Arbitrage trading: the long and the short of it

Yong Chen†
Texas A&M University

Zhi Da‡
University of Notre Dame

Dayong Huang§
University of North Carolina at Greensboro

First version: December 1, 2014
This version: March 12, 2015

† Mays Business School, Texas A&M University, ychen@mays.tamu.edu.
‡ Mendoza College of Business, University of Notre Dame, zda@nd.edu.
§ Bryan School of Business, University of North Carolina at Greensboro, d_huang@uncg.edu.

Arbitrage trading: the long and the short of it

March 12, 2015

Abstract

We measure arbitrage trading on both the long- and the short-sides by merging hedge fund equity holdings with short interest. Over time, aggregate hedge fund holdings track aggregate short interest well, and both have grown dramatically since the early 1990s. In the cross section, the difference between abnormal hedge fund holdings and abnormal short interest, which captures net arbitrage trading activity on a stock, strongly predicts future stock returns. When examining a broad set of asset pricing anomalies, we find anomaly returns to come exclusively from about 30% of the anomaly stocks that are traded by arbitrageurs. These stocks are also hard to arbitrage on average. Overall, our findings confirm that mispricing arises from limits to arbitrage and arbitrage trading is informative about mispricing.

Keywords: Hedge fund holdings, short interest, asset pricing anomalies, limits to arbitrage

JEL Classification: G11, G23

1. Introduction

Arbitrageurs play a crucial role in finance. By simultaneously taking long and short positions in different assets, they help to eliminate relative mispricing and therefore enforce market efficiency. As a result, their trading pins down the expected return on these assets, according to the seminal arbitrage pricing theory (APT) of Ross (1976). On the other hand, investors' behavioral biases may lead to persistent mispricing when arbitrageurs face limits-to-arbitrage (Shleifer and Vishny (1997) among others).

Tracking arbitrage trading has been a challenging empirical task due to the lack of data on arbitrageurs. In recent years, as hedge funds emerge as a group of likely arbitrageurs and their stock holdings data become available, a series of papers have inferred the long-side of arbitrage trading by investigating their quarterly holdings (e.g., Brunnermeier and Nagel (2004), Griffin and Xu (2009), Cao, Chen, Goetzmann, and Liang (2014)). At the same time, short sellers are often considered as informed traders (e.g., Desai, Ramesh, Thiagarajan, and Balachandran (2002) and Asquith, Pathak, and Ritter (2005)). Since short positions are involved in an arbitrage trade, several researchers have started to track the short-side of arbitrage trading by examining short interest on stocks over time (see Hanson and Sunderam (2014), Hwang and Liu (2014), Wu and Zhang (2014), among others).

The innovation of our paper is to combine hedge fund holdings on the long-side with the short interest on the short-side to infer the *net* arbitrage trading on a stock. The advantage of our approach is straightforward. A correctly priced stock can be traded by arbitrageurs for hedging purpose. It might be bought in some arbitrage transactions and sold short in others. Alternatively, arbitrageurs may disagree on the valuation of a stock, and thus some may purchase it while others sell it short. At the end of 2012, there are more than 2,300 stocks with both hedge fund holdings and short interest and they cover more than 90% of the U.S. equity universe in terms of market capitalization. For these stocks, focusing on either the long-side or the short-side alone will give imprecise inference about arbitrageurs' view on the stocks in aggregate. However, the *net* position should represent a better proxy for arbitrage trading and a more powerful predictor of future stock returns. Indeed, we confirm this conjecture in our empirical analysis. In particular, we find that stocks with large hedge fund holdings but simultaneously heavy short interest do not earn any abnormal return in the future.

We combine a comprehensive dataset on hedge funds' quarterly holdings with data on short interest during the period from 1990 to 2012.[1] Over time, aggregate hedge fund holdings track aggregate short interest well, and both experienced exponential growth since the early 1990s as plotted in Figure 1. The percentages of shares outstanding that are held by hedge funds or sold short are both less than 1% in earlier 1990s but peak around 5% in 2008 before levering off afterwards. The common trend shared by both the long- and the short-side of arbitrage trading confirms the increasing arbitrage activities documented by Hanson and Sunderam (2014) who use only short interest. Even in the cross section, we find similar distribution in hedge fund holdings and short interest. For example, these two values exhibit similar means (3.72% vs. 3.49%), medians (2.37% vs. 2.35%), and standard deviations (3.97% vs. 3.66%) across stocks. The similarity in their distributions supports the notion that on average, hedge fund holdings and short interest reveal the two legs of the same arbitrage trade.

Stocks can be held by hedge funds or sold short for many reasons other than arbitrage. For example, hedge funds may hold certain stocks to neutralize portfolio risk. Stocks may be sold short to hedge against a convertible bond purchase. To better measure arbitrage *trading*, we define abnormal hedge fund holding (AHF) and abnormal short interest (ASR) as their values in the current quarter minus their moving averages in the four prior quarters. Figure 2 shows that, at the aggregate level, AHF and ASR track each other well. This is particularly true during crisis periods

[1] We exclude small stocks and penny stocks from our main analysis to minimize measurement errors and market microstructure-related noise. We confirm that our return results are similar when we add back small stocks.

when mispricing is prevalent. Finally, AHFSR, defined as the difference between AHF and ASR, is our *net* arbitrage trading measure which captures trade imbalance of arbitrageurs. For example, an AHFSR of 1% (-1%) on a stock means that arbitrageurs, as a group, have purchased (sold) and additional 1% of the stock during the most recent quarter relative to their past averages.

Consistent with the existing literature, we find both abnormal hedge fund holding (AHF) and abnormal short interest (ASR) to predict returns. On the long-side, stocks in the highest AHF quintile outperform those in the lowest quintile by 0.44% per month in the next quarter. On the short-side, stocks in the highest ASR quintile underperform those in the lowest quintile by 0.41% per month in the next quarter. Most important, by focusing on *net* arbitrage trading, AHFSR generates the highest return spread in the same sample.[2] Stocks in the highest AHFSR quintile outperform those in the lowest quintile by 0.68% per month in the next quarter. The return spread is highly significant (t-value = 7.93) but declines quickly over time. It drops to 0.42% per month in the second quarter and further to 0.18% per month in the third. The finding suggests that arbitrage is effective in eliminating mispricing.

The strong return predictability of our net arbitrage trading measure (AHFSR) holds in a battery of robustness checks. We obtain similar results when we restrict both hedge fund holdings and short interest to be strictly positive (i.e., excluding zero values) and when we include small stocks in the sample. Further, the return predictability is not explained away by the common risk factors. It is equally strong during the first and second half of our sample period. It also survives Fama-MacBeth cross-sectional regressions that control for other well-known stock return predictors.

[2] In a contemporaneous study, Jiao, Massa, and Zhang (2015) also find that an increase (decrease) in hedge fund holdings accompanied with a decrease (increase) in short interest is informative about future stock return and firm fundamental. They do not define AHFSR and relate it to arbitrage trading and asset pricing anomalies as we focus on in our paper.

The advantage of our net arbitrage trading measure can also be illustrated by a double sort on AHF and ASR. Holding one variable constant, sorting on the second variable still generates large and significant return spreads, suggesting that arbitrage activities on both legs are informative. Interestingly, we find stocks with high-AHF-high-ASR to have about the same future returns as stocks with low-AHF-low-ASR, confirming that future returns are really driven by the net arbitrage trading activity. Finally and not surprisingly, stocks with high-AHF-low-ASR earn much higher future returns than stocks with low-AHF-high-ASR (1.18% vs. 0.40% per month).

Next, after showing that AHFSR captures arbitrage trading well, we apply it to investigate asset pricing anomalies in the cross section. We examine a total of 10 well-known asset pricing anomalies: the book-to-market ratio (BM) of Fama and French (2008), the gross profitability (GP) of Novy-Marx (2012), operating profit (OP) of Fama and French (2015), momentum (MOM) of Jegadeesh and Titman (1993), market capitalization (MC) of Fama and French (2008), asset growth (AG) of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), investment growth (IK) of Xing (2008), net stock issue (NS) of Fama and French (2008), accrual (AC) of Fama and French (2008) and net operating assets (NOA) of Hirshleifer et al. (2004). We verify that the long-minus-short future return spreads averaged across these 10 anomalies are positive and significant in our sample. The return spreads are 0.28%, 0.25%, 0.20%, and 0.15% per month during the first, second, third, and fourth quarters, respectively.[3]

More important, the anomaly returns are completely driven by about 30% of stocks in the long and short portfolios that are traded by the arbitrageurs. We classify an anomaly stock to be traded by the arbitrageurs if (1) it is in the long portfolio and recent bought by arbitrageurs (its AHFSR belongs to the top 30%) or (2) it is in the short portfolio and recently sold short by arbitrageurs (its AHFSR belongs to the bottom 30%). This subset of anomaly stocks earn return spreads of 0.88%, 0.61%, 0.34%, and 0.27% per month during the first, second, third, and fourth

[3] The magnitude is smaller compared to other studies since we use quintile sort instead of the more common decile sort and we exclude small stocks from our main sample.

quarters, respectively, after portfolio formation. In sharp contrast, the other 70% of anomaly stocks that are not traded by arbitrageurs do not earn any significant return spreads in the next four quarters. The fact that abnormal future returns only appear among anomaly stocks traded by the arbitrageurs and these abnormal returns decline quickly during the first year suggests that arbitrageurs are effective in eliminating mispricing.

We then examine stock characteristics that are related to limits-to-arbitrage. We find that anomaly stocks that are not traded by arbitrageurs in general are easy to arbitrage: they have lower idiosyncratic volatility, higher stock price and are more liquid. Since they are easy to arbitrage, they are unlikely to be mispriced, consistent with their lack of abnormal future returns. Overall, our findings confirm that mispricing arises from limits-to-arbitrage and arbitrage trading is informative about mispricing.

Our paper contributes to a growing literature that studies hedge fund holdings and short interest as return predictors and proxies for arbitrageurs' activities.[4] Brunnermeier and Nagel (2004) find that hedge funds ride with the bubble during the Tech Boom period. Griffin and Xu (2009) find that change in hedge fund ownership has predictive power for future stock returns in the cross-section. Griffin, Harris, Shu and Topaloglu (2011) show evidence suggesting that hedge funds destabilized the market during the tech bubble period. Agarwal, Jiang, Tang, and Yang (2013) find that hedge funds' confidential holding is informative about future stock returns. Cao, Chen, Goetzmann, and Liang (2014) find that, compared with other institutional investors, hedge funds hold and trade mispriced stocks, and mispriced stocks with higher hedge fund ownership realize higher future returns and are more likely to get mispricing corrected. Reca, Sias, and Turtle (2015) also find that hedge fund demand shocks predict future stock returns.

[4] Lou and Polk (2013) infer arbitrage activity from the comovement of stock returns.

There is also a large body of literature that examines the information embedded in short interest. Prior research has studied, both theoretically and empirically, the impact of short sales on security returns. Miller (1977) argues that in the presence of heterogeneous beliefs, binding short sale constraints prevent stock prices from fully reflecting negative opinions of pessimistic trades, leading to overpricing and low subsequent returns. Diamond and Verrecchia (1987) show that given the high costs (i.e., no access to proceeds) of short selling, short sales are more likely to be informative trades. Consistent with these theories, several empirical papers document a negative association between short interest and abnormal stock return (e.g., Asquith and Meulbroek 1995; Desai et al., 2002; and Boehmer, Jones, and Zhang, 2008). Using institutional ownership of stocks as a proxy for stock loan supply, Asquith, Pathak, and Ritter (2005) and Nagel (2005) examine the impact of short sale constraints on stock returns. Asquith, Pathak, and Ritter (2005) find that for small stocks with high short interest, low institutional ownership is associated with more negative future returns, which confirms the effect of binding short constraints on stock prices. However, they show that only 5% of the stocks trading on the NYSE, Amex, and Nasdaq have institutional ownership smaller than short interest, suggesting that short sale constraints are not pervasive. Nagel (2005) finds that short sale constraints help explain several cross-sectional stock return anomalies related to book-to-market ratio, analyst forecast dispersion, turnover, and return volatility. More recently, Drechsler and Drechsler (2014) find that short-rebate fee is a more informative signal about overpricing and arbitrageurs' trading on the short-leg.

To our best knowledge, our paper is the first to combine information about arbitrageurs' trading on both the long- and the short-side. Different from prior research that focuses on either the long- or the short-side, our study provides a more complete view about the effect of arbitrageurs' trading activities. We propose a simple measure of the net arbitrage trading. We find this net arbitrage trading measure to better predict future returns. When using the measure to study well-known return anomalies, we find strong and direct evidence supporting the notion that mispricing arises from limits-to-arbitrage and arbitrage trading is informative about mispricing.

