ARTICLE IN PRESS

Journal of Financial Economics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Smart money, dumb money, and capital market anomalies ☆

Ferhat Akbas [a], Will J. Armstrong [b], Sorin Sorescu [c,*], Avanidhar Subrahmanyam [d]

[a] School of Business, University of Kansas, Lawrence, KS 66045-7601, USA
[b] Rawls College of Business, Texas Tech University, Box 42101, Lubbock, TX 79409-2101, USA
[c] Mays Business School, Texas A&M University, Department of Finance, 360 Wehner Building, College Station, TX 77843-4218, USA
[d] Anderson School, University of California at Los Angeles, 110 Westwood Plaza, Los Angeles, CA 90095-1481, USA

ARTICLE INFO

Article history:
Received 7 July 2014
Received in revised form
17 March 2015
Accepted 18 March 2015

JEL classification:
G11
G14
G23

Keywords:
Stock return anomalies
Mutual funds
Hedge funds
Fund flows
Mispricing

ABSTRACT

We investigate the dual notions that "dumb money" exacerbates well-known stock return anomalies and "smart money" attenuates these anomalies. We find that aggregate flows to mutual funds (dumb money) appear to exacerbate cross-sectional mispricing, particularly for growth, accrual, and momentum anomalies. In contrast, hedge fund flows (smart money) appear to attenuate aggregate mispricing. Our results suggest that aggregate flows to mutual funds can have real adverse allocation effects in the stock market and that aggregate flows to hedge funds contribute to the correction of cross-sectional mispricing.

1. Introduction

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 dumb money generally is presupposed to make prices depart from fundamentals (Lou, 2012), whereas arbitrage by the smart money 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, forthcoming), suggesting that prices could depart from fundamentals for periods of time, and the persistence of such anomalies indicates that smart money is not fully able to erase them. Even though these notions prevail in financial thought, no direct documentation yet exists of the role of dumb and smart money in causing or correcting anomalies. In this paper, we provide evidence 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 stocks. Retail

☆ We are grateful to an anonymous referee for insightful and constructive suggestions. We also thank Jason Chen, Yong Chen, Francesco Franzoni, Hagen Kim, Lu Zhang, 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.
* Corresponding author. Tel.: +1 979 458 0380; fax: +1 979 845 3884.
E-mail address: ssorescu@tamu.edu (S. Sorescu).
[1] See, for example, "The Smart Way to Follow Dumb Money," by S. Jakab, available at http://online.wsj.com/news/articles/SB10001424052702304543904577396361227824738.

http://dx.doi.org/10.1016/j.jfineco.2015.07.003
0304-405X/© 2015 Published by Elsevier B.V.

investors appear to contribute to these distortions in several ways. Sirri and Tufano (1998) show that retail investors 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. Frazzini and Lamont (2008) show that retail investors tend to direct dumb money to mutual funds that hold overvalued stocks. 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 returns and subsequent return reversal (Coval and Stafford, 2007).

Lou (2012) shows that high-performing mutual funds tend to attract relatively higher flows, which are then reinvested by fund managers into their existing stock holdings. Similarly, mutual funds with poor performance tend to liquidate existing holdings to meet redemptions. Price pressure from the purchases of recent winners (or liquidation of recent losers) causes return continuation. The combination of performance chasing by investors and tendency by mutual fund managers to invest into existing holdings leads to a positive contemporaneous relation between stock-level fund flows and individual stock returns. In turn, this relation allows an understanding of the well-known momentum anomaly: the tendency of past winners to outperform past losers.

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 stock returns, in spite of significant arbitrage trading strategies carried out by quant-oriented hedge funds over the past two decades. Motivated by the above observations, we examine the inter-temporal relation between two time series: the aggregate mutual fund flows and an aggregate measure of monthly cross-sectional equity mispricing that includes several well-known equity return anomalies. Moreover, to understand the channel through which fund flows affect cross-sectional mispricing, we examine the relation between flows and returns to each individual anomaly.

We use, as a proxy for aggregate mispricing, the metric proposed by Stambaugh, Yu, and Yuan (2012, forthcoming). We identify each month stocks that are most likely to be overvalued or undervalued based on 11 characteristics that are known to predict the cross section of stock returns. We then compute the return on a hedge strategy that is long undervalued stocks and short overvalued stocks. This return is 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 move toward fundamentals. By contrast, the strategy should produce negative returns when stock prices diverge from fundamental values and cross-sectional mispricing is exacerbated.

