Smart Money, Dumb Money, and Capital Market Anomalies

July 5, 2014

Ferhat Akbas, Will J. Armstrong, Sorin Sorescu, Avanidhar Subrahmanyam*

We provide direct evidence for the dual notions that "dumb money" exacerbates well-known stock return anomalies, and "smart money" attenuates these anomalies. We use, as measure of cross-sectional mispricing, the performance of a long-short portfolio constructed with factors that predict stock returns in the cross-section. We find that aggregate flows to mutual funds ("dumb money") appear to exacerbate cross-sectional mispricing. In contrast, aggregate flows to hedge funds ("smart money") appear to attenuate mispricing. Our results suggest that aggregate flows to mutual funds may have real adverse allocation effects in the stock market, while aggregate flows to hedge funds contribute to the correction of cross-sectional mispricing.

*Akbas is from the School of Business, University of Kansas, Lawrence, Kansas 66045-7601. Armstrong is from Rawls College of Business, Texas Tech University, Lubbock, Box 42101, Texas 79409-2101. Sorescu is from Mays Business School, Texas A&M University, Department of Finance, 360 Wehner Building, College Station, Texas 77843-4218. Subrahmanyam is from the Anderson School, University of California at Los Angeles, 110 Westwood Plaza, Los Angeles, CA 90095-1481. Corresponding author: Sorin Sorescu (ssorescu@tamu.edu), Department of Finance, Mays Business School, Texas A&M University, 4218 TAMU, College Station, Texas 77843-4218, telephone (979) 458-0380. The authors wish to thank Jason Chen, Yong Chen, Hagen Kim, Brad Gilbert from Teachers Retirement System of Texas, John Claisse from Albourne America LLC, Tom Tull from the Employee Retirement System of Texas, and Michael Mulcahy from Bridgeway for many useful comments.

Smart Money, Dumb Money, and Capital Market Anomalies

Abstract

We provide direct evidence for the dual notions that "dumb money" exacerbates well-known stock return anomalies, and "smart money" attenuates these anomalies. We use, as measure of cross-sectional mispricing, the performance of a long-short portfolio constructed with factors that predict stock returns in the cross-section. We find that aggregate flows to mutual funds ("dumb money") appear to exacerbate cross-sectional mispricing. In contrast, aggregate flows to hedge funds ("smart money") appear to attenuate mispricing. Our results suggest that aggregate flows to mutual funds may have real adverse allocation effects in the stock market, while aggregate flows to hedge funds contribute to the correction of cross-sectional mispricing.

In the popular press and in academia, financial market price movements are often justified by alluding to the terms "dumb money" and "smart money."[1] Price pressure from the former class generally is presupposed to make prices depart from fundamentals, whereas arbitrage by the latter class makes prices converge to fundamental values (Frazzini and Lamont, 2008). There is extensive documentation of stock market anomalies (McLean and Pontiff, 2013, Stambaugh, Yu, and Yuan, 2012, 2013), which suggests that prices may indeed depart from fundamentals for periods of time, and the persistence of such anomalies indicates that smart money is not fully able to erase these anomalies. Even though these notions prevail in financial thought, there is as yet no direct documentation of the role of smart and dumb money in causing or correcting anomalies. In this paper, we provide clear evidence for the notion that dumb money exacerbates stock market anomalies and smart money attenuates them. We use mutual fund flows as a proxy for dumb money (Lou, 2012) and hedge fund flows as a proxy for smart money (Jagannathan, Malakhov, and Novikov, 2010).

Flows to mutual funds have been shown to create distortions in capital allocation across US stocks. Retail investors appear to contribute to these distortions in two ways. First, they tend to "chase performance" by directing money to mutual funds with strong recent performance, while failing to redeem capital from funds with poor recent performance (Sirri and Tufano, 1998). Second, they tend to direct money – "dumb money" – to mutual funds that hold overvalued stocks (Frazzini and Lamont, 2008). When mutual fund managers receive new flows from retail investors they usually increase positions in existing stock holdings. As a result, in the cross-section of mutual funds, net money inflows are associated with higher contemporaneous stock returns and subsequent return reversal (Coval and Stafford, 2007).

[1] See, for example, "The Smart Way to Follow Dumb Money," by S. Jakab, available at: http://online.wsj.com/news/articles/SB10001424052702304543904577396361227824738.

Taken together, these studies imply that money flows to mutual funds could have a real allocation impact at the aggregate stock market level because they exert the “wrong” type of price pressure on stocks that are already mispriced – the type that exacerbates cross-sectional mispricing. This could explain the persistence through time of cross-sectional predictability in US stock returns, in spite of significant arbitrage trading strategies carried out by quant-oriented hedge funds over the past two decades. To our knowledge, this important implication has not been tested in the literature and the real allocation impact of mutual fund flows remains an open question. We study the effects of aggregate mutual fund flows on the cross-sectional mispricing of US stocks by examining the inter-temporal relation between two time-series: the aggregate mutual fund flows and the aggregate level of monthly cross-sectional mispricing.

We use, as a proxy for cross-sectional mispricing, the metric proposed by Stambaugh, Yu, and Yuan (2012, 2013). Specifically, we identify each month, a group of stocks most likely to be overvalued and another group most likely to be undervalued based on eleven characteristics that are known to predict the cross-section of stock return. We then compute the returns on a “hedge” long-short investment strategy that is long undervalued stocks and short overvalued stocks. The return on this hedge strategy serves as a time-variant metric of the aggregate level of cross-sectional mispricing.[2] The strategy should produce positive returns when aggregate mispricing is being corrected and cross-sectional stock prices converge toward equilibrium. By contrast, the strategy should produce negative returns during months when stock prices diverge from equilibrium and cross-sectional mispricing is exacerbated.

[2] Because our focus is to identify stocks that are the most mispriced in the cross-section, we use the Stambaugh, Yu, and Yuan (2012, 2013) measure as a proxy for cross-sectional mispricing rather than as a performance measure.

Aggregate flows to mutual funds vary through time as a result of changing investors' sentiment and aggregate fear proxied by the VIX (see, e.g. Ben-Rephael, Kandel, and Wohl, 2012, Ederington and Golubeva, 2011), or as a result of past returns to arbitrage strategies (Akbas et al., 2014). We take advantage of this inter-temporal variation to evaluate the impact of fund flows on the aggregate cross-sectional mispricing metric, itself time varying. If aggregate flows to mutual funds contribute to exacerbating cross-sectional mispricing we expect to see a negative contemporaneous relation between the two time series.

Our results strongly support this hypothesis. We find that cross-sectional mispricing increases with mutual fund flows, as evidenced by a negative relation between flows and returns to the Stambaugh-Yu-Yuan mispricing metric. This suggests that mutual fund flows, in the aggregate, are associated with either an increase in the price of overvalued stocks, a decrease in the price of undervalued stocks, or both. In subsequent tests we examine the independent effects of aggregate mutual fund flows on the long and short leg components of the mispricing metric. We find that mutual fund flows do not affect the returns of the long leg. By contrast, mutual fund flows are associated with a significant price pressure in the returns of the short leg component. Because stocks in the short leg are – by construction – stocks that are likely to be overvalued, we conclude that aggregate mutual fund flows exacerbate cross-sectional mispricing because they are invested disproportionately into stocks that are already overvalued.

If mutual fund flows are disproportionately flowing into stocks that are already overvalued, and if the resulting price pressure further exacerbates these stocks' overvaluation, we would expect these stocks to experience a price reversal following periods of high aggregate mutual fund flows. A reversal in the price of overvalued stocks represents a price convergence toward the efficient market benchmark, translating into positive returns to the long-short hedge strategy (which remains short these overvalued stocks). Thus, our hypothesis also predicts a positive relation between aggregate mutual fund flows and future returns to the mispricing metric. Our results support this prediction. Moreover, we show that this relation once again comes exclusively from overvalued stocks (or the short leg of the hedge strategy). Thus, the impact of mutual fund flows on cross-sectional mispricing appears to operate principally through the purchase of overvalued stocks rather than the sale of undervalued stocks.

We next ask if any "smart money" is present in the market. We define "smart money" as aggregate fund flows that that take long positions in undervalued stocks or short positions in overvalued stocks – the opposite of what mutual funds do. The cross-sectional mispricing that is exacerbated by mutual fund flows should create an opportunity for more sophisticated investors to enter the market and take the opposite positions. As suggested by Jagannathan, Malakov, and Novikov (2010), and by Kokkonen and Suominen (2014), hedge funds are one such group of sophisticated investors, and we expect that aggregate hedge fund flows will have the opposite effect from mutual fund flows. We evaluate this hypothesis by regressing the returns to the mispricing metric on aggregate hedge fund flows (instead of mutual fund flows). The relation is now significantly positive (instead of negative). This suggests that hedge fund flows exert the "right" type of price pressure on mispriced stocks – the type that brings price convergence toward fundamental value and corrects cross-sectional mispricing. This conclusion is corroborated by the absence of any reversal in the mispricing metric during subsequent periods.