The rest of the paper is organized as follows. Section 2 describes our data and sample. Section 3 examines our net arbitrage trading measure (AHFSR) as a stock return predictor. Section 4 uses AHFSR to study asset pricing anomalies. Section 5 concludes.

2. Data and Sample Construction

We start our sample in 1990 as hedge fund holding and short interest data are relatively sparse before that.[5] At the end of each quarter, we exclude from our main sample stocks with a price of less than $5 per share and a market capitalization of less than the 20th percentile of NYSE size breakpoint. We exclude small stocks and penny stocks from our main analysis for two reasons. First, hedge funds only need to report common stock positions greater than 10,000 shares or $200,000 in market value. As a result, hedge fund holdings on small stocks and penny stocks are often underestimated. Second, excluding these stocks helps to alleviate the associated market microstructure noise. Our final sample still represents over 85% of the CRSP universe on average.

2.1. Hedge Fund Holdings

Our data on equity holdings of hedge funds are identical to that used in Cao, Chen, Goetzmann, and Liang (2014). The data are constructed by manually matching the Thomson Reuters 13F institutional ownership database with a comprehensive list of the names of hedge fund companies. The hedge fund company names are collected from six hedge fund databases, including TASS, HFR, CISDM, Bloomberg, Barclay Hedge, and Morningstar databases, augmented with additional sources. Hedge funds, as private investment vehicles, were historically exempt from registering with the SEC as an investment company. However, similar to other institutional investors, hedge fund management companies with more than $100 million in assets under management are required to file quarterly reports disclosing their holdings of registered equity

[5] For example, the aggregate hedge fund holdings and short interest, as a percentage of total market capitalization of the CRSP universe, was less than 1% on average for the quarters prior to 1990.

securities. Common stock positions greater than 10,000 shares or $200,000 in market value are subject to disclosure. As a result of this reporting requirement, hedge fund holdings of small stocks are likely to be neglected in the reporting, and thus excluding small stocks from our main analysis helps to alleviate this bias. 13F filings contain long positions in stocks while short equity positions are not required to be reported.

In our study, a hedge fund is defined as a management company included in a hedge fund database, a firm that self-identifies as a hedge fund, or a firm that imposes a threshold of high-net-worth investors and a performance-based compensation. After collecting a list of hedge fund company names from various sources, we match them with institution names in the 13F data. To address the fact that a hedge fund manager may not appear in any hedge fund databases because of the voluntary nature of reporting to a hedge fund database, a manually checking is used based on a variety of online resources. Further, following Brunnermeier and Nagel (2004) and Griffin and Xu (2009), each identified company is manually checked to ensure that hedge fund management is its primary business using two criteria: first, more than 50% of its clients are either high-net-worth individuals or invested in "other pooled investment vehicle (e.g., hedge funds)", and second, the adviser is compensated by a performance-based fee. Our final sample of hedge fund companies includes 1,517 hedge fund management firms that collectively manage over 5,000 individual hedge funds.

For each stock in our sample, we then compute its quarterly hedge fund holdings (HF) as the number of shares held by all hedge funds at the end of the quarter divided by the total number of shares outstanding. If the stock is not held by even a single hedge fund in that quarter, its HF is set to zero. Since stocks can be held by hedge funds for reasons other than arbitrage (to neutralize portfolio risk from a short position), to better measure hedge fund *trading*, we define abnormal hedge fund holding (AHF) in the current quarter HF minus the average HF in the past four quarters.

2.2 Short Interest

Mid-month short interest is obtained from the Compustat Short Interest file from 1990 to 2012. These monthly short interest data are reported by the NYSE, Amex, and NASDAQ. The Compustat Short Interest file covers NASDAQ stocks only from 2003, and following the literature, we supplement our sample with short interest data on NASDAQ prior to 2003 obtained from the exchanges.

For each stock in our sample, we compute its quarterly short interest (SR) as the number of shares sold short at the end of the quarter divided by the total number of shares outstanding. If the stock is not covered by our short interest files, its SR is set to zero. Again, we define abnormal short interest (ASR) in the current quarter HF minus the average SR in the past four quarters.

2.3 Asset Pricing Anomalies

We consider 10 well-documented popular anomalies largely following Fama and French (2008) and Stambaugh, Yu and Yuan (2012), in our investigation of the relation of hedge fund holdings and short interest to anomalous stock returns.

The first anomaly is the book-to-market ratio (BM) of Fama and French (1996, 2008). It is well known that firms with higher book-to-market ratio have higher returns in the future and these returns do not disappear after adjusting risk using the CAPM of Sharpe (1964) and Lintner (1965). The second anomaly is the operating profit (OP) of Fama and French (2015), who show that firms with higher operating profits have higher future returns. The third anomaly is the gross profitability (GP) of Novy-Marx (2013), who shows that firms with higher gross profit have higher future returns. The fourth anomaly is the momentum (MOM) of Jegadeesh and Titman (1993). In our setting, at the end of each quarter, we compute firm returns in the past 13 months by skipping the immediate month prior to the end of the quarter, divide them into winners and losers, and hold them in the next quarter. The fifth anomaly is the market capitalization (MC) of Fama and French (1996, 2008). On average, the larger the firm size, the lower its expected return. This size anomaly, similar to the book-to-market ratio anomaly, has a long history, survives the CAPM risk adjustment, and has been proposed as a factor in the three-factor model of Fama and French (1996), the five-factor model of Fama and French (2015), and the four-factor model of Hou, Xue, and Zhang (2014). The sixth anomaly is the asset growth (AG) of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), who show that firms with higher growth rates of asset have lower future return. The seventh anomaly is the investment growth (IK) of Xing (2008), who shows that firms with higher investment have lower future returns. The eighth anomaly is the net stock issue (NS) examined in Ritter (1991) and Loughran and Ritter (1995), and Fama and French (2008). The larger the net stock issue, the lower the future returns. The ninth anomaly is the accrual (AC) examined in Sloan (1996) and Fama and French (2008), who find a negative relationship between accrual and future stock returns. The tenth anomaly is the net operating assets (NOA) of Hirshleifer et al. (2004). They show that firms with larger operating assets have lower future returns.

We follow Fama and French (2008), Novy-Marx (2013), Hou Xue, and Zhang (2014) to compute the book-to-market ratio, market capitalization, net stock issue, and accrual. The calculation of gross profit follows Novy-Marx (2013). The calculation of operating profit follows Fama and French (2015). The calculation of momentum follows Jegadeesh and Titman (1993). The calculation of asset growth follows Cooper et al. (2008). The calculation of investment growth follows Xing (2008). The calculation of net operating assets follows Hirshleifer et al. (2004). Details of anomaly construction are in appendix A.

For each anomaly, we construct quintile portfolios at the end of each quarter. We then compute the monthly long-minus-short portfolio return spreads for the next quarter.

2.4 Sample Description

After excluding small stocks and penny stocks, our baseline sample contains about 1,600 stocks per quarter. As shown in Figure 1(a), the number of stocks started around 1,400 in 1990, reached a peak of 2,000 during the Tech bubble, and then leveled off to 1,400 again afterwards.

Since only small stocks and penny stocks are excluded, our baseline sample still covers more than 86% of the CRSP universe in terms of market capitalization.

Figure 1(a) also plots the cross-sectional coverages of the hedge fund holdings (HF) data and the short interest (SR) data over time. While most of the stocks in our sample have positive short interest, the coverage of hedge fund holdings was relatively small at the beginning. For example, in 1990, out of the 1,400 stocks in our sample, less than 1,000 have positive hedge fund holdings.[6] Nevertheless, the hedge fund holdings coverage has increased rapidly. Since 2000, most of the stocks in our sample have both positive hedge fund holdings and short interest. In other words, for almost all stocks in our sample in the last decade, observing both the long- and the short-leg of arbitrage activity should lead to better inference of the *net* arbitrage trading. Figure 1(b) plots the percentage of market cap coverage of hedge fund holdings and short interest. The market cap coverage is large. Stocks with positive hedge fund ownership account for more than 90% of our baseline sample in terms of market cap even in 1990.

Figure 2(a) plots aggregate hedge fund holdings and aggregate short interest over time. As can be seen, aggregate hedge fund holdings and short interest track each other well. They were both less than 1% in earlier 1990s but increased to around 5% in 2008. The abnormal hedge fund holdings (AHF) and abnormal short interest (ASR) also track each other well (with a correlation of 0.26) as shown in Figure 2(b). AHF and ASR are particularly highly correlated during crisis periods when mispricing is widespread. For example, their correlations exceed 60% during the two four-year-periods surrounding the Tech bubble (1999-2002) and the recent financial crisis (2006-2009). Finally, Figure 2(c) plots the difference between hedge fund holdings and short interest, i.e., HFSR, and the difference between abnormal hedge fund holdings and abnormal short interest, i.e., AHFSR. Here, AHFSR can be viewed as a measure of trade imbalance of arbitrageurs. An aggregate AHFSR of 1% (-1%) means that arbitrageurs, as a group, have purchased (sold) an

[6] The coverage of the hedge fund holdings was even smaller prior to 1990, which is why we start our sample in 1990.

additional 1% of the market during the most recent quarter relative to the average of the previous four quarters. We find aggregate AHFSR to fluctuate between -0.5% and 0.5% most of the time. One exceptionally large value (below -1%) of AHFSR occurred during late 2008 as arbitrageurs were fleeing the market due to funding liquidity constraints.

Table 1 Panel A summarizes the cross-sectional distribution of our main variables. We find a similar distribution in HF and SR. For example, across stocks, they have similar means (3.72% vs. 3.49%), medians (2.37% vs. 2.35%), and standard deviations (3.97% vs. 3.66%). AHF and ASR have very similar distributions as well. They have similar means (0.19% vs. 0.18%), medians (0.04% vs. 0.02%), and standard deviations (1.90% vs. 1.83%). The similarity in distribution supports the notion that HF and SR reflect the two legs of the same arbitrage trades on average.

Panel B of Table 1 reports the cross-sectional correlations among our variables. We find a positive correlation of 0.23 between hedge fund holdings (HF) and short interest (SR) across stocks. In other words, a stock with high HF is also likely to have a high SR. It is therefore useful to isolate the net trading on the long- and short-side. When we examine the correlations among AHF, ASR and AHFSR, we find AHFSR to be positively correlated with AHF (0.64) and negatively correlated with ASR (-0.68). These correlations are far from being perfect, suggesting that the net arbitrage trading is quite different from the arbitrage trading on either the long- or the short-side.

3. Arbitrage Trading and Future Returns

Since the majority of stocks are both held by hedge funds and simultaneously sold short, we argue that the *net* arbitrage trading between the long- and the short-side should be a more powerful predictor of future stock returns. In this section, we test the predictor power of our measure of net arbitrage trading for stock returns.

3.1 Portfolio sorts

We first examine whether net arbitrage capital forecasts future stock returns using a portfolio formation approach. Our hedge fund holding data is at quarterly frequency, hence we form portfolios at the end of each quarter and track portfolio returns in the following quarters. Specifically, at the end of each quarter, we rank stocks in our baseline sample based on their values of AHF, ASR or AHFSR, and assign them into quintiles. The lowest quintile includes stocks that have low values of AHF, ASR or AHFSR, and the highest quintile includes stocks that have high values of AHF, ASR and AHFSR. After the formation of the portfolios, we track the excess returns of each portfolio in the following quarters. We compute the excess return of a portfolio by equally averaging excess returns of all stocks that belong to that portfolio in that quarter. We first present the excess returns of these quintile portfolios, and then adjust risk exposures to the three factors of Fama and French (1996), the three factors of Fama and French augmented with the Carhart (1997) momentum factor, and the five factors of Fama and French (2015), which expand their original three factors to include a profitability factor and an asset growth factor.

Table 2 presents results from portfolio formation. Table 2A presents results from using the base sample. Panel A presents results of the AHF quintile portfolios. The results show that stocks experiencing a large increase in hedge fund holding (AHF-quintile 5) have monthly excess return of 1.05% with a t-value of 2.97, while stocks experiencing a large decrease in hedge fund holdings (AHF-quintile 1) have monthly excess return of 0.61% with a t-value of 1.77. The high-minus-low AHF portfolio (AHF-HML) has monthly excess return of 0.44% with a t-value of 4.98. The finding is consistent with the view that hedge fund holdings have return predictive power (see Griffin and Xu (2009), and Cao, Chen, Goetzmann, and Liang (2014) among others).