Aggregate flows to mutual funds vary through time as a result of changing investors' sentiment and aggregate fear (see, e.g. Ben-Rephael, Kandel, and Wohl, 2012 and Ederington and Golubeva, 2011) or as a result of past returns to arbitrage strategies (Akbas, Armstrong, Sorescu, and Subrahmanyam, forthcoming). We take advantage of this intertemporal 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, then we would expect to see a negative contemporaneous relation between the two time series.

Our results 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. In subsequent tests 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 likely to be overvalued (by construction), we conclude that mutual fund flows exacerbate cross-sectional mispricing because they are invested disproportionately into stocks that are already overvalued.

If mutual funds disproportionately purchase stocks that are already overvalued, and if the resulting price pressure further exacerbates these stocks' overvaluation, we would expect the stocks to experience a price reversal following periods of high aggregate mutual fund flows, as prices converge toward the efficient market benchmark. This would yield positive future returns to the long–short hedge strategy (which remains short these overvalued stocks). Our results support this prediction as well. Moreover, we show that this relation once again comes exclusively from overvalued stocks (or the short leg of the hedge strategy).

We next ask if any smart money is present in the market. We define smart money as aggregate fund flows that take long positions in undervalued stocks or short positions in overvalued stocks, the opposite of what mutual funds do. The smart money description does not apply only to hedge fund investors but could also apply to hedge fund managers, in which better compensation incentives, combined with the ability to take short positions, could result in smarter investment decisions. 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, Malakhov, and Novikov (2010) and by Kokkonen and Suominen (2014), hedge funds are one such group of sophisticated investors, and we expect that the effect of aggregate hedge fund flows on mispricing will be the opposite of mutual fund flows. We find that the effect of hedge fund flows on mispricing is significantly positive. 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

[2] Because our focus is on identifying stocks that are the most mispriced in the cross section, we use the Stambaugh, Yu, and Yuan (2012,2014) measure as a proxy for cross-sectional mispricing instead of as a performance measure.

absence of any predictive relation between hedge fund flows and future returns of the mispricing metric. Once hedge fund flows have corrected mispricing, prices remain corrected.

Examining the long and short components separately, we observe that hedge funds' corrective effect is driven primarily by overvalued stocks. That is, aggregate hedge fund flows appear to be most effective when they take short positions in overvalued stocks instead of long positions in undervalued stocks. This result is consistent not only with our previous results obtained in the case of mutual funds, as well as with a long-standing literature on short sales, which shows that short transactions are generally more 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.[3]

Turning now to individual anomalies, our analysis suggests that anomalies related to growth, momentum, and accrual are the main channels through which mutual fund flows exacerbate cross-sectional mispricing. In contrast, mutual fund flows do not appear to affect cross-sectional pricing of anomalies related to real investments. In the case of hedge funds, we find little evidence that hedge fund investors exhibit a demand for stock characteristics associated with individual anomalies, despite these flows having a significant corrective effect on the aggregate mispricing metric. This asymmetry could reflect a higher noise level in hedge fund flow data or the notion that hedge funds trade on more complex signals of mispricing. Additional analysis supports this conjecture. We combine anomalies into two groups: those based on real investment and those based on accruals, growth, and momentum. We find that mutual fund flows exacerbate mispricing and hedge fund flows tend to correct mispricing among the latter group. In contrast, neither flow type is related to anomalies that are based on real investments. We also show that hedge fund flows are strongly related to a composite portfolio based on anomalies that were known early in our sample period, suggesting that hedge fund managers tend to trade on anomalies of which they are aware.

Our paper has implications for the market efficiency literature. A significant puzzle in the literature is the persistence of cross-sectional return predictability despite the 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 might not completely vanish (see, e.g., Shleifer and Vishny, 1997).

We propose an explanation for the prevalence of anomalies that is complementary to the one provided by limits-to-arbitrage.[4] The effect we find occurs earlier in the sequence of events that leads to cross sectional mispricing. We conjecture that the cross-sectional mispricing is itself 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 in which at least some traders are sophisticated.