Examining the long and short components separately, we see that hedge fund's "corrective" effect is driven primarily by overvalued stocks. That is, aggregate hedge fund flows appear to be most effective when they are used to take short positions in overvalued stocks rather than long positions in undervalued stocks. This result is consistent not only with our previous results obtained in the case of mutual funds, but also with a long-standing literature on short sales, which documents that short transactions are generally informed (Boehmer, Jones, and Zhang (2008)). Thus, our paper shows that in the aggregate, hedge fund flows act as arbitrage capital that corrects cross-sectional mispricing. In contrast, aggregate mutual fund flows seem to impede the arbitrage function and exacerbate cross-sectional mispricing.

Our paper has important implications for the market efficiency literature. A significant puzzle in the literature is the persistence of cross-sectional return predictability despite the increasingly large number of hedge fund strategies that trade on various anomalies documented in the academic literature. As these anomalies became common knowledge among sophisticated traders, we would have expected them to vanish. The limits-to-arbitrage literature provides one explanation for why the anomalies may not completely vanish (see, e.g. Shleifer and Vishny, 1997).

We propose here an explanation for the prevalence of anomalies that is complementary to the one provided by limits-to-arbitrage.[3] The effect we document in this paper occurs earlier in the sequence of events that leads to cross-sectional mispricing. We conjecture that the cross-sectional mispricing is itself constantly fueled by performance-chasing retail investor money that enters the market through the mutual funds industry and by mutual fund managers' tendency to invest these new flows into existing stock holdings (Coval and Stafford, 2007, Wermers, 2003). In other words, our paper provides an explanation for the market imperfection that fuels mispricing, while the limits-to-arbitrage literature explains why mispricing does not instantaneously vanish in a market where at least some traders are sophisticated.

[3] See McLean and Pontiff (2013) for evidence supporting the post-publication persistence of 82 characteristic-based anomalies documented in the literature.

Our analysis thus suggests that despite cross-sectional return predictability now being common knowledge among sophisticated investors, judicious hedge-fund strategies that seek to exploit this predictability should continue to earn positive alphas so long as "dumb" money continues to enter the stock market via the mutual fund industry. Indeed, an aggregate consequence of this "dumb" money is to create a market for "smart money" – hedge funds – where investors can earn alpha by merely trading against the price pressure induced by these aggregate "dumb" flows. Our results also indicate that a net transfer of wealth occurs in the stock market from mutual funds to hedge funds investors.

Our paper also contributes to the literature on the price impact of fund flows at the aggregate level. Warther (1995), Edwards and Zhang (1998), Fant (1999), and Edelen and Warner (2001) document a significant positive contemporaneous relation between aggregate mutual fund flows and equity market returns, but argue that this relation is caused by an information effect rather than price pressure effect. On the other hand, using net exchange flows to proxy for investor sentiment, Ben-Rephael et al. (2012) show that aggregate stock market returns initially increase when investors move money from bond funds into equity funds; however, these returns completely reverse over the subsequent ten months. They conclude that aggregate mutual fund flows appear to exert temporary price pressure in the stock market. Our findings corroborate both of the above hypotheses in different contexts. Our data support the price pressure hypothesis rather than the information hypothesis for the case of mutual funds. In contrast, the information hypothesis is corroborated by our hedge fund results.

Finally, our paper contributes to the literature on the "dumb money" effect documented by Frazzini and Lamont (2008). We demonstrate the existence of "dumb money" effect at the aggregate level: the new money flowing into mutual funds appears to be, at least in part, originating from the "dumb" investors described in Frazzini and Lamont's paper. We also complement Frazzini and Lamont's methodology. The conclusion in their paper is based on a negative relation between fund-specific flows and the subsequent performance of fund-specific stock holdings. By contrast, our conclusion is based on the relation between aggregate fund flows and an exogenous, aggregate measure of cross-sectional mispricing. Hence our results both corroborate and strengthen Frazzini and Lamont's "dumb money" conclusion by documenting a dumb money effect at the aggregate level using a different methodology.

The rest of the paper is organized as follows. Section 1 describes our data and empirical methodology. Section 2 presents descriptive statistics. Section 3 documents the contemporaneous relation between mutual fund flows, hedge fund flows, and cross-sectional mispricing. Section 4 documents the predictive relation between fund flows and cross-sectional mispricing. Section 5 presents results from robustness tests, and Section 6 concludes.

1. Data and Variable Construction

To test our hypothesis we require measures of aggregate cross-sectional mispricing, aggregate mutual fund flows, and aggregate hedge fund flows. We describe these measures in this section, along with several control variables that we use in our empirical tests.

A. Measuring Mispricing

Identifying a metric for aggregate cross-sectional mispricing is of critical importance in testing our hypothesis. The metric should be able to isolate a subset of stocks that are most undervalued and another subset that are the most overvalued.

We use the mispricing measure developed by Stambaugh, Yu, and Yuan (2012, 2013). This measure is based on the large number of cross-sectional return anomalies documented in the finance literature that cannot be fully explained by standard risk models. If at least some of the anomalies-based return predictability is due to mispricing, then we can obtain an aggregate measure of mispricing by identifying two subsets of stocks: those classified as the most overvalued and those classified as the most undervalued by the cross-sectional return predictability literature. By tracking the returns of these two subsets during the following calendar month we can determine if mispricing becomes attenuated or exacerbated. For example, if stocks identified as overvalued at the end of month t have positive returns during month $t+1$, we would conclude that mispricing is exacerbated during month $t+1$. The same conclusion would follow if we observed a negative return during $t+1$ for stocks that were undervalued at the end of month t. By contrast, if stocks that are mispriced at the end of month t move during month $t+1$ in a direction opposite to mispricing, we would conclude that mispricing is attenuated.

Following Stambaugh, Yu, and Yuan (2012, 2013) and Cao and Han (2010), we identify stocks with a relatively higher level of mispricing across all eleven return predictability factors. Stambaugh, Yu, and Yuan (2013) show that a combination of high investor sentiment and high short-sale constraints results in the temporary overpricing of stocks. They also show that returns to these individual eleven measures have low correlations with each other, yet are relatively highly correlated with the aggregate returns to a long-short strategy that combines the eleven measures into a single signal. This suggests that each of the eleven factors captures a different facet of cross-sectional mispricing. Therefore, rather than focusing on individual predictability factors, we follow Stambaugh, Yu, and Yuan and use all eleven characteristics to identify stocks that are overvalued or undervalued at the end of each calendar month.[4] Additional details of the eleven predictability characteristics are provided in the Appendix.

We first rank all stocks in our sample each month across the eleven anomaly-based measures. The ranking is performed such that stocks with higher ranks during the subsequent month are expected to have higher average returns while stocks with lower ranks are expected to have lower average returns. For example, stocks with higher past returns will have higher ranks for momentum, while stocks with higher accruals are assigned lower ranks for the accruals anomaly. Each month we compute a stock level "score" as the equal-weighted average of each stock's decile rank for each of the eleven anomaly measures. Higher scores imply higher future return potentials. Stocks are then sorted into decile portfolios based on their scores. We expect that stocks with extreme measures across the eleven characteristics are among the set of stocks most likely to be mispriced.

Following this monthly ranking, we construct a hypothetical "hedge" long-short portfolio that takes long positions in the most undervalued stocks (those with the higher scores) and short positions in the most overvalued stocks (those with the lower scores). The returns to the "long leg" of the strategy are the average monthly returns of stocks deemed to be most undervalued. For the "short leg," the returns are those of stocks deemed to be most overvalued. The returns to the long-short strategy are obtained as the difference between the monthly return series of the long and short legs, respectively.

Recall that the strategy assigns overvalued stocks to the short leg. If the return to the

[4] Using all eleven characteristics – as opposed to individual characteristic – to build the mispricing metric is also justified by the fact that hedge funds normally do not trade on single return predictability attributes.

short leg is positive, it means that these stocks continue to go up in price and become even more overvalued. For undervalued stocks the exacerbation of mispricing operates through the long leg of the strategy, which contains undervalued stocks. The return to the long leg should normally be positive when mispricing corrects itself, but will become negative during months when mispricing deepens. Thus, during months when aggregate mispricing is exacerbated, the long component of the strategy will have negative returns, the short component will have positive returns, and the returns of the long-short strategy will be negative. Conversely, during months when aggregate mispricing is attenuated, the long component of the strategy will have positive returns, the short component will have negative returns, and the returns of the overall long-short strategy will be positive.

The sample used to construct the mispricing metric includes all common stocks listed on NYSE, AMEX and NASDAQ over the period from January 1991 to December 2012. The sample period starts in 1991 to coincide with the availability of monthly mutual fund flow data. We exclude stocks with end of month price of five dollars per share to match the subset of stocks in which mutual funds are likely to invest (Falkenstein, 1996, Brown, Wei, and Wermers, 2014).