Panel B presents results of ASR quintile portfolios. Stocks experiencing large increase in short interest (ASR-quintile 5) have monthly excess return of 0.49% with a t-value of 1.32, while stocks experiencing large decrease in short interest (ASR-quintile 1) have monthly excess return of 0.90% with a t-value of 2.70. The high-minus-low ASR portfolio (ASR-HML) has monthly excess return of -0.41% with a t-value of -4.21. The finding confirms the return predictive power in short interest as documented in the extant literature (see Asquith, Pathak, and Ritter (2005) among others).

Panel C combines AHF and ASR and uses AHFSR to sort the same set of stocks into quintiles. Our results show that stocks recently bought by arbitrageurs as a group (AHFSR-quintile 5) have monthly excess return of 1.11% with a t-value of 3.23, while stocks recently sold by arbitrageurs as a group (AHFSR-quintile 1) have monthly excess return of 0.43% with a t-value of 1.20. The high-minus-low AHFSR portfolio (AHFSR-HML) has monthly excess return of 0.68% (or, about 8.2% per year) with a t-value of 7.93. Therefore, the return spread is both economically and statistically significant.

Next, we discuss alphas (i.e., risk-adjusted returns) of these quintile portfolios. The alphas seem to be large in magnitude at extreme quintiles. This is especially true for stocks that have high AHF and stocks that have high ASR. In particular, for the three asset pricing models we consider, high AHF stocks have alphas of 0.28% (t-value = 3.11), 0.34% (t-value = 3.92), and 0.19% (t-value = 2.12), while high ASR stocks have alphas of -0.32% (t-value = -3.09), -0.17% (t-value = -1.91), and -0.35% (t-value = -3.31), respectively. This is not too surprising, since both hedge fund holdings variable (HF) and the short interest variable (SR) are bounded below by zero, and thus an increase in HF or SR tends to be more informative than a decrease.

When AHF and ASR are combined into AHFSR, alphas are large in magnitude for both high and low AHFSR portfolios. High AHFSR stocks have alphas of 0.36%, 0.42%, and 0.27%, and low AHFSR stocks have alphas of -0.34%, -0.22%, -0.38%, respectively. The alphas of high-minus-low portfolios are also larger and statistically significant for the AHFSR portfolio when comparing to those of AHF and ASR portfolios. Across the three factor models, the alphas of AHFSR-HML portfolios are 0.70%, 0.64%, 0.65% with t-values of 8.17, 7.57, 7.20, compared with the alphas of AHF-HML portfolios being 0.40%, 0.38%, 0.35% with t-values of 4.50, 4.27

and 3.80, and the alphas of ASR-HML portfolios being -0.50%, -0.42%, -0.44% with t-values -5.42, -4.71 and -4.58.

Further, we track the excess returns of these quintile portfolios in subsequent quarters in addition to the immediate following quarter.[7] In panel D, the variable Q1 indicates the high-minus-low return in the next quarter, the variable Q2 indicates the high-minus-low returns in the second quarter after portfolio formation, and the variable Q3 indicated high-minus-low returns in the third quarter after portfolio formation, and so on. Our results show that, for all three measures of arbitrage capital, excess returns decrease over time. The high-minus-low excess returns from AHFSR quintile portfolio is the largest at 0.68% per month in the immediate following quarter after portfolio formation. It drops to 0.42% in the second quarter after portfolio formation, further drops to 0.18% in the third quarter after portfolio formation, and finally drops to zero in the fourth quarter after portfolio formation. Overall, the results confirm that *net* arbitrage trading consistently predicts future return better than arbitrage activities on either the long-side or the short-side alone. The fact that these abnormal returns decline quickly during the first year suggests that arbitrageurs are effective in eliminating mispricing.

Table 2B through table 2F provides various robustness checks for our main results. In table 2B, we lift the restriction on firm size that we apply on our base sample. Specifically, we do not exclude those stocks whose market capitalization are less than the 20 percentile size breakpoints of NYSE firms, at the time of portfolio formation. In table 2C, we start with our base sample, and further exclude firms whose hedge fund holdings or short interest equal to zero. In table 2D, we conduct a test using the first half of our base sample which covers January of 1990 to June of 2000, while in table 2E, we use the second half of our base sample which covers July of 2000 to December of 2012. In short, the results from these robustness checks are similar to those we obtain

[7] From a practical perspective, it is also useful to examine subsequent quarters since hedge fund holdings are often reported with a temporal delay averaged about 45 days. See Agarwal, Jiang, Tang, and Yang (2013) and Brown and Schwarz (2013) for analysis of the reporting delay and its implications.

in table 2A. That is, our results are not overly sensitive to application of size breakpoints, or deletion of firms having no hedge fund holdings or short interest, or the choice of the sample period.

So far we have assumed AHF and ASR to be comparable so that a simple difference between them produces a measure of net arbitrage trading. The assumption seems reasonable given that AHF and ASR have similar distributions in the cross-section (see Table 1A). Nevertheless, to account for the possibility that the true net arbitrage trading could be a nonlinear function in both AHF and ASR, we consider an alternative approach to examine the incremental contribution of AHF or ASR by performing two-dimensional independent sorting based on AHF and ASR.

At the end of each quarter, in our base sample, we form tercile portfolios based on AHF, and independently form tercile portfolios based on ASR. Then, nine AHF-ASR portfolios are taken from the intersections of these two sets of tercile portfolios. Our premise is that, in the high AHF tercile, some stocks may have high ASR, but other stocks may have low ASR. Similarly, in the low AHF tercile, some stocks may have low ASR, but other stocks may have high ASR. However, we posit that it is the net value that should matter. Specifically, the excess return of stocks that have both high AHF and high ASR is 0.81%, while it is 0.71% for stocks that have both low AHF and low ASR. Their corresponding alphas are both very close to zero, exactly as one would expect if it is the net arbitrage trading that really matters. In sharp contrast, the excess returns are 1.18% for stocks that have high AHF and low ASR, and 0.40% for stocks that have high ASR and low AHF. Therefore, the double-sort results provide strong support that the net arbitrage trading is the driving force of the predictability for future stock returns.

To summarize, some arbitrage capital buys a stock for some reason, and at the same time, while other arbitrage capital sells short the same stock for other reasons. Therefore, it would be incomplete to rely on only one side of the arbitrage capital to infer about arbitrageurs' views on future stock returns, and thus it is crucial to consider both hedge fund holdings (the long-side) and short interest (the short-side).

3.2 Cross-Sectional Regressions

Although intuitive, as discussed in Fama and French (2008), it is difficult for the portfolio approach to identify which variable has unique information in predicting future excess returns, because the portfolio approach can be contaminated by choices of percentiles in the breakpoints and the order of sorting variables and other factors. Here, we conduct the Fama and MacBeth (1973) cross-sectional regressions to further investigate the role of AHF, ASR and AHFSR in predicting stock returns.

Our sample is quarterly from 1990 to 2012 at the firm level. The Fama and MacBeth (1973) procedure has two stages. In the first stage, at the end of each quarter, we run a cross-sectional regression of stock's excess returns over next quarter on the end-of-quarter AHF, ASR, or AHFSR, along with control variables. All explanatory variables are winsorized at the 1% and 99% levels, and standardized at the end of each quarter. The control variables include the book-to-market ratio of Fama and French (2008), the gross profitability of Novy-Marx (2013), operating profit of Fama and French (2015), momentum of Jegadeesh and Titman (1993), market capitalization of Fama and French (2008), asset growth of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), investment growth of Xing (2008), net stock issue of Fama and French (2008), accrual of Fama and French (2008) and net operating assets of Hirshleifer et al. (2004). Monthly stock returns are obtained from the CRSP and compounded into quarterly returns. Annual accounting data used for the calculation of control variables are from the COMPUSTAT. These characteristics of each firm from the third quarter of year *t* to the second quarter of year *t+1* are based on its accounting information of the last fiscal year that ends in calendar year *t-1*. In the second stage, we average the regression coefficient estimates over quarters and compute their t-values based on Newey and West (1987) standard errors with four lags.

Table 3, the left-most panel presents estimation results from using the base sample. The regression coefficients on AHF, ASR, and AHFSR are all significant and have expected signs, even after controlling for other variables shown to predict stock returns as well. The coefficient on AHF is 0.0045 with a t-value 5.39. The coefficient on ASR is -0.0044 and its t-value is -4.10. The coefficient on AHFSR is 0.0066 and its t-value is 5.82. As shown in Table 1, the average cross-sectional standard deviation of AHFSR is 0.0266, this implies that if AHFSR increase by 0.0266, which is a one standard deviation increase, excess return in the next quarter would rise by 0.66%. Meanwhile, a one standard deviation increase in AHF is associated with 0.45% increase in excess return in the next quarter, and a one standard deviation increase in ASR is associated with 0.44% decrease in excess return in the next quarter. Therefore, combing information in AHF and ASR leads to greater forecasting power for future stock returns.

Next, we restrict our sample to including only stocks that have positive hedge fund holdings and short interest, and break our base sample into two equal sub-samples and repeat the Fama-Macbeth regressions. These similar results are presented in the second, third and right-most blocks of Table 3, respectively. As can be seen, our results hold in these alternative sensitivity tests.

A number of control variables are included in the Fama and MacBeth regressions. Overall, regression coefficients on the control variables have correct signs, but many of these control variables are statistically insignificant. Apart from the NYSE size filter and $5 price filter we apply, a possible explanation is that these anomalies compete with each other and render each other insignificance. For example, AG competes with IK, and OP competes with GP. It is not surprising that each of these variables by themselves can be significant. The other possible explanation is the sample period we use. In our sample from 1990 to 2012, the value spread from Kenneth French is indeed small at 0.25% per month. During the period, momentum trading strategy suffers from a crash in the first half of 2009. In fact, momentum is highly significant in the first half of our sample (the third block) but insignificant in the second half of our sample (the fourth block). Combined, momentum is insignificant in predicting future excess returns in our sample. Intriguingly, net operating asset is significant in our base sample. Nevertheless, further check reveals that it is only significant in the first half of our base sample.

In sum, performing the Fama and MacBeth (1973) regressions, we show that net trading activity of arbitrageurs as proxied by AHFSR strongly forecasts future stock returns compared to what either the long-side AHF or the short-side ASR predicts. The predictability of AHFSR is over and above that of many other firm-level variables that can potentially forecast stock returns as well.

4. Arbitrage Trading and Asset Pricing Anomalies

If AHFSR measures net arbitrage trading activity, we could use it to reveal more insights on how arbitrageurs trade on well-known return anomalies in the cross-section. As detailed in Section 2.3, we examine a total of 10 anomalies, namely the book-to-market ratio (BM), the gross profitability (GP), operating profit (OP), momentum (MOM), market capitalization (MC), asset growth (AG), investment-to-capital ratio (IK), net stock issue (NS), accrual (AC), and net operating assets (NOA).

Table 4A Panel A verifies that the long-minus-short future return spreads averaged across these 10 anomalies are positive and significant in our sample. The monthly return spreads are 0.28% (t-value = 3.47), 0.25% (t-value = 3.19), 0.20% (t-value = 2.48), and 0.15% (t-value = 1.97) per month during the first, second, third, and fourth quarters, respectively. The magnitude is somewhat smaller compared to previous studies since we use quintile sorts instead of the more common decile sorts and we exclude small stocks from our main sample. Sample period is also likely to play a role. For example, it is well known that the size premium has largely disappeared after the 1980s. The value spread is only 0.18% per month in our sample, and the value spread from Kenneth French’s website is also small at 0.25% per month in the same period and the difference is driven by small stocks that are excluded from our sample. Our sample period also includes the 2009 momentum crash, explaining a smaller momentum profit. Not surprisingly, when we control for return factors constructed based on some of the anomalies, the resulting average five-factor alphas become smaller. They are 0.14%, 0.13%, 0.11%, and 0.08% per month during the first, second, third, and fourth quarters, respectively. The average alphas are still significant during the first two quarters after portfolio formation as shown in Panel A of Table 4B. In addition, consistent with the findings in Stambaugh, Yu, and Yuan (2012), most of the anomaly alphas are coming from the short leg since overpricing is harder for to arbitrage away due to the short-sale constraints.

