	-1.2848		-1.2876	-1.0403	0.7878	-0.6504	0.8261	1.5482	1.1582	-1.2484	
PC5	0.0225	0.1420	0.0001	0.0001	0.0066	0.0492	0.0338	0.0003	0.0934	0.0609	0.0369	0.0630	0.1082	0.0009	0.0008	0.0081	0.0133	0.0972
	1.1259	**3.0176**	-0.0704	0.0733	0.6029	1.6874	1.3881	0.1362		-1.8893	-1.4526	1.9236	**-2.5831**	0.2172	-0.2044	0.6713	-0.8605	
oct98 - mar00																		
PC1	0.1635	0.1901	0.0051	0.0013	0.2305	0.0921	0.1260	0.2162	0.6423	0.0152	0.5273	0.0076	0.0052	0.1787	0.0278	0.0471	0.1598	0.7357
	1.7682	-1.9378	-0.2851	-0.1416	**-2.1890**	-1.2737	-1.5190	**2.1011**		0.4969	**-4.2245**	-0.3491	-0.2889	-1.8659	-0.6760	-0.8896	1.7446	
PC2	0.1505	0.1370	0.0321	0.0822	0.0808	0.0092	0.1049	0.1318	0.0517	0.5432	0.4323	0.0460	0.1114	0.0236	0.0336	0.0874	0.1449	0.4886
	1.6836	1.5939	-0.7282	1.1971	1.1861	0.3853	-1.3692	1.5583		**4.3616**	**3.4902**	-0.8781	1.4166	0.6222	-0.7454	-1.2381	1.6468	
PC3	0.0961	0.0000	0.0291	0.0247	0.0224	0.0010	0.1043	0.0140	0.4027	0.0614	0.0011	0.0465	0.0186	0.1778	0.0724	0.0032	0.0049	0.3334
	-1.3041	0.0021	-0.6922	0.6365	-0.6057	-0.1289	1.3651	-0.4760		-1.0227	-0.1348	-0.8830	0.5506	1.8599	1.1177	0.2249	-0.2812	
PC4	0.4870	0.1238	0.0219	0.0319	0.0534	0.2055	0.0941	0.3040	0.4710	0.0389	0.0048	0.0316	0.1204	0.0195	0.2060	0.0008	0.1019	0.6432
	-3.8974	-1.5036	0.5983	-0.7265	0.9499	**2.0346**	1.2889	**-2.6437**		-0.8046	-0.2785	-0.7225	1.4798	0.5645	**2.0375**	-0.1164	-1.3473	
PC5	0.0002	0.0286	0.0056	0.0123	0.0002	0.0471	0.0174	0.0145	0.0192	0.0359	0.0073	0.0042	0.0052	0.0088	0.0536	0.0088	0.0963	-0.2710
	0.0580	0.6861	0.3003	0.4456	0.0535	0.8894	-0.5317	0.4844		-0.7724	-0.3425	-0.2610	0.2890	-0.3778	-0.9520	0.3775	-1.3059	
apr00 - dec02																		
PC1	0.8393	0.0928	0.1625	0.2445	0.0159	0.1680	0.0342	0.6903	0.9146	0.7694	0.1423	0.0883	0.1900	0.0119	0.0814	0.0280	0.7593	0.8878
	-12.7219	1.7812	**-2.4523**	**3.1674**	-0.7066	**2.5018**	1.0470	**-8.3119**		**-10.1709**	**2.2678**	-1.7324	**2.6964**	-0.6111	1.6577	0.9449	**-9.8879**	
PC2	0.0598	0.0240	0.2389	0.2072	0.2018	0.2328	0.2238	0.0536	0.6791	0.1339	0.0101	0.2570	0.1789	0.0550	0.1701	0.2913	0.0452	0.5240
	1.4047	0.8731	**3.1191**	**-2.8466**	**-2.7996**	**-3.0669**	**-2.9900**	1.3246		**2.1888**	0.5618	**3.2746**	**-2.5985**	-1.3428	**-2.5210**	**-3.5699**	1.2108	
PC3	0.0493	0.2027	0.0296	0.1089	0.1465	0.0888	0.0073	0.1735	0.3872	0.0039	0.2630	0.0004	0.0089	0.0195	0.1334	0.0272	0.0021	0.3293