B. Measuring Aggregate Mutual Fund Flow

To construct our measure of aggregate mutual fund flow we obtain monthly total net assets and returns from the CRSP Survivor-Bias-Free US Mutual Fund Database for all existing mutual funds. We filter our sample and select only those funds a code of "equity objective," as detailed in Huang, Sialm, and Zhang (2011). To be retained in the sample in a given month we require that each fund have non-missing values for each of the variables used to construct the aggregate measure. Our measure of monthly aggregate mutual fund flow, *AGGMFFLOW*, is computed as:

$$MFFLOW_t = \frac{\sum_{i=1}^{N}[TNA_{i,t} - TNA_{i,t-1}(1 + MRET_{i,t})]}{\sum_{i=1}^{N} TNA_{i,t-1}} \qquad (1)$$

where $TNA_{i,t}$ is the total net assets of mutual fund *i* at time *t*, and $MRET_{i,t}$ is the period return of mutual fund *i* at time *t*, net of fees. Monthly total net assets are available from the end of 1990; therefore, our measure of monthly aggregate mutual fund flow is available from January 1991 to December 2012. Our filtered sample of monthly data used to construct the aggregate measure includes 1,557,985 fund-month observations.

C. Measuring Aggregate Hedge Fund Flow

We construct our aggregate hedge flow measure using net assets and returns from the Lipper TASS database. Our focus is on hedge funds that primarily trade U.S. equities, so we start with U.S. dollar-denominated hedge funds that report returns on a monthly frequency. Consistent with Cao, Chen, Liang, and Lo (2013) we remove funds with strategies that are not primarily based on U.S. equities (e.g., we remove funds whose main strategy is identified as Fixed Income Arbitrage, Managed Futures, and Emerging Markets). We also remove funds whose primary strategy is classified as Fund of Funds to avoid double counting. To be retained in the sample in a given month we require that each fund have non-missing values for each of the variables used to construct the aggregate measure. Our hedge fund sample includes both active and dead funds and starts in January 1994 to minimize survivorship bias.[5] Our measure of monthly aggregate hedge fund flow, *HFFLOW*, is computed as:

$$HFFLOW_t = \frac{\sum_{i=1}^{N}[TNA_{i,t} - TNA_{i,t-1}(1 + HRET_{i,t})]}{\sum_{i=1}^{N} TNA_{i,t-1}} \qquad (2)$$

[5] Fung and Hsieh (2000) provide a detailed discussion of biases in the hedge fund databases. TASS began retaining dead funds in their database starting in 1994 so we begin our sample at this time to minimize survivorship bias. We are less concerned with the selection and incubation biases as we are not looking at individual fund performance, but rather aggregate flows to equity hedge funds.

where $TNA_{i,t}$ is the total net assets of hedge fund i at time t, and $HRET_{i,t}$ is the period return of hedge fund i at time t, net of fees. This measure is available from January 1994 to December 2012. Our filtered sample of monthly data used to construct the aggregate measure includes 279,504 fund-month observations.

D. Control Variables

To appropriately measure the effects of aggregate fund flows on the mispricing metric, we include two control variables that capture the effects of aggregate liquidity and three commonly used risk factors.

First, we note that the return of the long-short strategy should be higher when investors can easily trade to correct mispricing, and the ease of trade should vary with aggregate liquidity. Periods when the market is relatively less liquid should result in more trading frictions that slow down the mispricing correction process. We control for aggregate liquidity using the following two measures:

- AGGILLIQ, the *aggregate illiquidity* computed as the monthly equal-weighted average illiquidity of all common stocks listed on the NYSE with a price greater than \$5 at the end of the previous month. This measure captures the variation in price impact of trade, and we expect relatively less correction of aggregate mispricing during months when the cost of trading is relatively high.

- AGGTURN, the *aggregate turnover* computed as the monthly equal-weighted average turnover of all common stocks listed on the NYSE with a price greater than \$5 at the end of the previous month. This measure captures another dimension of liquidity. When aggregate turnover is high, it is easier for investors to trade in and out of stocks at low costs. Conversely, correction of aggregate mispricing should be more difficult during months with lower turnover.

We control for risk using two standard models, the capital asset pricing model (CAPM), and the three-factor model proposed by Fama and French (1993) which includes the excess return of the stock market (RMRF), the value factor (HML), and the size factor (SMB). While there is ongoing debate as to whether the Fama-French factors -- especially HML -- represent mispricing or risk, the three factor model is now a standard risk control method in the literature.

2. Descriptive statistics

Table 1 provides descriptive statistics of the key variables. Panel A provides univariate statistics for our sample. LONG represents the returns to the portfolio constructed using stocks that are deemed to be undervalued and SHORT represents the returns to the portfolio constructed using stocks that are deemed to be overvalued. Over the course of our sample period, the average monthly excess return on the long portfolio is +143 basis points, while the average monthly excess return on the market portfolio is +60 basis points. During this same period the average monthly excess return to the short portfolio is -45 basis points. The monthly return to the long-short strategy is 188 basis points, which, based on its standard deviation and the sample size, is reliably different from zero, suggesting that our mispricing metric performs quite well in this sample. These results provide prima-facie validation that the LONG and SHORT portfolios include primarily stocks that are undervalued and, respectively, overvalued. Likewise, the LONG-SHORT results provide internal validity for our aggregate measure of cross-sectional mispricing. Also shown in Panel A are measures of aggregate mutual fund flows (MFFLOW) and hedge fund flows (HFFLOW). Both of these variables exhibit sufficient intertemporal variation to allow for meaningful statistical inferences in our empirical tests.

Panel B of Table 1 provides correlations measured over the full sample period. Mutual fund flows and hedge fund flows are positively correlated with each other (ρ=+0.173). Although this correlation is significant, its economic magnitude is sufficiently low to allow for time periods when the two measures might move in opposite directions. Our proxy for mispricing (L-S) is negatively correlated with the market return (ρ=−0.45), suggesting that mispricing is more prone to being corrected during bear markets as opposed to bull markets. MFFLOW is positively correlated with market returns (ρ=+0.308) and negatively correlated with L-S (ρ=−0.159), while HFFLOW is not significantly correlated with the market (ρ=+0.04) and is positively correlated with L-S (ρ=+0.125). These correlations provide a first glimpse of what we will soon document in our main results, namely that flows to mutual funds are "dumb money" that temporarily exacerbate cross-sectional mispricing, while flows to hedge funds are "smart money" that tend to reduce this mispricing.

We also observe differences between MFFLOW and HFFLOW with respect to measures of liquidity. Recall that AGGILLIQ is a measure that reflects the price impact of trade, while AGGTURN is a measure that captures the ease of trade dimension of liquidity. MFFLOW has a strong positive correlation with aggregate illiquidity (ρ=+0.58) and a strong negative correlation with aggregate turnover (ρ=−0.59) suggesting that mutual fund flows are associated with less liquid markets across both dimensions, perhaps because the flows themselves contribute to a high price impact of trade in the underlying stocks. By contrast, the correlation of HFFLOW with aggregate liquidity is lower in economic magnitude and inconsistent across the two measures of liquidity.

Table 2 shows the performance of the mispricing metric, in the form of returns to the long-short strategy. Both raw and abnormal returns (alpha) are shown. The first three columns present the raw returns. The numbers are the same as those presented in Table 1, and are repeated here for completeness. The remaining nine columns present abnormal returns computed using three different asset pricing models: the market model, the Fama-French three-factor model, and a four-factor model that also includes momentum. In all cases the intercepts of the Long-Short strategy are positive and highly significant, with alphas ranging from +180 basis points per month (t-statistic= +8.88) to +211 basis points (t-statistic= +8.06). Having accounted for risk, we observe an interesting asymmetry between the performance of the LONG and SHORT portfolios. While the LONG alphas are positive and significant as expected, most of the alpha in the Long-Short portfolio appears to come from the SHORT side. This suggests that among mispriced stocks, those that are overvalued are more mispriced than those that are undervalued. This asymmetry will also be noted in other results later in the paper and is consistent with our "dumb money" hypothesis that mutual fund flows exacerbate mispricing primarily through net investments in overvalued stocks rather than redemptions of undervalued stocks.

Another interesting aspect of Table 2 is that there is a strong, negative relation between L-S and both the market factor and the size factor, and a positive and marginally significant relation between L-S and the value factor. This suggests, as noted earlier, that mispricing is corrected primarily during bear markets, and also during periods when small stocks underperform large stocks, and during periods when value stocks outperform growth stocks. To control for these relations, we include the market, size, and value factors in each of the subsequent tables.

The extreme magnitudes of the L-S alphas observed in Table 2 provide internal validity for the eleven factors of Stambaugh, Yu, and Yuan serving as a valid measure of aggregate cross-sectional mispricing.