We then identify stocks in the long- and short-anomaly portfolios that are also traded by the arbitrageurs in the same direction. We classify an anomaly stock to be traded by the arbitrageurs if it is in the long portfolio and recently bought by arbitrageurs (its AHFSR belongs to the top 30%), or it is in the short portfolio and recently sold short (its AHFSR belongs to the bottom 30%). Table 4C shows that these stocks account for only about 30% of both the long- and the short-portfolio. Strikingly, the anomaly returns are completely driven by these stocks that are traded by the arbitrageurs. As shown in Panel B of Table 4A, this subset of anomaly stocks earn return spreads of 0.88% (t-value = 7.10), 0.61% (t-value = 4.88), 0.34% (t-value = 2.68), and 0.27% (t-value = 2.18) per month during the first, second, third, and fourth quarters, respectively. The corresponding five-factor alphas are 0.70% (t-value = 6.31), 0.45% (t-value = 3.90), 0.25% (t-value = 1.98), and 0.22% (t-value = 1.73). Hence, the alpha shows a quick decline over time during the first year. When we examine the alphas on the long- and short-legs separately, we again find the alphas to mostly come from the short-leg. While the alpha on the long-leg is small and significant only in the first quarter, the alpha more than doubles on the short-leg and persists for a longer time.

In sharp contrast, the other 70% of anomaly stocks that are not traded by arbitrageurs do not earn any significant return spreads or alphas in any of the next four quarters as reported in Panel C of Table 4A. This is true for both the long- and the short-legs. The fact that future abnormal returns only appear among anomaly stocks traded by the arbitrageurs and these abnormal returns decline quickly during the first year suggests that arbitrageurs are effective in eliminating mispricing.[8] A close examination of Tables 4A and 4B confirms that our findings are not driven by one or two anomalies. Instead, they appear consistently and uniformly across each of the 10 anomalies examined.

So far, our findings suggest that anomaly stocks are not created equal. The anomaly stocks traded by arbitrageurs seem to be more mispriced and arbitrageurs' trading eliminates the mispricing quickly. A natural question follows: How do anomaly stocks traded by arbitrageurs differ from those that are not traded? We compare these two subsets of anomaly stocks by examining their stock price, idiosyncratic volatility, and the Amihud (2002) illiquidity measure at the portfolio level. The Amihud measure is transformed into percentiles among NYSE/AMEX or NASDAQ firms separately.

Table 5 reports results of the comparisons. Across almost all the anomalies and for both the long- and the short-portfolios, anomaly stocks that are traded have significantly lower prices, have significantly higher idiosyncratic volatilities, and are significantly more illiquid according to the Amihud measure. Pontiff (1996) and Shleifer and Vishny (1997) argue that idiosyncratic volatility is a major arbitrage cost. In addition, it is well known that hedge funds often hold illiquidity assets (e.g., Getmansky, Lo, and Marakov (2004)). The evidence is consistent with a notion that the anomaly stocks are harder to arbitrage, explaining why they are mispriced to start with.

Taken together, we find that net arbitrage activity contains useful prospective information about stock returns. Furthermore, anomaly stocks have large arbitrage costs, and their anomalous

[8] Akbas, Armstrong, Sorescu, and Subrahmanyam (2014) find that aggregate money flow into the hedge fund industry attenuates stock return anomalies.

returns are significantly associated with net arbitrage activity. These findings confirm that mispricing arises from limits-to-arbitrage and arbitrage trading is informative about mispricing.

5. Conclusion

Arbitrageurs play a crucial role in finance, but measuring their activities has been a challenge empirically. By merging a comprehensive dataset on hedge funds' quarterly holdings with data on short interest, we track arbitrage trading on both the long- and the short-sides. Over time, aggregate hedge fund holdings track aggregate short interest well and both experienced fast growth since the early 1990s. In the cross section, the difference between abnormal hedge fund holding and abnormal short interest, which captures the net arbitrage trading activities on a stock, strongly predicts its next-quarter return. When examining a broad set of asset pricing anomalies, we find anomaly returns to come exclusively from about 30% of the anomaly stocks that are traded by arbitrageurs. These stocks are also harder to arbitrage on average. Overall, our findings confirm that mispricing arises from limits-to-arbitrage and arbitrage trading is informative about mispricing.

Our simple measure of arbitrage trading can be applied in many other applications. For example, one could relate the arbitrage trading on an anomaly to its future performance. One could examine the return spread between stocks with high- and low-AHFSR as a pricing factor in the spirit of the original Arbitrage Pricing Theory. We leave these and other interesting applications for future research.

Appendix: Details of the Construction of Anomalies.

This appendix provides details on the constructions of the 10 anomalies we examine in the paper. Following the timing convention in Fama and French (2008), Novy-Marx (2013) and Hou, Xue, and Zhang (2014), the financial/accounting ratios for each stock from July of year *t* to June of year *t+1* (third quarter of year *t* to second quarter of year *t+1*) is its financial ratios for the last fiscal year ending in calendar year *t-1*. At the end of each quarter, we sort all stocks into quintiles based on stocks' financial ratios. Monthly excess returns in the next three months are calculated as equal-weighted averages of excess returns of individual firms in each portfolio. The portfolio is rebalanced at every March, June, September and December.

1. Book-to-market ratio (BM). Book equity is stockholders' book equity, plus balance sheet deferred taxes (Compustat item ITCB) and investment tax credit (TXDB) if available, minus the book value of preferred stock. We employ tiered definitions largely consistent with those used by Davis, Fama, and French (2000), Novy-marx (2013) and Hou, Xue, and Zhang (2014) to construct stockholders' equity and book value of preferred stock. Stockholders equity is as given in Compustat (SEQ) if available, or else common equity (CEQ) plus the book value of preferred stock, or else total assets minus total liabilities (AT - LT). Book value of preferred stock is redemption value (PSTKRV) if available, or else liquidating value (PSTKL) if available, or else par value (PSTK). Book-to-market ratio at year *t-1* is computed as book equity for the fiscal year ending in calendar year $t-1$ divided by the market capitalization at the end of December of $t-1$. Stocks with missing book values or negative book-values are deleted.
2. Gross Profit to Asset (GP). Following Novy-Marx (2013), we measure gross profits-to-assets at year *t-1* as gross profit at year *t-1* (Compustat item GP) divided by total assets at year *t-1* (AT).

3. Operating Profit (OP). Following Fama and French (2015), we measure operating profit at year *t-1* as year *t-1* gross profit (Compustat item GP), minus selling, general, and administrative expenses (XSGA) if available, minus interest expense (XINT) if available, all divided by year *t-1* book equity. Stocks with missing book value or negative book-value are deleted.
4. Momentum (MOM). Following Jegadeesh and Titman (1993). At the end of March, June, September and December (time *t*), we compute each stock's cumulative return from month *t-13* to *t-2*, and form quintile portfolios for the next three months. We compute equal-weighted monthly returns in each portfolio for month *t+1* to *t*+3, and the portfolio is rebalanced at the end of month *t+3*.
5. Market Capitalization (ME). Following Fama and French (2008), ME is defined as the market capitalization at the end of June in each year. It is the product of number of shares and shares outstanding from the CRSP. This ME covers the following four quarters.
6. Asset Growth (AG). Following Cooper, Gulen and Schill (2008), we compute asset growth at year *t-1* as total assets (AT) for the fiscal year ending in calendar year *t-1* divided by total assets for the fiscal year ending in calendar year *t-2*, minus one.
7. Investment growth (IK). Following Xing (2008), we measure investment growth for year *t-1* as the growth rate in capital expenditure (CAPX) from the fiscal year ending in calendar year *t−2* to the fiscal year ending in *t−1*.
8. Net Stock Issue (NS). Following Fama and French (2008), we compute net stock issue at year *t-1*, as the split-adjusted shares outstanding for fiscal year ending in calendar year *t-1* divided by the split-adjusted shares outstanding for fiscal year ending in calendar year *t-2*, minus one. The split-adjusted shares outstanding are calculated as shares outstanding (CSHO) times the adjustment factor (AJEX).
9. Accrual (AC). Accruals at year *t-1* are defined following Fama and French (2008), as the change in operating working capital per split-adjusted share from *t-2* to *t-1* divided by book equity per split-adjusted share at *t-1*. Operating working capital is computed as current

assets (ACT) minus cash and short-term investments (CHE), minus, the difference of current liability (LCT) and debt in current liabilities (DLC) if available.

10. Net Operating Assets (NOA). Following Hirshleifer et al. (2004), we define net operating assets (NOA) at year *t-1*, as operating assets minus operating liabilities at year *t-1* scaled by total assets at year *t-2* (Compustat item AT). Operating assets are total assets (AT) minus cash and short-term investment (CHE). Operating liabilities are total assets minus debt included in current liabilities (item DLC, zero if missing), minus long-term debt (item DLTT, zero if missing), minus minority interests (item MIB, zero if missing), minus book value of preferred stocks as described in the definition of book equity (zero if missing), and minus common equity (CEQ).

References

Agarwal, Vikas, Wei Jiang, Yuehua Tang, and Baozhong Yang, 2013, Uncovering hedge fund skill from the portfolio holdings they hide, *Journal of Finance* 59, 1271–1289.

Akbas, Ferhat, William Armstrong, Sorin Sorescu, and Avanidhar Subrahmanyam, 2014, Smart money, dumb money, and equity return anomalies, Working paper, Texas A&M University.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross section and time series effects, *Journal of Financial Markets* 5, 31–56.

Asquith, P., and L. Meulbroek, 1995. An empirical investigation of short interest, Working paper, M.I.T.

Asquith, Paul, Parag Pathak, and Jay Ritter, 2005, Short Interest, Institutional Ownership, and Stock Returns, *Journal of Financial Economics* 78, 243-276.

Boehmer, Ekkehart, Charles Jones, and Xiaoyan Zhang, 2008, Which Shorts are Informed? *Journal of Finance* 63, 491-527.

Brown, Stephen, and Christopher Schwarz, 2013, Does market participants care about portfolio disclosure? Evidence from hedge fund 13F filings, Working paper, NYU and UC Irvine.

Brunnermeier, Markus, and Stefan Nagel, 2004, Hedge funds and the technology bubble, *Journal of Finance* 59, 2013-2040.

Brunnermeier, Markus, and Lasse Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

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

Cao, Charles, Yong Chen, William Goetzmann, and Bing Liang, 2014, The role of hedge fund in the security price formation process, Working paper.

Cooper, M., H. Gulen, and M. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–52.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389-406.

Desai, H., K. Ramesh, S. R. Thiagarajan, and B. V. Balachandran, 2002, An Investigation of the Informational Role of Short Interest in the Nasdaq Market, *Journal of Finance* 57, 2263-2287.

Diamond, Douglas, and Robert Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277-311.

Drechsler, Itamar, and Qingyi Drechsler, 2014, Theshorting premium and asset pricing anomalies, Working paper, NYU.

Fama, Eugene, and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–87.

Fama, Eugene, and Kenneth French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Fama, Eugene, and Kenneth French, 2014, A five-factor asset pricing model, *Journal of Financial Economics*, forthcoming.

Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium–empirical tests, *Journal of Political Economy* 81, 607–636.

Getmansky, M., Lo, A. W., Makarov, I., 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529–609.

Griffin, John, Jeffrey Harris, Tao Shu, and Selim Topaloglu, 2011, Who drove and burst the tech bubble?, *Journal of Finance* 66, 1251-1290.

Griffin, John, 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.

Hanson, Samuel, and Adi Sunderam, 2014, The Growth and Limits of Arbitrage: Evidence from Short Interest, *Review of Financial Studies* 27, 1238-1286.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets, *Journal of Accounting and Economics* 38, 297-331.

Hou, Kewei, Chen Xue, and Lu Zhang, 2014, Digesting anomalies: An investment approach, *Review of Financial Studies*, forthcoming.

Hwang, Byoung-Hyoun, and Baixiao Liu, 2014, Short sellers' trading anomalies, Working paper, Cornell University.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jiao, Y., M. Massa, and H. Zhang, 2015, Short selling meets hedge fund 13F: An anatomy of informed demand, Working paper, INSEAD.

Lintner, John, 1965, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics* 47, 13–37.

Lou, Dong, and Christopher Polk, 2014, Comomentum: Inferring Arbitrage Activity from Return Correlations, Working paper, London School of Economics.

Loughran, T., and J. R. Ritter, 1995, The New Issues Puzzle, *Journal of Finance* 50, 23–51.

Miller, E., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151–1168.

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

Newey, W.K., and K. D. West, 1987, A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium. *Journal of Financial Economics* 108, 1-28.

Pontiff, Jeffrey, 1996, Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* 111, 1135-1151.

Reca, Blerina B., Richard W. Sias, and Harry J. Turtle, 2015, Hedge fund crowds and mispricing, *Management Science*, forthcoming.