Our analysis thus suggests that despite cross-sectional return predictability now being common knowledge among sophisticated investors, judicious strategies that seek to exploit this predictability should continue to earn positive alphas so long as dumb money enters the stock market via the mutual fund industry. An aggregate consequence of this dumb money is to create a market for smart money, i.e. hedge funds, in which investors can earn alpha by merely trading against the price pressure induced by these dumb flows.

Our paper is related to work 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 they argue that this relation is caused by an information effect, not a price pressure effect. Using net exchange flows to proxy for investor sentiment, Ben-Rephael, Kandel, and Wohl (2012) show that aggregate stock market returns initially increase when investors move money from bond funds into equity funds but completely reverse over the subsequent 10 months, which suggests temporary price pressures. Our findings corroborate both of these hypotheses in different contexts. Our data support the price pressure hypothesis instead of the information hypothesis for the case of mutual funds. In contrast, the information hypothesis is corroborated by our hedge fund results.

We also contribute to the literature on the dumb money effect shown by Frazzini and Lamont (2008) and by Lou (2012). We demonstrate the existence of a 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

[3] In related work, Shive and Yun (2013) argue that patient institutional investors are able to profit from front-running mutual fund trades, but they do not link their analyses to stock return anomalies. DeVault, Sias, and Starks (2014) show that periods of positive sentiment lead to increases in institutional demands. Edelen, Ince, and Kadlec (2014) show that institutional money is generally on the wrong side of return anomalies, i.e., institutions exacerbate mispricing. These last two studies do not focus on hedge fund flows, however.

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

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 identifying a dumb money effect at the aggregate level, using a different methodology.

Finally, we add to the results of Lou (2012). While in Lou's paper mutual fund flows cause momentum due to rigid investment rules by fund managers, our paper shows that mutual fund flows can also exacerbate cross-sectional mispricing using a broader measure of mispricing.[5] An implication that is common to both Lou's paper and ours is that we both raise concerns about a previous conclusion in the literature that smart money is present in the mutual fund sector. For example, Gruber (1996), Zheng (1999), and Keswani and Stolin (2008) find a smart money effect in mutual funds, whereby retail money tends to flow to funds that do well over the next quarter. The Lou (2012) results suggest that the smart money effect shown by Gruber (and subsequent papers) is driven by a mechanical price pressure resulting from fund managers' tendency to invest new flows into existing holdings, not by the managers' ability to identify undervalued stocks. To support his conclusion, Lou shows that the superior, one-quarter performance reverses over longer horizons. Our paper complements Lou (2012) by proposing that smart money does exist, but not in the mutual fund sector. We find evidence of smart money in the hedge fund sector.

2. 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.

2.1. Measuring mispricing

We use the aggregate mispricing measure developed by Stambaugh, Yu, and Yuan (2012, forthcoming). This measure is based on 11 cross-sectional return anomalies shown 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 classified 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 observe 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 conclude that mispricing is attenuated.

Following Stambaugh, Yu, and Yuan (2012, forthcoming) and Cao and Han (2010), we identify stocks with a relatively higher level of mispricing across all 11 return predictors. Stambaugh, Yu, and Yuan (forthcoming) 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 predictors have low correlations with each other, yet are relatively highly correlated with the aggregate returns to a long–short strategy that combines the measures into a single signal. This suggests that each of the 11 components captures a different facet of cross-sectional mispricing. Therefore, while we do consider individual predictability characteristics, we follow Stambaugh, Yu, and Yuan (forthcoming) and also construct an aggregate mispricing metric to identify stocks that are overvalued or undervalued at the end of each calendar month.[6] Additional details of the predictability characteristics are provided in the Appendix.

To construct the aggregate mispricing metric, we first score all stocks in our sample each month according to their future returns predicted by each of the anomalies, and we group the scores into deciles. Each stock, therefore, is ascribed 11 different deciles rankings each month, one for each anomaly. The scoring is performed such that stocks with higher scores are expected to have higher average returns during the subsequent month and stocks with lower scores are expected to have lower average returns. For example, stocks with higher past returns have higher scores for momentum, and stocks with higher accruals are assigned lower scores for the accruals anomaly. We then compute an aggregate score for each stock, each month. The aggregate score is the equal-weighted average of the decile ranks previously computed. Higher aggregate scores imply higher future return potentials and we expect stocks with the most extreme aggregate scores to be among the most mispriced.