	1.2686	**-2.8069**	0.9720	-1.9465	**2.3063**	1.7387	0.4767	**2.5509**		0.3476	**-3.3259**	-0.1167	-0.5274	0.7843	**2.1845**	0.9313	0.2532	
PC4	0.0245	0.1291	0.0000	0.0046	0.0257	0.0360	0.0662	0.0034	0.1008	0.0000	0.0000	0.0268	0.0063	0.0487	0.0175	0.0020	0.0571	0.1848
	0.8831	**2.1436**	0.0244	0.3773	0.9046	-1.0753	-1.4822	-0.3258		-0.0074	0.0039	-0.9232	0.4422	1.2599	0.7438	0.2511	1.3706	
PC5	0.0017	0.0115	0.0560	0.0095	0.0228	0.0027	0.0005	0.0033	0.0367	0.0037	0.0377	0.0447	0.1009	0.1224	0.0128	0.0119	0.0386	0.4084
	-0.2303	0.6013	-1.3564	0.5451	0.8512	-0.2904	0.1208	0.3226		-0.3397	-1.1021	1.2044	-1.8654	**-2.0789**	-0.6349	0.6109	1.1156	
jan03 - jun07																		
PC1	0.3954	0.4052	0.0072	0.0001	0.0000	0.1412	0.1323	0.6377	0.7372	0.5472	0.5145	0.0143	0.0002	0.0000	0.1245	0.0732	0.7291	0.8837
	5.8322	**-5.9522**	-0.6125	-0.0760	0.0226	**2.9242**	**2.8159**	**9.5663**		**7.9272**	**-7.4232**	-0.8701	-0.1128	-0.0116	**2.7197**	**2.0273**	**11.8292**	
PC2	0.1313	0.0881	0.0049	0.0070	0.0765	0.0469	0.0415	0.0266	0.2022	0.0005	0.0071	0.0000	0.0000	0.0304	0.0070	0.0053	0.0087	-0.0941
	2.8031	**-2.2420**	0.5037	-0.6069	**-2.0756**	-1.5998	-1.5012	1.1912		0.1633	-0.6104	-0.0036	0.0267	-1.2770	-0.6069	-0.5243	-0.6737	
PC3	0.1848	0.0095	0.0008	0.0048	0.0015	0.0009	0.0548	0.0006	0.3714	0.1774	0.0307	0.0124	0.0380	0.0027	0.0037	0.0793	0.0044	0.5140
	-3.4338	0.7067	-0.2097	0.4993	-0.2824	-0.2168	1.7361	0.1713		**-3.3486**	1.2839	-0.8072	1.4324	-0.3780	0.4381	**2.1156**	0.4795	
PC4	0.0165	0.0003	0.0468	0.0696	0.0088	0.0473	0.0885	0.0000	0.1376	0.0150	0.0004	0.0377	0.0190	0.0583	0.0663	0.0037	0.0000	0.0907
	-0.9350	0.1262	-1.5977	1.9729	0.6812	-1.6061	**-2.2475**	-0.0141		-0.8899	-0.1429	1.4266	-1.0036	-1.7944	-1.9222	-0.4374	-0.0333	
PC5	0.0268	0.0007	0.0000	0.0056	0.0256	0.0089	0.0057	0.0110	0.0435	0.0468	0.0029	0.0053	0.0207	0.0289	0.0127	0.0810	0.0088	0.0637
	1.1966	0.1897	0.0129	-0.5395	-1.1690	0.6852	-0.5456	-0.7609		-1.5987	0.3875	0.5262	-1.0475	1.2429	0.8177	**2.1403**	-0.6777	
jul07 - apr09																		
PC1	0.8049	0.3268	0.0001	0.0430	0.1903	0.1906	0.1681	0.6365	0.8417	0.7683	0.2497	0.0012	0.1750	0.1876	0.3195	0.1798	0.8539	0.9405
	-9.0848	**3.1159**	-0.0524	0.9476	**2.1679**	**2.1705**	**2.0103**	**-5.9178**		**-8.1436**	**2.5798**	-0.1529	**2.0598**	**2.1491**	**3.0641**	**2.0942**	**-10.8119**	