Ritter, Jay, 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3-27.

Ross, Stephen, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shleifer, Andrei, and Robert Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

Sloan, R., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289-315.

Stambaugh, Robert, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

Wu, Julie, and Andrew Zhang, 2014, Have short sellers become more sophisticated? Evidence from market anomalies, Working paper, University of Georgia.

Xing, Yuhang, 2008, Interpreting the value effect through the *Q*-theory: An empirical investigation, *Review of Financial Studies* 21, 1767-1795.

Table 1. Summary Statistics

This table presents summary statistics for the following variables: hedge fund holdings (HF), defined as the ratio between shares owned by hedge funds and the number of outstanding shares; short ratio (SR), defined as the ratio between shares shorted and the number of shares outstanding; the difference between HF and SR (HFSR), abnormal hedge fund holdings (AHF), defined as the percentage change of current HF from the average of HF in the previous four quarters; abnormal short ratio (ASR), defined as the percentage change of current SR from the average of SR in the previous four quarters; the difference between AHF and ASR (AHFSR). The summary statistics include the mean, 5 percentile, 25 percentile, median, 75 percentile, 95 percentile, and standard deviations (P5, P25, P50, P75, P95, STD). Monthly stock returns are from the CRSP. Annual accounting data used for the calculation of characteristics are from the COMPUSTAT. The characteristics of each firm from the July of year t to the June of year t+1 are based on its accounting information of the last fiscal year that ends in calendar year t-1. At the end of each quarter, we delete firms whose market capitalizations are below the 20 percentile market capitalization of NYSE firms, which are available at the website of Kenneth French. In panel A, at the end of each quarter, we first compute the above statistics across firms, and then take average across quarters. % of CRSP represents the total market capitalization of our sample as a portion of the market capitalization of the full CRSP universe. In panel B, at the end of each quarter, we compute the correlations between HF, SR, HFSR, AHF, ASR, AHFSR and firm characteristics, and present average correlations over quarters. We consider the following characteristics: the book-to-market ratio (BM) of Fama and French (2008), the gross profitability (GP) of Novy-Marx (2013), operating profit (OP) of Fama and French (2015), momentum (MOM) of Jegadeesh and Titman (1993), market capitalization (MC) of Fama and French (2008), asset growth (AG) of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), investment-to-capital ratio (IK) of Xing (2008), net stock issue (NS) of Fama and French (2008), accrual (AC) of Fama and French (2008) and net operating assets (NOA) of Hirshleifer et al. (2004).

Panel A: Summary							
Var.	Mean	P5	P25	P50	P75	P95	STD
HF	0.0372	0.0034	0.0112	0.0237	0.0490	0.1201	0.0397
SR	0.0349	0.0038	0.0112	0.0235	0.0444	0.1120	0.0366
HFSR	0.0023	-0.0742	-0.0163	0.0013	0.0188	0.0823	0.0469
AHF	0.0019	-0.0267	-0.0055	0.0004	0.0075	0.0363	0.0190
ASR	0.0018	-0.0237	-0.0052	0.0002	0.0069	0.0338	0.0183
AHFSR	0.0000	-0.0450	-0.0103	0.0002	0.0106	0.0445	0.0266
% of CRSP	0.8673	0.7487	0.8550	0.8729	0.8926	0.9222	0.0450

Panel B: Correlation																
	HF	SR	HFSR	AHF	ASR	AHFSR	BM	GP	OP	MOM	MC	AG	IK	NS	AC	NOA
HF	1.00	0.23	0.63	0.42	0.06	0.24	0.00	0.06	-0.04	0.06	-0.17	0.08	0.06	0.08	-0.01	0.06
SR	0.23	1.00	-0.59	0.04	0.45	-0.30	-0.09	0.08	-0.08	-0.04	-0.20	0.19	0.11	0.11	0.05	0.10
HFSR	0.63	-0.59	1.00	0.30	-0.31	0.45	0.08	-0.02	0.03	0.07	0.00	-0.09	-0.04	-0.02	-0.05	-0.03
AHF	0.42	0.04	0.30	1.00	0.07	0.64	0.01	0.01	-0.01	0.02	-0.03	-0.02	0.00	-0.02	-0.01	-0.01
ASR	0.06	0.45	-0.31	0.07	1.00	-0.68	-0.03	0.01	-0.02	-0.02	-0.05	0.05	0.03	0.02	0.02	0.02
AHFSR	0.24	-0.30	0.45	0.64	-0.68	1.00	0.03	0.00	0.01	0.02	0.01	-0.04	-0.03	-0.03	-0.02	-0.02

Table 2. Returns and Alphas of Portfolios Formed on Arbitrage Capital

At the end of each quarter, we form quintile portfolios based on AHF, ASR, or AHFSR, and track each portfolio's monthly excess returns in the next quarter, which are the equal-weighted average of excess returns of firms in each portfolio. Quintile 5 has the highest AHF, ASR or AHFSR. We adjust risk exposure using the three factors of Fama and French (1996), the Fama-French three factor and the Carhart (1997) momentum factor, and the five factors of Fama and French (2015), and they are labelled as FF3, FF4 and FF5 respectively. Panel A presents results from portfolios formed on AHF, panel B presents results from portfolios formed on ASR, panel C presents results from portfolios formed on AHFSR, panel D presents excess returns of these portfolios in the future quarters that span the next four quarters. The left panel presents excess returns and alphas, and the right panel presents their t-values. Table 2A uses our base sample from 1990/01 to 2012/12, which deletes the firms whose market capitalizations are below the NYSE 20 percentile size breakpoints. Table 2B presents results from using a sample before applying the 20 percentile NYSE size breakpoints. Table 2C uses firms that have non-zero HF and SR in our base sample. Table 2D uses the first half of our sample. Table 2E uses the second half of our sample. Table 2F presents results from tercile portfolios independently formed on AHF and ASR in our base sample. Excess Returns are in percentages. Factors data are obtain from the webpage of Kenneth French.

Table 2A. Base Sample								
	Return and Alpha				t-value			
	Ret.	FF3	FF4	FF5	Ret.	FF3	FF4	FF5
Panel A: Quintile Portfolios Formed on AHF								
AHF1	0.61	-0.12	-0.04	-0.16	1.77	-1.23	-0.41	-1.65
AHF2	0.57	-0.12	-0.05	-0.23	1.95	-1.53	-0.61	-3.03
AHF3	0.62	-0.02	0.05	-0.19	2.36	-0.26	0.66	-2.51
AHF4	0.81	0.11	0.16	-0.04	2.76	1.43	2.01	-0.47
AHF5	1.05	0.28	0.34	0.19	2.97	3.11	3.92	2.12
AHF-HML	0.44	0.40	0.38	0.35	4.98	4.50	4.27	3.80
Panel B: Quintile Portfolios Formed on ASR								
ASR1	0.90	0.18	0.24	0.09	2.70	2.08	2.91	1.05
ASR2	0.86	0.18	0.21	0.01	3.05	2.38	2.68	0.10
ASR3	0.75	0.09	0.13	-0.05	2.79	1.19	1.73	-0.65
ASR4	0.68	0.00	0.07	-0.11	2.32	0.05	0.77	-1.27
ASR5	0.49	-0.32	-0.17	-0.35	1.32	-3.09	-1.91	-3.31
ASR-HML	-0.41	-0.50	-0.42	-0.44	-4.21	-5.42	-4.71	-4.58
Panel C: Quintile Portfolios Formed on AHFSR								
AHFSR1	0.43	-0.34	-0.22	-0.38	1.20	-3.53	-2.48	-3.80
AHFSR2	0.61	-0.08	-0.01	-0.20	2.07	-0.98	-0.06	-2.40
AHFSR3	0.69	0.06	0.10	-0.09	2.65	0.74	1.40	-1.21
AHFSR4	0.81	0.14	0.17	-0.04	2.85	1.75	2.07	-0.48
AHFSR5	1.11	0.36	0.42	0.27	3.23	4.26	5.02	3.23
AHFSR-HML	0.68	0.70	0.64	0.65	7.93	8.17	7.57	7.20
Panel D: Subsequent-quarter Returns and t-values after Portfolio Formation								
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AHF-HML	0.44	0.22	0.17	-0.04	4.98	2.77	2.14	-0.56
ASR-HML	-0.41	-0.26	-0.12	-0.11	-4.21	-2.57	-1.15	-1.21
AHFSR-HML	0.68	0.42	0.18	0.01	7.93	4.90	1.90	0.18

Table 2B. Full Sample								
	Return and Alpha				t-value			
	Ret.	FF3	FF4	FF5	Ret.	FF3	FF4	FF5
Panel A: Quintile Portfolios Formed on AHF								
AHF1	0.57	-0.17	-0.04	-0.20	1.66	-1.84	-0.46	-2.14
AHF2	0.61	-0.08	0.02	-0.18	2.13	-1.09	0.23	-2.48
AHF3	0.66	0.07	0.14	-0.05	2.79	0.85	1.74	-0.64
AHF4	0.80	0.13	0.19	0.00	2.90	1.82	2.58	-0.01
AHF5	0.96	0.21	0.31	0.12	2.87	2.56	4.01	1.44
AHF-HML	0.39	0.38	0.35	0.32	5.34	5.19	4.68	4.18
Panel B: Quintile Portfolios Formed on ASR								
ASR1	0.80	0.04	0.18	-0.01	2.32	0.41	2.27	-0.15
ASR2	0.83	0.18	0.23	0.03	3.13	2.27	2.94	0.47
ASR3	0.84	0.26	0.32	0.14	3.63	3.40	4.23	1.83
ASR4	0.69	0.04	0.10	-0.08	2.50	0.51	1.15	-0.99
ASR5	0.43	-0.36	-0.21	-0.38	1.21	-3.76	-2.60	-3.91
ASR-HML	-0.36	-0.40	-0.39	-0.37	-4.58	-5.02	-4.83	-4.39
Panel C: Quintile Portfolios Formed on AHFSR								
AHFSR1	0.41	-0.37	-0.21	-0.39	1.17	-3.89	-2.74	-4.07
AHFSR2	0.62	-0.05	0.04	-0.15	2.22	-0.63	0.53	-2.06
AHFSR3	0.77	0.19	0.24	0.06	3.34	2.53	3.16	0.87
AHFSR4	0.81	0.15	0.21	0.01	2.98	2.07	3.02	0.16
AHFSR5	0.99	0.23	0.34	0.15	2.92	2.81	4.55	1.86
AHFSR-HML	0.58	0.60	0.55	0.54	7.74	7.94	7.39	6.88

Table 2C. HF>0,SR>0, Base Sample								
	Return and Alpha				t-value			
	Ret.	FF3	FF4	FF5	Ret.	FF3	FF4	FF5
Panel A: Quintile Portfolios Formed on AHF								
AHF1	0.63	-0.10	-0.02	-0.13	1.81	-0.98	-0.17	-1.37
AHF2	0.61	-0.08	0.00	-0.18	2.06	-1.02	-0.06	-2.40
AHF3	0.69	0.05	0.12	-0.13	2.58	0.63	1.52	-1.80
AHF4	0.81	0.10	0.15	-0.05	2.69	1.26	1.79	-0.65
AHF5	1.08	0.30	0.36	0.20	2.99	3.07	3.76	2.04
AHF-HML	0.45	0.40	0.38	0.33	4.76	4.23	3.99	3.40
Panel B: Quintile Portfolios Formed on ASR								
ASR1	0.96	0.24	0.30	0.15	2.85	2.61	3.37	1.65
ASR2	0.86	0.19	0.21	0.01	3.01	2.38	2.67	0.17
ASR3	0.79	0.12	0.16	-0.03	2.85	1.57	2.07	-0.48
ASR4	0.71	0.03	0.09	-0.09	2.37	0.31	1.03	-1.02
ASR5	0.50	-0.30	-0.16	-0.34	1.35	-2.81	-1.67	-3.09
ASR-HML	-0.45	-0.54	-0.46	-0.49	-4.47	-5.63	-4.96	-4.86
Panel C: Quintile Portfolios Formed on AHFSR								
AHFSR1	0.47	-0.30	-0.18	-0.34	1.30	-3.03	-1.98	-3.35
AHFSR2	0.61	-0.08	0.00	-0.20	2.04	-0.94	-0.01	-2.31
AHFSR3	0.76	0.12	0.16	-0.04	2.82	1.53	2.15	-0.54
AHFSR4	0.83	0.16	0.19	-0.02	2.87	1.90	2.25	-0.25
AHFSR5	1.15	0.39	0.44	0.30	3.26	4.31	4.90	3.30
AHFSR-HML	0.67	0.69	0.62	0.64	7.25	7.36	6.75	6.46