Following this monthly scoring we construct a hypothetical 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

[5] Excluding the momentum factor reduces the economic significance of the influence of fund flows on mispricing. However, our results remain robust to excluding momentum in that the statistical significance is unaltered.

[6] Using all 11 characteristics, as opposed to individual characteristics, to build an aggregate mispricing metric is also justified by the fact that hedge funds normally do not trade on single return predictability attributes. Moreover, the aggregate metric is justified by the fact that the Sharpe ratio of the long–short returns based on this aggregate metric is higher than the Sharpe ratios of the individual anomalies. In separate (untabulated) tests we compare the Sharpe ratio of the aggregate mispricing metric with that of each individual anomaly. In all cases, the Sharpe ratio of the aggregate metric is higher than that of individual anomalies, and this difference is statistically significant (p-value < 0.05) in 10 out of 11 cases. These results corroborate the Stambaugh, Yu, and Yuan conjecture that the aggregate metric "diversifies away [the] noise in each individual anomaly and … increases precision" and justifies using the aggregate mispricing metric instead of individual anomaly metrics.

Please cite this article as: Akbas, F., et al., Smart money, dumb money, and capital market anomalies. Journal of Financial Economics (2015), http://dx.doi.org/10.1016/j.jfineco.2015.07.003

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.

The strategy assigns overvalued stocks to the short leg. If the return to the 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 it should become negative during months when mispricing deepens. Thus, during months when aggregate mispricing is exacerbated, the long component of the strategy has negative returns, the short component has positive returns, and the returns of the long–short strategy are negative. Conversely, during months when aggregate mispricing is attenuated, the long component of the strategy has positive returns, the short component has negative returns, and the returns of the overall long–short strategy are positive.

The sample used to construct the mispricing metric includes all common stocks listed on NYSE, AMEX, and Nasdaq over the period from January 1994 to December 2012. The sample period starts in 1994 to coincide with the availability of monthly hedge fund flow data. We exclude stocks with an end-of-month price of less than 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).

2.2. Measuring aggregate mutual fund flow

To construct our measure of aggregate mutual fund flow we obtain monthly total net assets and returns from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free US Mutual Fund Database for all existing mutual funds. We filter our sample and select only those funds with 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

$$\mathrm{MFFLOW}_t = \frac{\sum_{i=1}^{N}\left[\mathrm{TNA}_{i,t} - \mathrm{TNA}_{i,t-1}(1+\mathrm{MRET}_{i,t})\right]}{\sum_{i=1}^{N}\mathrm{TNA}_{i,t-1}}, \tag{1}$$

where $\mathrm{TNA}_{i,t}$ is the total net assets of mutual fund i at time t and $\mathrm{MRET}_{i,t}$ 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. However, data on hedge fund flows are available only beginning in January 1994. Therefore, our aggregate mutual fund flows are measured over the period from January 1994 to December 2012 (our results on mutual fund flows are robust to the 1991–2012 period). Our filtered sample of monthly data used to construct the aggregate measure contains 1,522,775 fund-month observations.

2.3. Measuring aggregate hedge fund flow

We construct an aggregate measure of hedge fund flows using net assets and returns from the Lipper TASS database. Our focus is on hedge funds that primarily trade US equities, so we start with hedge funds denominated in US dollars, which report returns on a monthly basis. Consistent with Cao, Chen, Liang, and Lo (2013), we remove funds with strategies that are not primarily based on US equities (e.g., funds whose main strategy is identified as fixed income arbitrage, managed futures, or 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.[7] Our measure of monthly aggregate hedge fund flow, HFFLOW, is computed as

$$\mathrm{HFFLOW}_t = \frac{\sum_{i=1}^{N}\left[\mathrm{TNA}_{i,t} - \mathrm{TNA}_{i,t-1}(1+\mathrm{HRET}_{i,t})\right]}{\sum_{i=1}^{N}\mathrm{TNA}_{i,t-1}}, \tag{2}$$

where $\mathrm{TNA}_{i,t}$ is the total net assets of hedge fund i at time t and $\mathrm{HRET}_{i,t}$ 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 has 279,504 fund-month observations.