PC2	0.0006	0.0319	0.0354	0.0684	0.0184	0.0518	0.0162	0.1801	0.4975	0.0517	0.1604	0.0166	0.0195	0.0082	0.0022	0.0708	0.0348	0.6049
	0.1060	0.8124	0.8567	-1.2118	-0.6129	-1.0455	0.5733	**2.0961**		-1.0447	1.9549	0.5803	-0.6301	0.4066	-0.2101	1.2343	0.8498	
PC3	0.0926	0.0687	0.4713	0.5048	0.1801	0.1182	0.1809	0.0390	0.5821	0.1011	0.1376	0.4770	0.4520	0.2163	0.0127	0.0341	0.0835	0.6529
	1.4290	-1.2147	**-4.2228**	**4.5156**	**2.0960**	1.6374	**2.1019**	0.9006		1.4997	-1.7860	**-4.2707**	**4.0618**	**2.3491**	0.5066	0.8397	1.3502	
PC4	0.0553	0.0335	0.0025	0.0271	0.1410	0.0007	0.0148	0.0634	0.0919	0.0101	0.0036	0.1854	0.0012	0.1661	0.0894	0.0925	0.0006	0.7680
	1.0817	0.8331	0.2237	-0.7459	-1.8117	-0.1178	-0.5488	1.1637		0.4509	-0.2703	**2.1333**	-0.1526	-1.9959	1.4013	1.4278	0.1134	
PC5	0.0189	0.0002	0.1316	0.0082	0.0004	0.0592	0.0679	0.0051	0.4848	0.0015	0.0050	0.0000	0.0185	0.0644	0.0016	0.0030	0.0047	-0.2075
	-0.6215	-0.0664	-1.7406	0.4078	0.0938	-1.1220	-1.2067	0.3200		-0.1755	-0.3180	0.0014	-0.6141	-1.1731	0.1777	0.2455	-0.3074	
may09 - dec12																		
PC1	0.7733	0.4130	0.2956	0.2659	0.4024	0.0841	0.0684	0.6657	0.8658	0.8233	0.2301	0.3226	0.3096	0.2086	0.0215	0.0292	0.8354	0.9181
	11.9690	**-5.4359**	**4.1982**	**-3.9007**	**-5.3183**	-1.9634	-1.7562	**9.1450**		**13.9904**	**-3.5428**	**4.4720**	**-4.3402**	**-3.3269**	-0.9606	-1.1235	**14.5995**	
PC2	0.0355	0.0145	0.0000	0.0066	0.1158	0.1498	0.1592	0.0502	0.4051	0.0100	0.0163	0.0317	0.0556	0.0037	0.0328	0.0957	0.0283	-0.0323
	-1.2431	-0.7874	-0.0253	-0.5290	**-2.3458**	**-2.7202**	**-2.8199**	-1.4899		-0.6506	0.8346	-1.1726	1.5725	0.3932	1.1928	**2.1079**	-1.1053	
PC3	0.0000	0.0008	0.0014	0.0111	0.0096	0.0041	0.0032	0.0396	0.0888	0.0612	0.2129	0.0462	0.0257	0.3556	0.2883	0.2308	0.0013	0.5764
	0.0237	0.1846	-0.2414	0.6875	-0.6374	-0.4144	0.3693	1.3162		-1.6540	**3.3702**	-1.4267	1.0529	**4.8143**	**4.1252**	**3.5498**	-0.2363	
PC4	0.0311	0.0033	0.0001	0.0016	0.0267	0.0000	0.0008	0.0020	-0.0691	0.0220	0.0090	0.0002	0.0319	0.0139	0.0012	0.0014	0.0461	0.2453
	1.1609	-0.3744	0.0697	0.2633	-1.0730	-0.0047	-0.1828	0.2929		-0.9729	0.6184	0.0878	-1.1765	-0.7692	-0.2218	-0.2388	-1.4246	
PC5	0.0231	0.0209	0.0091	0.0126	0.0019	0.0254	0.0029	0.0019	-0.0421	0.0282	0.0262	0.0020	0.0000	0.0010	0.0003	0.0104	0.0086	0.0874