Table 2D. First Half of the Sample, 1990/01-2000/06								
	Return and Alpha				t-value			
	Ret.	FF3	FF4	FF5	Ret.	FF3	FF4	FF5
Panel A: Quintile Portfolios Formed on AHF								
AHF1	0.77	-0.19	-0.01	-0.18	1.67	-1.28	-0.04	-1.21
AHF2	0.69	-0.23	-0.03	-0.28	1.91	-1.79	-0.29	-2.52
AHF3	0.70	-0.14	0.03	-0.21	2.24	-1.06	0.28	-1.82
AHF4	0.94	0.01	0.15	-0.05	2.64	0.12	1.33	-0.43
AHF5	1.15	0.10	0.19	0.05	2.45	0.72	1.34	0.37
AHF-HML	0.38	0.29	0.20	0.23	3.10	2.39	1.62	1.76
Panel B: Quintile Portfolios Formed on ASR								
ASR1	1.05	0.09	0.27	0.09	2.42	0.66	2.05	0.71
ASR2	1.04	0.14	0.26	0.05	3.00	1.19	2.20	0.46
ASR3	0.83	-0.03	0.08	-0.10	2.58	-0.26	0.70	-0.94
ASR4	0.72	-0.21	-0.04	-0.23	1.94	-1.61	-0.34	-1.94
ASR5	0.63	-0.41	-0.21	-0.45	1.32	-2.62	-1.41	-2.93
ASR-HML	-0.41	-0.50	-0.47	-0.55	-3.44	-4.44	-4.03	-4.56
Panel C: Quintile Portfolios Formed on AHFSR								
AHFSR1	0.54	-0.47	-0.27	-0.48	1.14	-3.10	-1.91	-3.15
AHFSR2	0.72	-0.21	-0.01	-0.24	1.98	-1.59	-0.07	-2.04
AHFSR3	0.85	0.02	0.17	-0.06	2.73	0.19	1.54	-0.58
AHFSR4	0.94	0.02	0.14	-0.05	2.65	0.20	1.12	-0.49
AHFSR5	1.21	0.20	0.31	0.17	2.65	1.46	2.29	1.35
AHFSR-HML	0.67	0.67	0.57	0.65	5.93	5.77	4.99	5.27

Table 2E. Second Half of the Sample, 2000/07-2012/12								
	Return and Alpha				t-value			
	Ret.	FF3	FF4	FF5	Ret.	FF3	FF4	FF5
Panel A: Quintile Portfolios Formed on AHF								
AHF1	0.47	-0.14	-0.13	-0.10	0.90	-1.26	-1.17	-0.89
AHF2	0.48	-0.05	-0.05	-0.08	1.04	-0.62	-0.57	-0.88
AHF3	0.54	0.07	0.08	-0.04	1.30	1.02	1.04	-0.56
AHF4	0.69	0.15	0.15	0.08	1.48	1.96	1.89	0.98
AHF5	0.96	0.38	0.40	0.44	1.84	4.12	4.60	4.52
AHF-HML	0.50	0.52	0.53	0.54	3.99	4.22	4.29	4.27
Panel B: Quintile Portfolios Formed on ASR								
ASR1	0.75	0.21	0.21	0.20	1.50	2.36	2.36	2.18
ASR2	0.69	0.17	0.16	0.06	1.57	2.09	2.00	0.74
ASR3	0.69	0.19	0.18	0.09	1.61	2.24	2.20	1.10
ASR4	0.66	0.15	0.15	0.12	1.45	1.65	1.63	1.22
ASR5	0.35	-0.30	-0.26	-0.16	0.62	-2.38	-2.51	-1.25
ASR-HML	-0.41	-0.51	-0.46	-0.37	-2.67	-3.47	-3.65	-2.40
Panel C: Quintile Portfolios Formed on AHFSR								
AHFSR1	0.34	-0.30	-0.27	-0.22	0.62	-2.66	-2.72	-1.85
AHFSR2	0.51	-0.01	-0.01	-0.04	1.12	-0.12	-0.09	-0.51
AHFSR3	0.56	0.08	0.08	-0.01	1.35	0.97	0.92	-0.10
AHFSR4	0.69	0.19	0.18	0.09	1.56	2.56	2.48	1.17
AHFSR5	1.03	0.45	0.46	0.49	2.00	5.09	5.29	5.13
AHFSR-HML	0.69	0.75	0.73	0.70	5.38	5.85	5.93	5.18

Table 2F. Double Sorting from AHF and ASR								
	Ret. and Alpha				t-value			
	AHF1	AHF2	AHF3	AHF-HML	AHF1	AHF2	AHF3	AHF-HML
Panel A: Excess Returns								
ASR1	0.71	0.72	1.18	0.47	2.12	2.55	3.47	4.41
ASR2	0.67	0.67	0.98	0.30	2.29	2.74	3.26	3.73
ASR3	0.40	0.45	0.81	0.42	1.13	1.48	2.25	4.75
ASR-HML	-0.31	-0.27	-0.36		-3.43	-2.80	-3.73	
Panel B: FF3 Alpha								
ASR1	0.01	0.05	0.42	0.41	0.11	0.56	4.36	3.91
ASR2	-0.02	0.07	0.28	0.30	-0.22	0.86	2.88	3.60
ASR3	-0.37	-0.26	0.03	0.40	-3.51	-2.72	0.31	4.56
ASR-HML	-0.38	-0.31	-0.39		-4.35	-3.31	-4.09	
Panel C: FF4 Alpha								
ASR1	0.08	0.08	0.47	0.39	0.88	0.87	4.94	3.66
ASR2	0.02	0.12	0.30	0.28	0.27	1.52	3.10	3.33
ASR3	-0.25	-0.14	0.12	0.38	-2.57	-1.60	1.32	4.25
ASR-HML	-0.34	-0.22	-0.35		-3.86	-2.43	-3.65	
Panel D: FF5 Alpha								
ASR1	-0.04	-0.13	0.29	0.33	-0.42	-1.52	2.93	3.04
ASR2	-0.11	-0.11	0.12	0.23	-1.24	-1.46	1.27	2.64
ASR3	-0.41	-0.38	-0.02	0.39	-3.97	-3.84	-0.22	4.25
ASR-HML	-0.37	-0.25	-0.31		-3.99	-2.49	-3.12	

Table 3. Fama-MacBeth Regression of Monthly Excess Return on AHF, ASR and AHFSR

This table presents the Fama and MacBeth (1973) regression results from regressing excess returns of the next quarter on the current end-of-quarter AHF, ASR and AHFSR. Control variables include: the book-to-market ratio (BM) of Fama and French (2008), the gross profitability (GP) of Novy-Marx (2013), operating profit (OP) of Fama and French (2015), momentum (MOM) of Jegadeesh and Titman (1993), market capitalization (MC) of Fama and French (2008), asset growth (AG) of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), investment growth (IK) of Xing (2008), net stock issue (NS) of Fama and French (2008), accrual (AC) of Fama and French (2008) and net operating assets (NOA) of Hirshleifer et al. (2004). We take logs of BM and MC.

Monthly stock returns are from the CRSP and are compounded into quarterly returns. Annual accounting data used for the calculation of control variables are from the COMPUSTAT. These characteristics of each firm from the July of year t to the June of year t+1 are based on its accounting information of the last fiscal year that ends in calendar year t-1. The calculation of t-values follows Newey and West (1987) using four lags. All explanatory variables are winsorized at 1% and 99% and standardized at the end of each quarter. The sample is quarterly from 1990 to 2012.

Fama-MacBeth Regression of Future Excess Return on AHF, ASR and AHFSR												
Sample	1990/01-2012/12			1990/01-2012/12, HF>0, SR>0			1990/01-2000/06			2000/07-2012/12		
AHF	0.0045***			0.0047***			0.0036***			0.0052***		
t-value	5.39			5.46			2.81			6.31		
ASR		-0.0044***			-0.0048***			-0.0054***			-0.0036***	
t-value		-4.1			-4.23			-2.81			-3.38	
AHFSR			0.0066***			0.0071***			0.0067***			0.0065***
t-value			5.82			6.17			3.24			6.87
BM	0.0036	0.0033	0.0033	0.0036	0.0032	0.0032	0.0004	-0.0001	-0.0002	0.0065	0.0063	0.0063
t-value	1.35	1.22	1.21	1.26	1.12	1.12	0.09	-0.02	-0.04	1.59	1.54	1.55
OP	0.0025	0.0024	0.0025	0.0026	0.0023	0.0025	-0.0014	-0.0018	-0.0017	0.0061*	0.0061*	0.0062*
t-value	0.95	0.87	0.91	0.94	0.83	0.88	-0.33	-0.41	-0.4	1.75	1.71	1.74
MOM	0.0001	0.0002	0.0001	-0.0005	-0.0003	-0.0004	0.0096***	0.0102***	0.0099***	-0.0078	-0.0081	-0.008
t-value	0.01	0.04	0.03	-0.14	-0.09	-0.11	3.16	3.29	3.25	-1.48	-1.55	-1.51
MC	-0.0023	-0.0029	-0.0027	-0.003	-0.0037*	-0.0034*	-0.0007	-0.0012	-0.0009	-0.0038**	-0.0045***	-0.0045***
t-value	-1.24	-1.57	-1.45	-1.55	-1.93	-1.74	-0.21	-0.34	-0.25	-2.36	-2.9	-2.81
AG	0.0007	0.0009	0.0009	0.0011	0.0013	0.0013	0.0025	0.0027	0.0027	-0.0016	-0.0014	-0.0014
t-value	0.35	0.45	0.44	0.54	0.61	0.62	0.77	0.82	0.81	-0.63	-0.56	-0.56
IK	-0.0004	-0.0003	-0.0004	-0.0006	-0.0005	-0.0005	0.0002	0.0003	0.0003	-0.0013	-0.0012	-0.0013
t-value	-0.47	-0.36	-0.42	-0.57	-0.47	-0.53	0.2	0.31	0.31	-1.05	-0.99	-1.07
GP	0.0027	0.0027	0.0027	0.0031	0.0031*	0.003	0.0036	0.0034	0.0034	0.0015	0.0016	0.0015
t-value	1.49	1.48	1.44	1.66	1.67	1.62	1.61	1.54	1.53	0.54	0.57	0.53
NS	-0.0016	-0.0017	-0.0016	-0.0016	-0.0017	-0.0017	-0.0015	-0.0016	-0.0016	-0.0015	-0.0016	-0.0015
t-value	-1.26	-1.4	-1.35	-1.15	-1.23	-1.21	-0.86	-0.95	-0.98	-0.88	-0.97	-0.89
AC	0.0007	0.0008	0.0008	0.0009	0.001	0.001	-0.0001	0	0.0001	0.0011	0.0013	0.0012
t-value	1.02	1.17	1.14	1.34	1.55	1.48	-0.09	0.03	0.06	1.14	1.24	1.19
NOA	-0.0059**	-0.0059**	-0.0059**	-0.0060**	-0.0060**	-0.0060**	-0.0105**	-0.0104**	-0.0104**	-0.0016	-0.0017	-0.0017
t-value	-2.43	-2.46	-2.44	-2.49	-2.53	-2.5	-2.46	-2.48	-2.44	-0.83	-0.86	-0.87
Const.	0.0237***	0.0237***	0.0236***	0.0246***	0.0245***	0.0245***	0.0253***	0.0253***	0.0253***	0.0215	0.0215	0.0215
t-value	2.85	2.85	2.84	2.94	2.93	2.93	3.01	3.00	3.00	1.62	1.61	1.61
R2	0.0829	0.0833	0.0837	0.0879	0.0886	0.0888	0.0727	0.0739	0.0738	0.0904	0.0901	0.091
Obs.	115461	115461	115461	101993	101993	101993	55355	55355	55355	61512	61512	61512

Table 4. Arbitrage Trading and Anomaly Returns in the Following Year

For each anomaly variable, at the end of each quarter, we construct quintile portfolios and compute monthly portfolio returns in the next four quarters (Q1, Q2, Q3 and Q4). LMS (Panel A) is the return difference between the long leg and short leg. Meanwhile, we independently form three AHFSR portfolios using 30% and 70% AHFSR cutoff values. At the end of each quarter, in the long leg, we identify firms that belong to the AHFSR group 3 (Trading) and firms that do not belong to AHFSR group 3 (Not Trading). Similarly, in the short leg, we identify those firms that belong to the AHFSR group 1 (Trading) and those firms that do not belong to the AHFSR group 1 (Not Trading). We track the monthly equal-weighted averages of these four portfolios. Return of trading (Panel B) is defined as the return difference between the returns of the long leg and the short leg when arbitrage capital trades, and return of not trading (Panel C) is defined as the return difference between the returns of the long leg and short leg when arbitrage capital does not trade. The difference between the group of trading and not trading is in Panel D. The upper table (Table 4A) presents returns. The middle table (Table 4b) presents associated t-values. The bottom table (Table 4C) presents, for each anomaly, the total number of stocks on the long or short leg, the numbers and proportions of stocks that are traded by the arbitrage capital on the long and short leg, in our sample. Column “Avg.” represents results from a portfolio that invests equally across 10 anomalies. We also report the alpha of the long-minus-short of these composite portfolios, Alpha (LMS), from using the Fama-French (2015) five factor model, the alphas of the long composite portfolio and short portfolio, Alpha(L) and Alpha(S). We consider the following anomalies: the book-to-market ratio (BM) of Fama and French (2008), the gross profitability (GP) of Novy-Marx (2013), operating profit (OP) of Fama and French (2015), momentum (MOM) of Jegadeesh and Titman (1993), market capitalization (MC) of Fama and French (2008), asset growth (AG) of Cooper et al. (2008), Hou, Xue, and Zhang (2014), and Fama and French (2015), investment growth (IK) of Xing (2008), net stock issue (NS) of Fama and French (2008), accrual (AC) of Fama and French (2008) and net operating assets (NOA) of Hirshleifer et al. (2004). The sample is monthly stock returns from 1990 to 2012.