3. Results for Contemporaneous Relations

A. Aggregate Mutual Fund Flows

Turning now to the test of our main hypothesis, we begin by examining the contemporaneous relation between aggregate *mutual fund* flows and the returns to the long-short strategy. Recall that our main hypothesis predicts that these flows exacerbate mispricing, so we expect them to be negatively related to aggregate the mispricing metric. This relation is shown in Table 3. The table also examines the separate LONG and SHORT components in addition to the combined L-S mispricing metric.

Model 1 shows the simple relation without control variables. Mutual fund flows (MFFLOW) have a positive and significant relation with the returns to both the LONG (+2.204, t-statistic=+4.12) and SHORT (+3.172, t-statistic=3.78) components of the mispricing metric, with the magnitude of the SHORT component being significantly higher. This positive coefficient of MFFLOW obtained with the SHORT measure suggests that mutual fund flows accentuate mispricing of overvalued stocks, because these stocks go up in value during months when money flows into mutual funds at the aggregate level. The positive coefficient of MFFLOW obtained with the LONG measure could suggest that mutual fund flows correct mispricing of undervalued stocks, albeit to a lesser extent than they exacerbate mispricing for overvalued stocks; however, this is not robust to the remaining specifications shown in the table.

As expected, the coefficient of MFFLOW obtained with the L-S measure is negative and significant (−0.968, t =−2.22), implying that mutual fund flows, in the aggregate, exacerbate mispricing in the cross-section of US stocks.

Model 2 includes controls for excess market return, aggregate illiquidity, and aggregate turnover. The results are similar to those of Model 1, except that the coefficient of MFFLOW is no longer significant in the LONG leg of the strategy. Model 3 includes additional controls for value and size. The results are similar to those of Model 2.[6] Again, we find no significant relation between flows and the LONG component of the mispricing metric. Overall, the results in Table 2 suggest that aggregate mutual fund flows exacerbate cross-sectional mispricing, and this effect operates primarily through stocks belonging to the SHORT component of the mispricing metric – those who are the most overvalued. In other words, mutual fund flows contribute to cross-sectional mispricing through the purchase of overvalued stocks rather than the sale of undervalued stocks.[7]

B. Aggregate Hedge Fund Flows

We next examine the relation between mispricing metric and aggregate *hedge fund* flows. As discussed previously, we expect flows to hedge funds to be "smart money" that reduces cross-sectional mispricing. That is, when net new money flows into hedge funds, the money should be

[6] In untabulated results, we also include the Betting Against Beta factor (BAB) of Frazzini and Pedersen (2014) as an additional control variable in the Model 2 and 3 specifications. Using the L-S portfolio as the dependent variable, we find that the loading on the MFFLOW variable is quantitatively similar to the results presented in Table 3. The coefficient on the BAB factor is positive and insignificant in both model specifications.

[7] One important distinction between mutual funds and other institutions is that mutual funds are generally prohibited from shorting stocks. The evidence presented in Table 3 suggests that this institutional difference does not drive our results. If mutual funds represent smart money, we should observe evidence of positive price pressure in stocks that are undervalued and negative price pressure associated with the selling of existing positions (if any) in overvalued stocks. Contrary to this, we find that aggregate mutual fund inflows exert greater positive price pressure on overvalued stocks than on undervalued stocks consistent with net inflows being disproportionately invested in overvalued stocks.

used to purchase undervalued stocks, (short) sell overvalued stocks, or both. If our hypothesis is correct, we expect to find a positive contemporaneous relation between hedge fund flows and the aggregate mispricing metric.

Table 4 repeats the analysis in Table 3, except that we now also include hedge fund flows (HFFLOW) in addition to mutual fund flows. We use both flows in the same model since they are positively correlated (Table 1) and the evidence presented in Table 3 suggests that variation in the mispricing metric may be due, at least in part, to dumb money flows. Our hypothesis is that the coefficient on MFFLOW in the long-short strategy should remain negative, while the coefficient on HFFLOW should be significantly positive. The results strongly support this hypothesis. For all three models the coefficients of MFFLOW are significantly negative and those of HFFLOW are significantly positive, all at very high levels of statistical significance.[8]

As we did in Table 3, we also examine the separate LONG and SHORT legs of the mispricing strategy. The MFFLOW results continue to be driven by the short leg; the relation between flows and the returns of SHORT is positive, suggesting that mutual fund flows are used to take long positions on overvalued stocks. The HFFLOW results are also driven by the SHORT leg, but with the opposite sign. This suggests that flows from hedge funds sector are invested primarily in the form of short positions on overvalued stocks. Overall, we conclude that aggregate flows to mutual funds can be qualified as "dumb" money while aggregate flows to hedge funds appear to be "smart" money.

[8] We obtain similar results when MFFLOW is excluded from the regression specification. Repeating the analysis in Table 4 for the L-S portfolio we obtain coefficient estimates on the HFFLOW variable of 0.257, 0.275, and 0.204 (t-statistics = 2.40, 2.66, and 1.94) for Models 1, 2 and 3, respectively.

Though not critical to our main hypothesis, an interesting relation emerges between MFFLOW and HFFLOW in Table 4—the effects of mutual fund flows on cross-sectional mispricing appear to dominate the effects of hedge fund flows, both economically and statistically. Indeed, using the Table 4, Model 3 setting with the full set of controls, we find that a one standard deviation move in mutual fund flows decreases the L-S portfolios return by −1.02%, while a one standard deviation move in hedge fund flows increases that return by +0.56%.

Given the significant profitability of the L-S portfolio documented in Table 2, an interesting question is why smart money flows do not immediately take advantage of dumb money flows?[9] The limits-to-arbitrage literature suggests that market frictions may prevent hedge funds from completely eliminating mispricing. De Long et al. (1990) show that noise traders are a source of risk for arbitrageurs in that their trades can cause mispricing to deepen resulting in short-term losses on arbitrage positions. Shleifer and Vishny (1997) show that periods of poor performance may lead investors to withdraw capital, and that arbitrageurs may reduce arbitrage intensity in anticipation of this possibility. Abreu and Brunnermeier (2002) show that arbitrage may be delayed as arbitrageurs reduce their capital intensity until a critical mass of capital is directed towards the mispricing. Together, this literature suggests that hedge funds may not be able to immediately erase the price effects of dumb money flows.

[9] This relation may be partially due to the fact that data in hedge fund databases are self-reported and participation is not required. Accordingly, the hedge fund flow measure is constructed from a limited set of funds over a limited period of time and thus may not fully characterize aggregate flows.

4. Results for Forward Predictability and Reversal

We now seek to determine if the cross-sectional mispricing associated with *mutual fund* flows reverses itself during subsequent months. We do this by examining the relation between mutual fund flows and *future* returns of those same exact stocks that were originally included in the contemporaneous metric of mispricing. Retail investor flows are characterized as "dumb money" in Frazzini and Lamont (2008) because the price pressure exerted by these flows causes contemporaneous movements in stock prices that subsequently reverse. In our context, if flows to mutual funds are "dumb" as proposed by Frazzini and Lamont, the exacerbation they create in cross-sectional mispricing should correct itself in subsequent months. Thus, we expect to find a positive and significant relation between current mutual fund flows and the *future* returns to the L-S strategy that tracks the *same* stocks selected at t=0. These are the stocks that in Tables 3 and 4 were shown to experience an increase in cross-sectional mispricing.

Turning now to *hedge fund* flows, if these flows represent "smart money" that reduces aggregate mispricing, we expect to find no relation between current flows and future returns to L-S mispricing strategy. That is, if the price effect associated with hedge fund flows represents a correction toward fundamental value, it should not reverse in subsequent months. This is in contrast with the price effect of mutual fund flows, which – because it captures an increase in mispricing – is expected to reverse when stock prices ultimately converge to fundamental value.

A. Aggregate Mutual Fund Flows

Table 5 shows the relation between MFFLOW measured in month *t*, and the cumulative three-month return of the long-short strategy measured during months (*t+1*, *t+3*). If mutual funds exacerbate cross-sectional mispricing and if the mispriced stocks experience a return reversal in the subsequent three months, we expect to find a positive relation between current MFFLOW and future returns to the long-short strategy. The results in Table 5 confirm this reversal. The coefficient on MFFLOW is significantly positive for all three models where L-S is used as dependent variable.

Of particular interest are the results obtained with the short leg of the strategy. Recall that we had previously concluded that mutual fund flows are disproportionately invested in long positions of overvalued stocks, creating temporary upward price pressure for these stocks. If so, we would expect to see a reversal in the price of these stocks during the subsequent three months. Again, Table 5 confirms this conjecture: we find a negative relation between current MFFLOWS and future returns to the short component of the long-short strategy. Since stocks in the short component were overvalued when initially purchased by mutual funds, this negative relation suggests that these same stocks are now converging in price toward equilibrium. While we also observe return reversal in the long leg of the strategy, the disproportionately higher magnitude of return reversal in the short leg leads to the positive relation between aggregate mutual fund flows and future returns to the long-short strategy.