Table 4A: Returns

	BM	GP	OP	MOM	MC	AG	IK	NS	AC	NOA	Avg.	Alpha(LMS)	Alpha(L)	Alpha(S)
Panel A: LMS Returns														
Q1	0.18	0.32	0.35	0.17	0.20	0.35	0.22	0.46	0.16	0.38	0.28	0.14	-0.03	-0.17
Q2	0.21	0.24	0.27	0.06	0.25	0.36	0.28	0.38	0.08	0.39	0.25	0.13	-0.02	-0.15
Q3	0.17	0.22	0.19	-0.15	0.19	0.32	0.25	0.33	0.14	0.38	0.20	0.11	-0.01	-0.11
Q4	0.11	0.18	0.18	-0.22	0.20	0.16	0.22	0.24	0.13	0.29	0.15	0.08	0.02	-0.07
Panel B: Return of Trading														
Q1	0.69	1.06	1.03	0.85	0.73	1.00	0.78	0.99	0.69	1.00	0.88	0.70	0.22	-0.48
Q2	0.59	0.69	0.65	0.43	0.60	0.69	0.58	0.67	0.39	0.77	0.61	0.45	0.13	-0.32
Q3	0.21	0.50	0.49	0.02	0.15	0.42	0.33	0.57	0.22	0.52	0.34	0.25	0.09	-0.16
Q4	0.16	0.27	0.41	-0.09	0.26	0.30	0.35	0.35	0.17	0.49	0.27	0.22	0.06	-0.15
Panel C: Return of Not Trading														
Q1	-0.07	0.02	0.04	-0.20	0.04	-0.01	-0.05	0.19	-0.12	0.11	-0.01	-0.12	-0.14	-0.02
Q2	0.05	0.06	0.09	-0.15	0.15	0.19	0.12	0.24	-0.10	0.22	0.09	-0.02	-0.09	-0.07
Q3	0.16	0.12	0.07	-0.25	0.20	0.26	0.20	0.22	0.10	0.32	0.14	0.04	-0.04	-0.08
Q4	0.09	0.15	0.07	-0.30	0.21	0.07	0.16	0.20	0.11	0.21	0.10	0.02	0.00	-0.02
Panel D: Difference between Trading and Not Trading														
Q1	0.76	1.04	0.99	1.05	0.69	1.01	0.83	0.80	0.82	0.89	0.89	0.82	0.36	-0.46
Q2	0.54	0.63	0.56	0.57	0.46	0.50	0.46	0.43	0.49	0.55	0.52	0.47	0.22	-0.25
Q3	0.06	0.38	0.41	0.27	-0.05	0.16	0.13	0.35	0.11	0.21	0.20	0.21	0.13	-0.08
Q4	0.07	0.12	0.33	0.21	0.04	0.24	0.19	0.15	0.07	0.28	0.17	0.20	0.07	-0.13

Table 4B: t-values														
	BM	GP	OP	MOM	MC	AG	IK	NS	AC	NOA	Avg.	Alpha(LMS)	Alpha(L)	Alpha(S)
Panel A: LMS Returns														
Q1	0.82	1.86	2.03	0.54	0.97	2.25	1.92	2.70	1.71	2.50	3.47	2.27	-0.49	-1.92
Q2	0.99	1.37	1.44	0.20	1.30	2.31	2.32	2.19	0.86	2.53	3.19	1.99	-0.36	-1.58
Q3	0.78	1.21	1.11	-0.62	0.99	1.90	1.82	1.85	1.51	2.45	2.48	1.46	-0.07	-1.04
Q4	0.55	0.99	1.02	-0.96	0.99	0.97	1.64	1.36	1.41	1.86	1.97	1.18	0.19	-0.62
Panel B: Return of Trading														
Q1	2.68	5.12	4.11	2.51	3.09	5.22	4.63	4.47	5.10	5.13	7.10	6.31	2.99	-4.37
Q2	2.29	3.24	2.70	1.33	2.59	3.54	3.43	2.99	2.71	4.29	4.88	3.90	1.54	-2.63
Q3	0.83	2.24	2.19	0.08	0.65	2.00	1.86	2.52	1.47	2.78	2.68	1.98	0.99	-1.20
Q4	0.64	1.15	1.76	-0.34	1.04	1.49	1.96	1.51	1.17	2.78	2.18	1.73	0.60	-1.16
Panel C: Return of Not Trading														
Q1	-0.35	0.11	0.25	-0.64	0.21	-0.07	-0.44	1.23	-1.21	0.72	-0.07	-1.75	-2.36	-0.21
Q2	0.25	0.33	0.54	-0.52	0.76	1.22	0.98	1.50	-1.04	1.39	1.16	-0.30	-1.36	-0.75
Q3	0.74	0.65	0.44	-1.02	1.06	1.54	1.46	1.26	1.04	1.94	1.72	0.55	-0.57	-0.81
Q4	0.44	0.84	0.44	-1.34	1.12	0.39	1.13	1.18	1.09	1.21	1.27	0.28	-0.03	-0.22
Panel D: Difference between Trading and Not Trading														
Q1	5.19	6.48	5.76	7.25	4.41	6.98	5.11	5.19	5.81	5.57	7.85	6.92	5.52	-5.93
Q2	3.84	3.93	3.53	3.74	3.12	3.62	2.85	3.03	3.36	3.61	4.69	3.99	3.33	-3.27
Q3	0.38	2.23	2.73	1.66	-0.33	1.02	0.82	2.28	0.74	1.26	1.72	1.67	1.96	-1.00
Q4	0.47	0.65	2.13	1.33	0.29	1.45	1.14	0.90	0.44	1.66	1.44	1.54	0.98	-1.55

Table 4C: Number and Percentage of Stocks Traded by Arbitrage Capital

	BM	GP	OP	MOM	MC	AG	IK	NS	AC	NOA
# of Stocks Traded by HF, Long Leg	92	102	95	106	99	104	92	91	87	89
Portion of Stocks Traded by HF, Long Leg	0.29	0.32	0.30	0.33	0.31	0.33	0.33	0.29	0.34	0.28
# of Stocks Traded by HF, Short Leg	95	83	108	102	71	116	101	113	90	109
Portion of Stocks Traded by HF, Short Leg	0.30	0.26	0.33	0.31	0.22	0.37	0.36	0.36	0.35	0.34
# of Stocks on the Long or Short Leg	322	322	322	322	322	315	281	315	254	314

Table 5. Differences in Characteristics of Stocks Traded and Not Traded by Arbitrage Capital

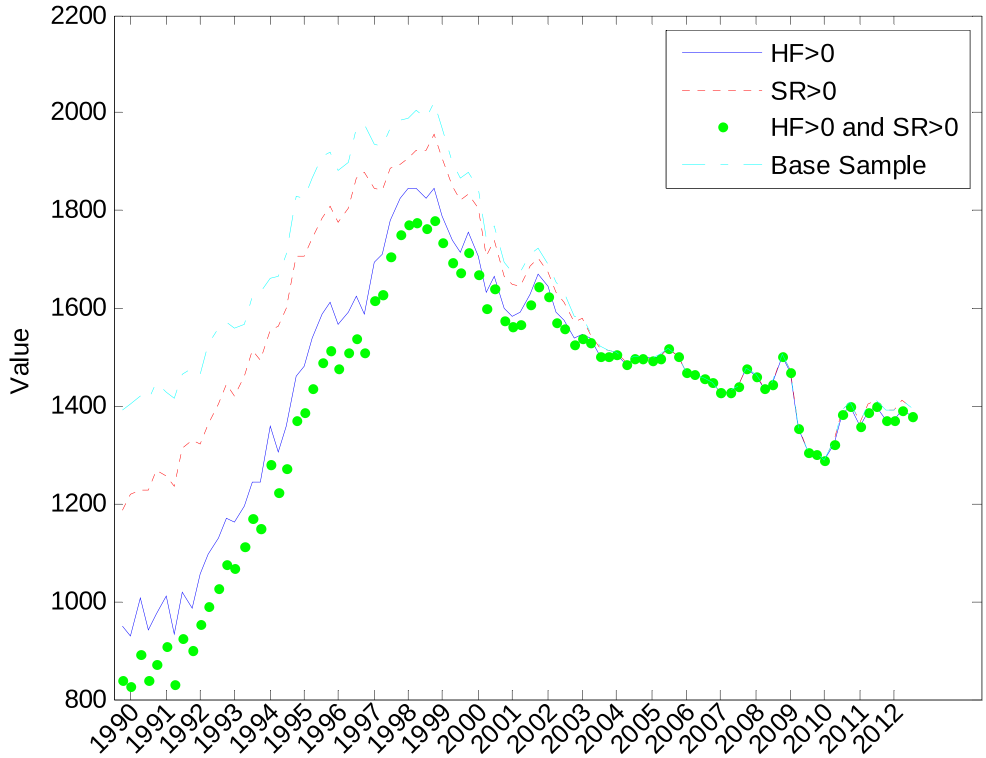
At the end of each quarter, on the long leg of each anomaly, we identify a portfolio of stocks that have high AHFSR ranking and these stocks are treated as the group of traded; we also identify a portfolio of stocks that does not have the high AHFSR rankings, and this is the group of stocks that are not traded. For these two portfolios, we compute portfolio-level price, idiosyncratic volatility, and he Amihud (2002) measure, by equal averaging across stocks in each portfolio. The Amihud measure is transformed into percentiles among NYSE/AMEX or NASDAQ firms separately. Panel A presents the difference of these characteristics (Price, IVOL, the Amihud measure) between the "Trading" portfolio and "Not Trading" portfolio for the short-leg of each anomaly, Panel B repeats this analysis for the long-leg, and panels C and D report associated t-values.

Variable	BM	GP	OP	MOM	MC	AG	IK	NS	AC	NOA	Average
Panel A: Trading - Not Trading, Short Leg											
Price	-4.6378	-3.9056	-2.8359	-0.2065	-3.1528	-3.0403	-1.9179	-2.1326	-1.1004	-1.5826	-2.4513
IVOL	0.0203	0.0157	0.0172	0.0064	0.0091	0.0113	0.0121	0.0151	0.0101	0.0106	0.0128
Amihud	34.34	23.74	21.62	20.56	39.38	27.96	23.51	25.99	20.86	22.75	26.07
Panel B: Trading - Not Trading, Long Leg											
Price	-3.4778	-1.9372	-1.2219	-0.6085	-1.6197	-1.8045	-1.9789	-2.2597	-1.6995	0.1077	-1.6500
IVOL	0.0111	0.0126	0.0132	0.0049	0.0159	0.0116	0.0111	0.0110	0.0093	0.0197	0.0120
Amihud	15.94	25.65	26.55	26.88	10.90	19.78	20.64	21.48	22.78	26.94	21.75
Panel C: Trading - Not Trading, Short Leg, t-values											
Price	-8.92	-6.72	-7.04	-0.63	-4.52	-5.48	-5.10	-5.62	-3.22	-5.53	-8.70
IVOL	19.79	11.96	18.35	8.52	12.72	14.09	16.52	19.12	13.67	18.51	20.96
Amihud	67.91	40.33	36.17	30.47	49.78	51.29	39.09	44.74	32.14	42.19	79.39
Panel D: Trading - Not Trading, Long Leg, t-values											
Price	-9.81	-5.48	-2.32	-1.27	-5.37	-4.90	-3.88	-4.84	-3.52	0.17	-6.51
IVOL	15.52	17.46	15.96	5.77	16.56	15.22	16.06	16.00	11.22	15.21	20.75
Amihud	24.79	49.55	44.19	53.27	18.13	25.97	30.08	26.48	33.81	40.69	55.96

Figure 1. Number of Stocks and Sample Coverage

At the end of each quarter, we count the number of firms that have larger than zero values of HF, larger than zero values of SR, larger than zero values of HF and SR, and the total number of firms in our base sample, and plot them over quarters in the upper figure. We compute the market capitalization of these firms as a proportion of the market capitalization of the CRSP universe, and plot them in the bottom figure. Our sample does not include firms whose market capitalization is less than 20% size percentile of the NYSE firms. The sample is quarterly from 1990 to 2012.