B. Aggregate Hedge Fund Flows

Table 6 repeats the analysis in Table 5 by including HFFLOW in addition to MFFLOW. The relation between MFFLOW and future returns to the long-short strategy remains positive and significant as in Table 5. This contrast with the negative contemporaneous relation in Tables 3 and 4, and signals a price reversal in the stocks purchased with mutual fund flows.

Our main focus in Table 6 is on HFFLOW. Recall that in Table 4 the contemporaneous relation between HFFLOW and the long-short strategy was significantly positive. From that we had inferred that hedge fund flows are used to take the "correct" positions on stocks that are misvalued, particularly overvalued. In Table 6 we ask if the price of these same stocks reverses, as it did for the mutual fund flows. Again, we track, over the period from t+1 to t+3, the price of stocks initially selected according to the mispricing metric at t=0. When future returns to this modified L-S strategy are regressed on HFFLOW, the coefficient is not statistically significant, suggesting the absence of any price reversal. Looking at the long and short legs separately, we see that the coefficient on both legs is significantly negative. This means that stocks that have been purchased by hedge funds experience a slight price reversal (resulting in losses for the hedge funds). However, stocks that had been shorted by hedge funds continue to underperform, resulting in gains for hedge funds that exceed the losses on their long positions.

The results obtained with the short leg in Table 6 corroborate our previous conclusion that hedge fund flows are primarily "smart" money. If these flows are invested into short positions in overvalued stocks, hedge funds will earn significant "alpha" returns because the shorted stocks continue to underperform during the following three months. Once again, this conclusion is consistent with findings from the short-sale literature, which show that short transactions tend to be more informed than long transactions (see, e.g. Boehmer, Jones, and Zhang (2008)).

5. Robustness tests.

We perform several additional tests to assess the robustness of our results.

Detrended Fund Flows: One possible concern with our results is that they could be contaminated by the presence of a time trend in mutual and hedge fund flows – the total dollar amount invested in both types of funds has certainly increased during our sample period. We note that our measure of monthly fund flows approximates new dollar flows as a percentage of net total assets. However, to completely rule out this concern we repeat the analysis in Table 4 using detrended fund flows. To construct the detrended measure we regress aggregate mutual fund flows and, alternatively, aggregate hedge fund flows on a linear time trend and retain the residuals. The results, presented in Table 7, are even stronger than those presented in Table 4. The detrended MFFLOW variable is highly negatively correlated with the long-short strategy, with t-statistics ranging from −4.26 to −5.78. And, the detrended HFFLOW variable maintains a positive and significant relation with the L-S strategy, with t-statistics ranging from +2.61 to +3.46. We conclude that replacing our flow variables with the detrended series does not materially change our conclusion.

Orthogonalized Fund Flows: Another possible concern is that flow variables could be highly correlated with market returns, with aggregate illiquidity, or with turnover. To overcome this concern we regress aggregate mutual fund flows on RMRF, AGGILLIQ, AGGTURN, and HFFLOW and retain the residuals. We then regress aggregate hedge fund flows on RMRF, AGGILLIQ, AGGTURN, and MFFLOW and retain the residuals. We repeat the analysis in Table 4 with these orthogonalized measures and present the results in Table 8. As we did in Table 4, we observe a strong negative relation between the orthogonalized MFFLOW and contemporaneous returns to the long-short strategy. We also observe a positive and significant relation between the orthogonalized HFFLOW variable and the contemporaneous long-short strategy returns. We conclude that replacing our aggregate flow variables with the residual flow variables does not materially affect our conclusion.

Predicted vs. Residual Fund Flows: Both mutual and hedge fund flows are autocorrelated at the annual level. For mutual fund flows the first lag autocorrelation coefficient is +0.66; for hedge fund flows it is +0.39. The presence of autocorrelation brings up an interesting question: are the results driven by the predictable or by the unexpected components of fund flows? To answer this question we perform, for each measure of aggregate fund flows, full-sample regressions of flows on 12-month lagged flows and contemporaneous measures of the following variables: excess market return, aggregate illiquidity, implied volatility of the S&P 500 index, and six macroeconomic variables including growth in industrial production, growth in consumer consumption of durables, non-durables, and services, growth in employment, and an NBER recession dummy variable.[10] Table 9 repeats the analysis in Table 4 replacing the flow variables with their predicted and residual components. We find that predicted variables are generally insignificant, while loadings on residual components are similar to our main results, and are significant at the 1% level. This suggests that our results are not driven by average flows, but rather by periods when flows unexpectedly deviate from the trend.

Figure 1 presents additional evidence of the relation between residual fund flows and the correction (or deepening) of mispricing. We plot the average L-S return for portfolios formed according to the intensity of residual fund flows. Quintile 1 represents average L-S returns during months when residual fund flows are low, and Quintile 5 represents average returns during months when residual fund flows are high. Panel A presents the results when quintile portfolios are formed using residual mutual fund flows. Confirming the results in Table 9, the average return of the L-S return series is high when residual mutual fund flows are low with a

[10] We repeat the analysis in Table 9 augmenting the independent variables in the predictive regressions with the Baker and Wurgler (2006) orthogonal investor sentiment measure (available on Jeffrey Wurgler's website through 2010). Our results (untabulated) are not materially different from results presented in Table 9.

difference between the Q5 and Q1 portfolios of −1.98% per month (*p*-value=0.02). Panel B presents the results when quintile portfolios are formed using residual hedge fund flows. As in Table 9, we find that the average L-S return is high when residual hedge fund flows are high with a difference between the Q5 and Q1 portfolios of +2.49% per month (*p*-value=0.01).

In analyses not tabulated for brevity, we further examine the relation between the strategy returns and residual fund flows. We start by estimating a first order VAR regression using the long-short strategy returns, residual mutual fund flows, and residual hedge fund flows. We then obtain the innovations from the VAR model for each series and examine their contemporaneous correlations. Consistent with the results in Table 9, we find that unexpected shocks of the long-short strategy return series are negatively correlated (coefficient estimate =−0.157, *p*-value=0.0214) with shocks to the mutual fund flows, and are positively correlated with shocks to hedge fund flows (coefficient estimate =+0.154, *p*-value=0.0243).

Volatility and Sentiment: For our last set of robustness tests we account for the possibility that our flow variables could be mere proxies for investor concerns about expected market volatility or investor sentiment in general. To rule out this concern, we independently include in our analysis a measure of market volatility proxied by the level of VIX (implied volatility of the S&P 500 index). We also include a measure of investor sentiment computed as in Baker and Wurgler (2006) using proxies that are orthogonalized against a set of macro variables. Both measures are computed at the time of portfolio formation.

In Table 10 we repeat the analysis of Table 4, except that we also include the level of VIX and level of investor sentiment as additional variables. VIX is not significantly related to the long-short strategy returns in any of the model specifications. Investor sentiment is significantly related to the long-short strategy returns when no control variables are included, but this relation is subsumed by control variables in the Model 2 and Model 3 results. The coefficients on the MFFLOW and HFFLOW remain materially unchanged across all regression specifications.

6. Summary and conclusion

Using mutual and hedge fund flows as proxies for "smart" and "dumb" money, respectively, we document their impacts on cross-sectional equity return anomalies. At the aggregate level mutual fund flows appear to exacerbate mispricing in the cross-section of US stocks. In general, monthly mutual fund flows are associated with a simultaneous *increase* in the price of stocks that are already overvalued at the beginning of the month, causing these stocks to become even more overvalued by the end of the month. This conclusion is corroborated by a reversal in the price of these same exact stocks during the subsequent three months.

In contrast to mutual fund flows, hedge fund flows appear to reduce cross-sectional mispricing. Monthly flows to hedge funds are associated with a simultaneous *decrease* in the price of stocks that are overvalued at the beginning of the month. Consistent with this mispricing correction hypothesis we find no reversal in the price of these stocks during the subsequent three-month period.

We conclude that aggregate mutual fund flows fit the "dumb money" description of Frazzini and Lamont (2008), while aggregate flows to hedge funds are better suited for the "smart money" label we introduce in our paper. Our research has not yet explored implications of our findings for aggregate investor welfare. For example, while there may be a wealth transfer from naïve investors to hedge funds, the economy as a whole may benefit from more efficient prices and consequently, better allocation of real investment. This topic requires further investigation in empirical and theoretical research.

APPENDIX

Our composite measure of aggregate cross-sectional mispricing is based on the following eleven anomalies shown to predict returns in the cross-section of US stocks (e.g., Stambaugh, Yu, and Yuan, 2012):

- Failure Probability: Campbell, Hilscher, and Szilagyi (2007) show that stocks with a high probability of failure have lower future returns.
- O-score: Ohlson (1980) shows that stocks with higher O-Scores (higher probability of bankruptcy) have lower future returns compared to those with lower scores.
- Net Stock Issuances: Ritter (1991) and Loughran and Ritter (1995) show that stocks that issue equity underperform the stocks of nonissuers.
- Composite Equity Issuance: Daniel and Titman (2006) show that firms with higher equity issuance underperform those with lower measures. Composite Equity issues increases with SEOs and share-based acquisitions, and decreases with share repurchases and dividends.
- Accruals: Sloan (1996) shows that stocks with high accruals underperform stocks with low accruals.
- Net Operating Assets: Hirshleifer et al. (2004) show that stocks with higher net operating assets underperform those with lower net operating assets
- Momentum: Jegadeesh and Titman (1993) show that stocks with higher past performance are shown to outperform stocks with lower past perform
- Gross Profitability: Novy-Marx (2010) shows that stocks with higher gross profitability have higher future returns.

- Asset Growth: Cooper, Gulen, and Schill (2008) show that stocks with higher asset growth have lower future returns.
- Return on Assets: Chen, Novy-Marx, and Zhang (2010) show that stocks with higher return on assets have higher future returns.
- Investment-to-Assets: Titman, Wei, and Xie (2004) show that stocks with higher past investment (scaled by total assets) have lower future returns.

REFERENCES

Abreu, Dilip, and Marcus K. Brunnermeier, 2002, Synchronization risk and delayed arbitrage, *Journal of Financial Economics* 66, 341-360.

Akbas, Ferhat, Will J. Armstrong, Sorin Sorescu and Avanidhar Subrahmanyam, 2014, Capital market efficiency and arbitrage efficacy, forthcoming, *Journal of Financial and Quantitative Analysis.*

Baker, Macolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, Journal of Finance 61, 1645-1680.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl, 2012, Measuring investor sentiment with mutual fund flows, *Journal of Financial Economics* 104, 363-382.

Brown, Nerissa C., Kelsey D. Wei, and Russ Wermers, 2014, Analyst recommendations, mutual fund herding, and overreaction in stock prices, *Management Science*, 1-20.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899-2939.

Cao, Charles, Yong Chen, Bing Liang, and Andrew W. Lo, 2013, Can hedge funds time market liquidity? *Journal of Financial Economics* 109, 493-516.

Cao, Jie, and Bing Han, 2010, Idiosyncratic Risk, Costly Arbitrage, and Cross-section of Stock Returns, *Working paper, The University of Texas at Austin*.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 16091652.

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

Daniel, Kent D., and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605-1643.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-738.

Ederington, Louis, and Evgenia Golubeva, 2011, The impact of stock market volatility expectations on investor behavior: Evidence from aggregate mutual fund flows. *Working paper*.

Falkenstein, Eric, G., 1996, Preference for stock characteristics as revealed by mutual fund portfolio holdings, *Journal of Finance* 51, 111-135.

Fama, Eugene and Kenneth French, 2006, Profitability, investment, and average returns, *Journal of Financial Economics* 82, 491518.

Frazzini, Andrea, and Owen A. Lamont, 2008, Dumb money: Mutual Fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299-322.

Frazzini, Andrea and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

Fung, William and David A. Hsieh, 2000, Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, *Journal of Financial and Quantitative Analysis* 35, 291-307.

Harvey, Campbell R., Yan Liu, and Heging Zhu, 2013, … And the Cross-Section of Expected Returns, *Duke University Working Paper*, November 9, 2013.

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.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2011, Risk shifting and mutual fund performance,*Review of Financial Studies* 24, 2575-2616.

Jagannathan, Ravi, Alexey Malakhov, Dmitry Novikov, 2010, Do hot hands exist among hedge fund managers? An empirical evaluation, *Journal of Finance* 65, 217-255.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for market efficiency, *Journal of Finance* 48, 6591.

Kokkonen, Joni and Matti Suominen, 2014, Hedge Funds and Stock Market Efficiency, working paper, Aalto University.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3547-3489.

McLean, David, and Jeffrey Pontiff, 2013, Does academic research destroy stock return predictability?, *Working paper, Boston College.*

Novy-Marx, Robert, 2010, The other side of value: Good growth and the gross profitability premium, working paper, University of Chicago.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109-131.

Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 327.

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

Titman, S., Wei, K.C.J., Xie, F., 2004. Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677-700.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52 (1), 35-55.

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

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2013, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Working paper*.

Wermers, Russ, 2003, Is money really "Smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, *Working paper*.

Warther, Vincent A., 1995, Aggregate mutual fund flows and security returns, *Journal of Financial Economics* 39, 209-235.

Figure 1: Residual Fund Flows and Long minus Short (L-S) Strategy Performance

Figure 1 presents the equal-weighted average Long minus Short (L-S) return of the cross-sectional mispricing metric for quintile portfolios formed using residual fund flows. Q1 represents the average monthly L-S returns when residual fund flows are low, and Q5 represents the average monthly L-S returns when residual fund flows are high. Panel A presents results when portfolios are formed using residual mutual fund flows. Panel B presents results when portfolios are formed using residual hedge fund flows. Residual mutual (hedge) fund flows are estimated using a full sample regression of aggregate mutual (hedge) fund flows on 12 months of past mutual (hedge) fund flows and contemporaneous measures of the following variables: excess market return, aggregate illiquidity, implied volatility of the S&P 500 index (VIX), and six macroeconomic variables including growth in industrial production, growth in consumer consumption of durables, non-durables, and services, growth in employment, and an NBER recession dummy variable.

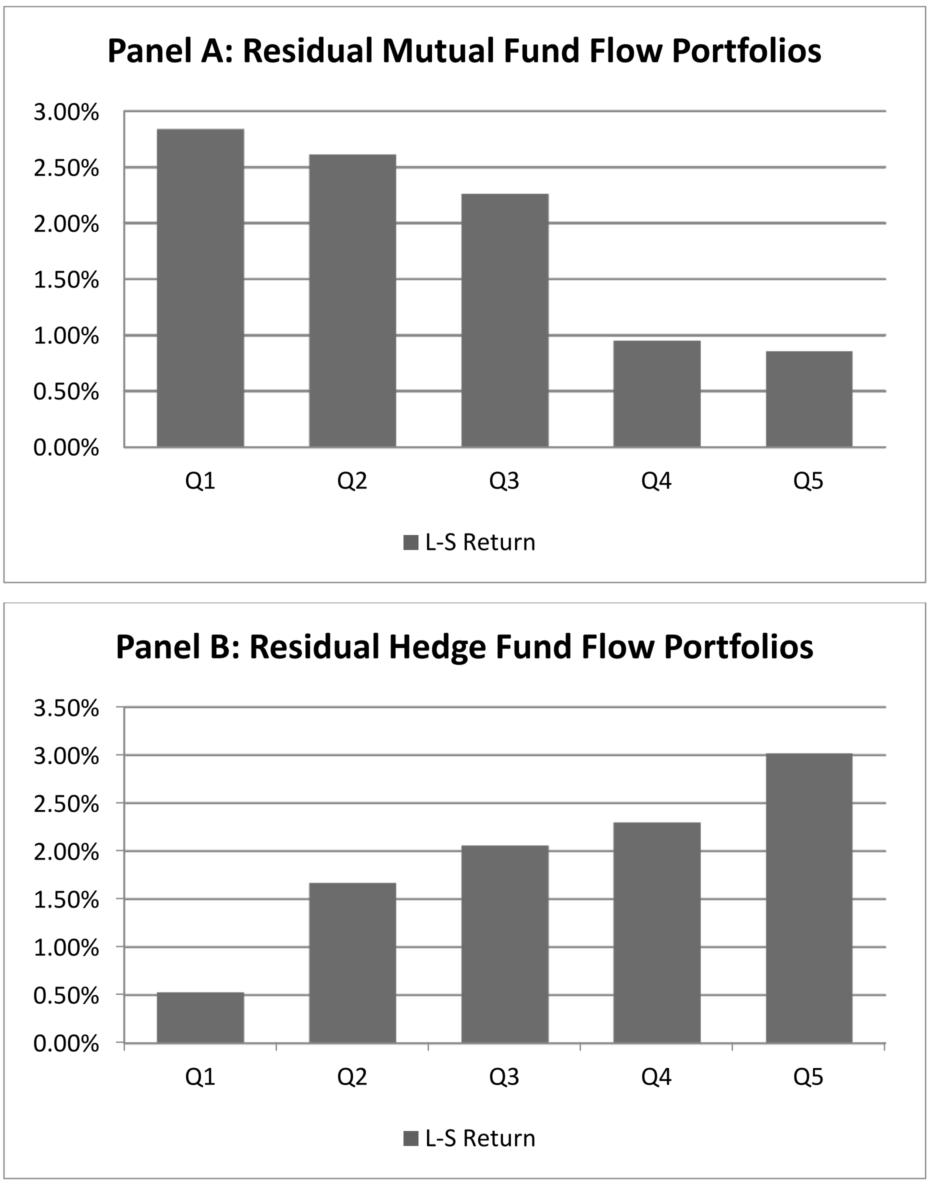

Table 1: Summary Statistics

Shown below are summary statistics of key monthly variables measured over the period 1991 to 2012 (hedge fund measures are available from 1994 to 2012). The key flow variables are MFFLOW and HFFLOW which respectively represent the mean monthly aggregate flow of equity mutual funds and equity hedge funds. Details on their construction are provided in Sections 1.B and 1.C, respectively. Aggregate control variables include monthly excess market returns (RMRF), aggregate illiquidity (AGGILLIQ), and aggregate turnover (AGGTURN). Details on their construction are provided in Section 1.D. We also report summary statistics of the distribution of returns to a composite anomaly-based trading strategy that is constructed using the eleven anomalies documented in Stambaugh, Yu, and Yuan (2012). LONG, SHORT, and L-S represents returns to the long, short, and long-short components of the mispricing metric, respectively. Details on the construction of the mispricing metric are provided in Section 1.A.