(a) Number of Stocks



(b) MktCap Coverage of HF>0, SR>0, HF >0 and SR>0 in Our Sample

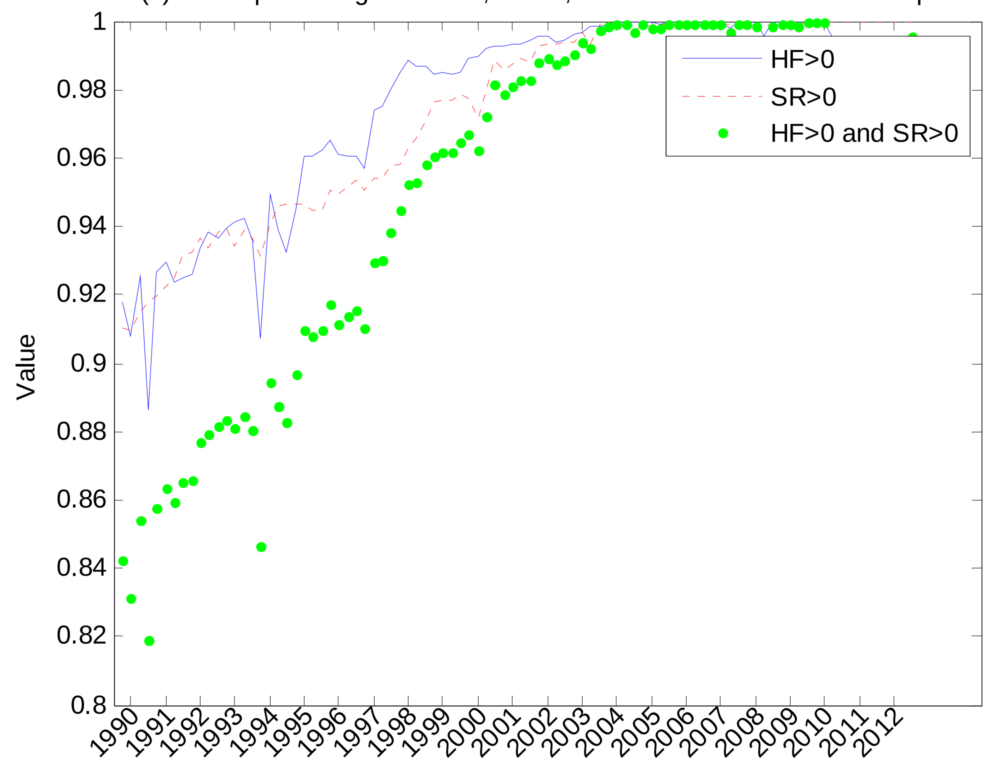
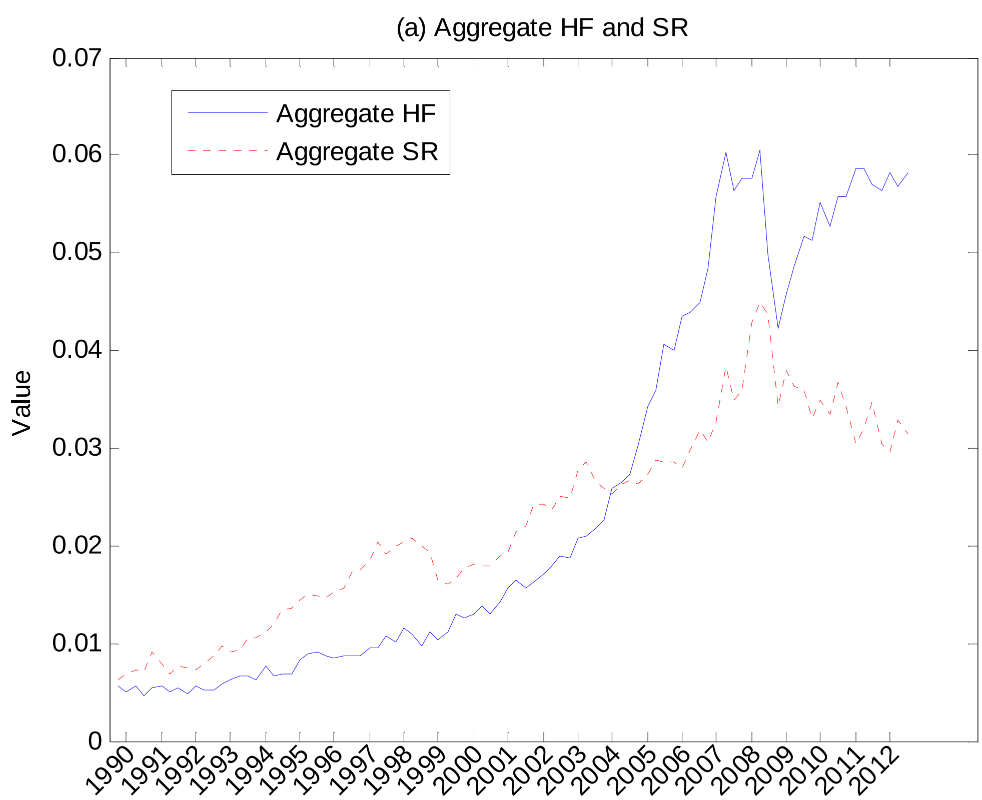
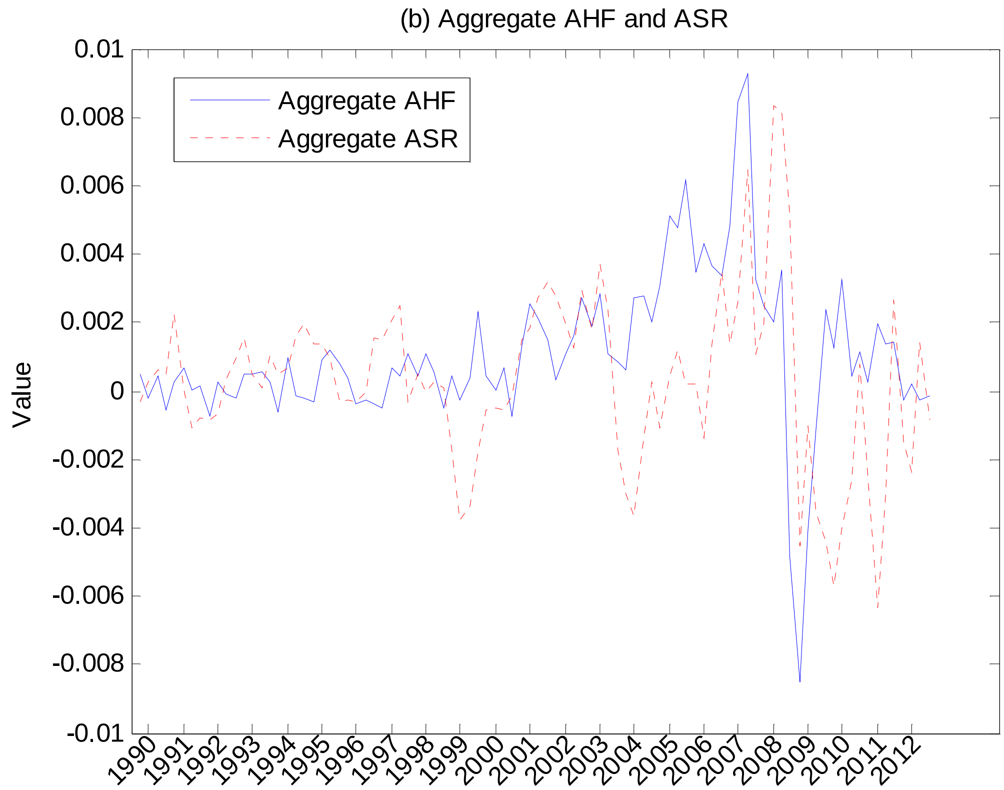
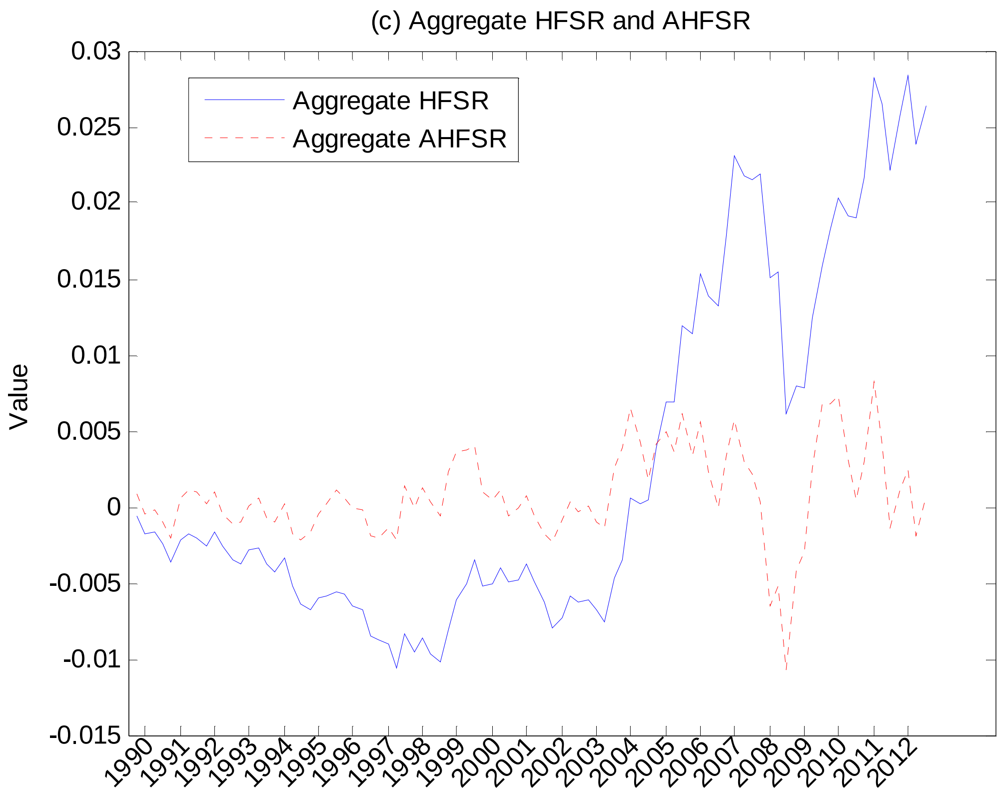


Figure 2. Aggregate Variables about Hedge Fund Holdings and Short Interest

We plot value-weighted averages of the following variables: hedge fund holdings (HF), defined as the ratio between shares owned by hedge funds and the number of outstanding shares; short ratio (SR), defined as the ratio between shares shorted and the number of shares outstanding; the difference between HF and SR (HFSR), abnormal hedge fund holdings (AHF), defined as the percentage change of current HF from the average of HF in the previous four quarters; abnormal short ratio (ASR), defined as the percentage change of current SR from the average of SR in the previous four quarters; the difference between AHF and ASR (AHFSR). The sample is quarterly from 1990 to 2012.




(b) Aggregate AHF and ASR
Aggregate AHF
Aggregate ASR
Value
0.01
0.008
0.006
0.004
0.002
0
-0.002
-0.004
-0.006
-0.008
-0.01
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012


(c) Aggregate HFSR and AHFSR
Aggregate HFSR
Aggregate AHFSR
Value
0.03
0.025
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012