Panel A: Descriptive Statistics, 1991 to 2012

Variable	N	Mean	Median	St.Dev.	Min	P10	P25	P75	P90	Max
MFFLOW	264	0.0045	0.0038	0.006	−0.015	−0.002	0.000	0.008	0.012	0.025
HFFLOW	228	0.0076	0.0095	0.019	−0.102	−0.011	0.000	0.019	0.025	0.074
RMRF	264	0.0060	0.0118	0.044	−0.172	−0.050	−0.020	0.034	0.061	0.113
AGGILLIQ	264	0.0527	0.0462	0.035	0.008	0.016	0.023	0.069	0.101	0.229
AGGTURN	264	0.1341	0.1089	0.077	0.048	0.059	0.070	0.190	0.250	0.403
LONG	264	0.0143	0.0171	0.051	−0.189	−0.054	−0.015	0.048	0.074	0.176
SHORT	264	−0.0045	0.0017	0.071	−0.273	−0.094	−0.049	0.039	0.077	0.185
L-S	264	0.0188	0.0160	0.038	−0.102	−0.022	−0.003	0.036	0.062	0.157

Panel B: Pairwise Correlations, 1991 to 2012

	MFFLOW	HFFLOW	RMRF	AGGILLIQ	AGGTURN	LONG	SHORT
HFFLOW	0.173						
	0.01						
RMRF	0.308	0.040					
	0.00	0.55					
AGGILLIQ	0.580	−0.171	0.074				
	0.00	0.01	0.23				
AGGTURN	−0.591	−0.179	−0.120	−0.633			
	0.00	0.01	0.05	0.00			
LONG	0.266	0.072	0.833	0.045	−0.125		
	0.00	0.28	0.00	0.47	0.04		
SHORT	0.279	−0.015	0.847	0.027	−0.029	0.854	
	0.00	0.82	0.00	0.66	0.64	0.00	
L-S	−0.159	0.125	−0.450	0.011	−0.116	−0.236	−0.707
	0.01	0.06	0.00	0.86	0.06	0.00	0.00

* p-values listed below correlation estimates

Table 2: Mispricing Metric: Returns to a Long-Short Strategy that uses Cross-Sectional Return Predictors, 1991 to 2012

Shown below are the mean excess returns, market model alphas, Fama and French 3-factor alphas, and Fama and French 4-factor alphas of a cross-sectional trading strategy that is used as a proxy for cross-sectional mispricing. Results are reported for the Long, Short, and Long minus Short (L-S) legs of the strategy. Details on the construction of the mispricing metric are provided in Section 1.A. The t-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Mean Excess Returns			Market Model Alphas			Fama-French 3-factor Alphas			Fama-French 4-factor Alphas		
	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S
Intercept	0.0143	−0.0045	0.0188	0.0085	−0.0126	0.0211	0.0076	−0.0120	0.0197	0.0055	−0.0125	0.0180
	4.51	−0.98	6.71	3.90	−5.17	8.06	4.75	−5.84	8.68	3.56	−6.72	8.88
RMRF				0.9705	1.3564	−0.3859	0.9386	1.2650	−0.3264	0.9410	1.2655	−0.3245
				19.29	22.16	−4.62	14.74	19.29	−5.25	15.53	19.05	−5.49
HML							0.2198	−0.0016	0.2214	0.2099	−0.0037	0.2137
							2.71	−0.01	1.72	2.61	−0.03	1.73
SMB							0.4588	0.6239	−0.1650	0.4570	0.6235	−0.1665
							2.83	3.19	−2.62	3.00	3.20	−2.61
UMD										0.1609	0.0347	0.1262
										2.03	0.46	1.65
N	264	264	264	264	264	264	264	264	264	264	264	264
$Adj - R^2$				0.693	0.717	0.200	0.774	0.803	0.267	0.798	0.803	0.292

Table 3: Aggregate Mutual Fund Flows and Cross-Sectional Mispricing, 1991 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long, Short, or Long minus Short (L-S) component of the cross-sectional mispricing metric. The independent variables, measured contemporaneously with the dependent variable, include the following: aggregate mutual fund flow (MFFLOW), excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). Details on the construction of the mispricing metric are provided in Section 1.A. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1			Model 2			Model 3		
	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S
MFFLOW	2.204	3.172	−0.968	0.111	1.204	−1.093	−0.284	0.970	−1.254
	4.12	3.78	−2.22	0.32	2.62	−2.31	−0.94	1.98	−2.55
RMRF				0.963	1.330	−0.367	0.943	1.244	−0.301
				19.12	21.62	−4.89	16.15	20.82	−5.43
AGGILLIQ				−0.084	−0.038	−0.047	−0.043	−0.034	−0.009
				−1.23	−0.44	−0.69	−0.88	−0.64	−0.13
AGGTURN				−0.037	0.112	−0.148	−0.044	0.084	−0.128
				−1.09	2.76	−4.56	−1.83	2.30	−3.99
HML							0.216	−0.008	0.224
							2.70	−0.07	1.88
SMB							0.463	0.607	−0.144
							2.76	3.10	−2.28
Intercept	0.004	−0.019	0.023	0.017	−0.031	0.048	0.017	−0.026	0.043
	1.05	−2.76	5.65	2.04	−3.16	5.84	3.04	−3.40	5.49
N	264	264	264	264	264	264	264	264	264
$Adj-R^2$	0.067	0.074	0.022	0.692	0.725	0.244	0.774	0.808	0.304

Table 4: Aggregate Mutual Fund Flows, Hedge Fund Flows and Cross-Sectional Mispricing, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long, Short, or Long minus Short (L-S) component of the cross-sectional mispricing metric. The independent variables, measured contemporaneously with the dependent variable, include the following: aggregate mutual fund flow (MFFLOW), aggregate hedge fund flow (HFFLOW), excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). Details on the construction of the mispricing metric are provided in Section 1.A. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1			Model 2			Model 3		
	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S
MFFLOW	3.065	4.594	−1.528	0.114	2.069	−1.955	−0.247	1.693	−1.941
	4.62	4.44	−2.68	0.27	3.86	−3.74	−0.60	2.53	−3.62
HFFLOW	0.053	−0.277	0.330	0.054	−0.310	0.364	0.061	−0.230	0.291
	0.23	−0.92	2.77	0.49	−2.35	3.56	0.59	−2.02	2.72
RMRF				0.960	1.322	−0.361	0.942	1.253	−0.311
				18.23	21.62	−4.70	16.21	21.43	−5.07
AGGILLIQ				−0.095	−0.372	0.277	0.056	−0.205	0.261
				−0.97	−3.71	2.70	0.67	−2.80	2.76
AGGTURN				−0.038	0.075	−0.113	−0.030	0.072	−0.102
				−1.14	1.82	−3.16	−1.17	1.85	−2.84
HML							0.232	0.022	0.210
							2.98	0.18	1.68
SMB							0.450	0.560	−0.109
							2.62	2.90	−1.66
Intercept	0.003	−0.019	0.022	0.018	−0.012	0.030	0.011	−0.017	0.028
	0.71	−2.52	5.56	2.08	−1.21	3.24	1.60	−2.10	3.13
N	228	228	228	228	228	228	228	228	228
$Adj-R^2$	0.088	0.098	0.048	0.695	0.751	0.305	0.770	0.815	0.348

Table 5: Aggregate Mutual Fund Flows and Future Cross-Sectional Mispricing, 1991 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the future cross-sectional mispricing, proxied by the 3-month forward looking return [t+1,t+3] of the strategy that trades on cross-sectional return predictability. The independent variables are aggregate mutual fund flow (MFFLOW), excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). Details on the construction of the mispricing metric are provided in Section 1.A. The t-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1			Model 2			Model 3		
	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S
MFFLOW	−1.690	−3.745	2.250	−5.927	−8.407	2.977	−5.485	−7.759	2.735
	−1.69	−2.22	2.04	−3.56	−3.41	2.09	−3.55	−3.20	1.87
RMRF				0.248	0.464	−0.218	0.151	0.379	−0.230
				1.81	2.25	−1.87	0.98	1.64	−2.07
AGGILLIQ				0.531	1.016	−0.537	0.458	0.923	−0.514
				1.88	2.45	−2.37	1.57	2.16	−2.26
AGGTURN				−0.267	−0.043	−0.206	−0.294	−0.065	−0.212
				−1.37	−0.15	−1.76	−1.43	−0.22	−1.77
HML							−0.442	−0.546	0.124
							−2.05	−1.67	0.56
SMB							0.040	−0.195	0.270
							0.22	−0.63	1.24
Intercept	0.045	0.009	0.037	0.071	−0.021	0.091	0.079	−0.013	0.091
	3.94	0.54	4.17	1.89	−0.38	3.25	1.98	−0.22	3.24
N	260	260	261	260	260	261	260	260	261
$Adj-R^2$	0.008	0.027	0.027	0.084	0.084	0.064	0.100	0.092	0.069

Table 6: Aggregate Mutual Fund Flows, Hedge Fund Flows and Future Cross-Sectional Mispricing, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the future cross-sectional mispricing, proxied by the 3-month forward looking return [t+1,t+3] of the strategy that trades on cross-sectional return predictability. The independent variables are aggregate mutual fund flow (MFFLOW), aggregate hedge fund flow (HFFLOW), excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). Details on the construction of the mispricing metric are provided in Section 1.A. The t-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1			Model 2			Model 3		
	LONG	SHORT	L-S	LONG	SHORT	L-S	LONG	SHORT	L-S
MFFLOW	−1.206	−5.175	4.391	−5.598	−9.621	4.775	−5.570	−9.325	4.474
	−1.04	−2.28	2.99	−2.49	−2.76	2.64	−2.54	−2.70	2.45
HFFLOW	−0.945	−1.323	0.320	−0.880	−1.155	0.203	−0.760	−1.020	0.184
	−2.19	−2.37	1.14	−2.47	−2.91	0.93	−2.08	−2.45	0.80
RMRF				0.254	0.530	−0.288	0.190	0.470	−0.292
				1.56	2.18	−2.34	1.11	1.79	−2.30
AGGILLIQ				0.508	0.589	−0.167	0.530	0.513	−0.058
				1.19	0.90	−0.47	1.19	0.73	−0.16
AGGTURN				−0.320	−0.190	−0.116	−0.332	−0.208	−0.110
				−1.56	−0.66	−0.99	−1.55	−0.68	−0.90
HML							−0.299	−0.498	0.228
							−1.33	−1.43	0.93
SMB							0.104	−0.196	0.334
							0.50	−0.60	1.47
Intercept	0.049	0.017	0.033	0.086	0.030	0.058	0.088	0.036	0.053
	4.00	0.98	3.66	2.01	0.49	1.90	1.94	0.55	1.72
N	224	224	225	224	224	225	224	224	225
$Adj-R^2$	0.030	0.072	0.083	0.100	0.103	0.100	0.105	0.106	0.111

Table 7: Detrended Aggregate Mutual Fund and Hedge Fund Flows and Cross-Sectional Mispricing, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long minus Short (L-S) return series of the cross-sectional mispricing metric. The independent variables of interest are detrended aggregate mutual fund flow (DT AGGMFFLOW) and detrended aggregate hedge fund flow (DT AGGHFFLOW). Detrended flow variables are computed as the residuals obtained by regressing flows on a time trend. Control variables include excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1 L-S	Model 2 L-S	Model 3 L-S
DT MFFLOW	−3.980	−2.270	−2.357
	−5.78	−4.40	−4.26
DT HFFLOW	0.351	0.350	0.274
	3.01	3.46	2.61
RMRF		−0.342	−0.286
		−4.35	−4.58
AGGILLIQ		0.151	0.134
		1.58	1.47
AGGTURN		−0.059	−0.045
		−1.90	−1.51
HML			0.223
			1.79
SMB			−0.108
			−1.66
Intercept	0.019	0.023	0.020
	6.74	2.74	2.58
N	228	228	228
$Adj-R^2$	0.176	0.313	0.359

Table 8: Orthogonalized Aggregate Mutual Fund and Hedge Fund Flows and Cross-Sectional Mispricing, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long minus Short (L-S) return series of the mispricing metric. The independent variables of interest are are aggregate mutual fund flow orthogonalized against the contemporaneous values of the other independent variables (ORTH AGGMFFLOW) and aggregate hedge fund flow orthogonalized against the contemporaneous values of the other independent variables (ORTH AGGHFFLOW). Control variables include excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1 L-S	Model 2 L-S	Model 3 L-S
ORTH MFFLOW	−1.703	−1.703	−1.755
	−2.76	−3.19	−3.14
ORTH HFFLOW	0.286	0.286	0.211
	2.62	2.73	1.88
RMRF		−0.428	−0.377
		−5.68	−6.60
AGGILLIQ		0.040	0.050
		0.44	0.55
AGGTURN		−0.105	−0.086
		−3.38	−2.76
HML			0.210
			1.68
SMB			−0.109
			−1.66
Intercept	0.020	0.036	0.031
	6.41	4.24	3.79
N	228	228	228
$Adj-R^2$	0.045	0.305	0.348

Table 9: Predicted vs. Residual Components of Aggregate Fund Flows and Cross-Sectional Mispricing, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long minus Short (L-S) return series of the cross-sectional mispricing metric. The independent variables of interest are predicted and residual aggregate mutual fund flow (PRED AGGMFFLOW and RES AGGMFFLOW) and predicted and residual aggregate hedge fund flow (PRED AGGHFFLOW and RES AGGHFFLOW). Predicted and residual mutual (hedge) fund flows are estimated using a full-sample regression of aggregate mutual (hedge) fund flows on 12 months of past mutual (hedge) fund flows and contemporaneous measures of the following variables: excess market return, aggregate illiquidity, implied volatility of the S&P 500 index (VIX), and six macroeconomic variables including growth in industrial production, growth in consumer consumption of durables, non-durables, and services, growth in employment, and an NBER recession dummy variable. Control variables include excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Model 1 L-S	Model 2 L-S	Model 3 L-S
PRED MFFLOW	−1.174	−0.370	−0.568
	−1.83	−0.40	−0.66
RES MFFLOW	−2.525	−2.976	−2.817
	−3.31	−4.43	−4.18
PRED HFFLOW	0.111	0.057	−0.020
	0.46	0.28	−0.10
RES HFFLOW	0.530	0.520	0.434
	2.43	3.23	2.87
RMRF		−0.433	−0.375
		−4.87	−5.15
AGGILLIQ		0.110	0.084
		0.63	0.53
AGGTURN		−0.116	−0.110
		−2.81	−2.81
HML			0.204
			1.68
SMB			−0.101
			−1.52
Intercept	0.022	0.035	0.035
	4.32	2.69	2.91
N	216	216	216
$Adj-R^2$	0.061	0.329	0.367

Table 10: Relation between Aggregate Fund Flows and Cross-Sectional Mispricing Controlling for Implied Volatility and Investor Sentiment, 1994 to 2012

Shown below are coefficient estimates of time-series regressions where the dependent variable is the monthly Long minus Short (L-S) return series of the cross-sectional mispricing metric. The independent variables of interest are aggregate mutual fund flow (MFFLOW), aggregate hedge fund flow (HFFLOW), implied volatility of the S&P 500 index (VIX), and the Baker and Wurgler (2006) measure of investor sentiment (SENTORTH, available from 1994 to 2010). VIX is measured at the end of the prior period. SENTORTH represents the orthogonal investor sentiment measure as of the end of the prior month. Control variables include excess market return (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Implied Volatility (VIX) L-S			Sentiment (SENTORTH) L-S		
MFFLOW	−1.601	−1.907	−1.878	−1.614	−1.905	−1.957
	−2.57	−3.55	−3.32	−2.90	−3.62	−3.40
HFFLOW	0.293	0.375	0.306	0.321	0.372	0.270
	2.49	3.37	2.70	2.63	3.06	1.83
VIX	−0.023	0.012	0.016			
	−0.57	0.42	0.56			
SENTORTH				0.018	0.009	0.006
				2.51	1.22	0.86
RMRF		−0.364	−0.315		−0.337	−0.292
		−4.78	−5.03		−3.94	−4.45
AGGILLIQ		0.261	0.239		0.250	0.222
		2.66	2.81		2.48	2.32
AGGTURN		−0.118	−0.108		−0.099	−0.098
		−3.35	−3.22		−2.43	−2.36
HML			0.212			0.206
			1.70			1.49
SMB			−0.108			−0.105
			−1.62			−1.55
Intercept	0.028	0.029	0.026	0.022	0.028	0.029
	2.59	2.87	2.61	6.01	2.66	2.78
N	228	228	228	205	205	205
$Adj-R^2$	0.045	0.303	0.346	0.120	0.307	0.346