2.4. Control variables

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

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:

1. AGGILLIQ, the aggregate illiquidity computed as the monthly equal-weighted average illiquidity of all common stocks listed on the NYSE with a share price greater than five dollars 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.
2. AGGTURN, the aggregate turnover computed as the monthly equal-weighted average turnover of all

[7] Fung and Hsieh (2000) provide a detailed discussion of biases in the hedge fund databases. TASS began retaining dead funds in its 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.

common stocks listed on the NYSE with a share price greater than five dollars 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 the three-factor model proposed by Fama and French (1993): 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 and French factors (especially HML) represent mispricing or risk, the three factor model is now a standard risk control method in the literature.[8]

3. 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, according to the aggregate mispricing metric. Over the course of our sample period, the average monthly excess return on the LONG portfolio is +138 basis points, and the average monthly excess return on the market portfolio is +52 basis points. During this same period the average monthly excess return to the SHORT portfolio is −60 basis points. The monthly return to the long–short strategy is +198 basis points, which, based on its standard deviation and sample size, is reliably different from zero, suggesting that our mispricing metric performs well in this sample. These results provide prima facie validation that the LONG and SHORT portfolios include primarily stocks that are, respectively, undervalued and overvalued. Likewise, the long minus short (L−S) 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 aggregate 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 ($\rho=+0.173$). Although this correlation is significant, its economic magnitude is sufficiently low to allow for time periods when the two measures move in opposite directions. Our proxy for mispricing (L−S) is negatively correlated with the market return ($\rho=-0.47$), suggesting that mispricing is more prone to being corrected during bear markets than during bull markets. MFFLOW is positively correlated with market returns ($\rho=+0.345$) and negatively correlated with L−S ($\rho=-0.177$), and HFFLOW is not significantly correlated with the market ($\rho=+0.040$) and is positively correlated with L−S ($\rho=+0.125$). These correlations provide a first glimpse of what we show in our main results, namely, that flows to mutual funds are dumb money that temporarily exacerbate cross-sectional mispricing, and that 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. AGGILLIQ is a measure that reflects the price impact of trade, and AGGTURN is a measure that captures the ease of trade dimension of liquidity. MFFLOW has a strong positive correlation with aggregate illiquidity ($\rho=+0.479$) and a strong negative correlation with aggregate turnover ($\rho=-0.544$), 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 intertemporal performance of hedge portfolios constructed based on the aggregate mispricing metric (Panel A) and based on the prediction of individual anomalies (Panel B). The first three columns in Panel A present the raw returns of the aggregate metric hedge portfolio. The numbers are the same as those presented in Table 1 and are repeated here for completeness. The remaining six columns present abnormal returns computed using two different asset pricing models: the Fama and 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 +187 basis points ($t=+8.28$) to +204 basis points ($t=+8.06$) per month. 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 also is 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 instead of redemptions of undervalued stocks.

Another interesting aspect of Panel A is that a strong, negative relation exists between L–S and both the market factor and the size factor and a positive and marginally significant relation exists between L–S and the HML factor. This suggests that mispricing is corrected primarily during bear markets, 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.

In Panel B of Table 2 we provide alphas for the individual anomalies, using both the three- and four-factor models. Under either specification, we find that the L-S portfolio alphas are significant for every anomaly at the 10% level or better. Overall, the significance and magnitudes of these alphas provide validity for the

[8] Our results are also robust to including the momentum factor, in addition to the three Fama and French (1993) factors.

Table 1
Summary statistics.

Shown below are summary statistics of key monthly variables measured over the period 1994–2012. 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 2.2 and 2.3, respectively. Aggregate control variables are monthly excess market returns (RMRF), aggregate illiquidity (AGGILLIQ), aggregate turnover (AGGTURN), returns to the value strategy (HML), and returns to the size strategy (SMB). Details on their construction are provided in Section 2.4. We also report summary statistics of the distribution of returns to a composite anomaly-based trading strategy that is constructed using the 11 anomalies shown in Stambaugh, Yu, and Yuan (2012). LONG, SHORT, and L-S represents returns to the long, short, and long–short return series of the mispricing metric, respectively. Details on the construction of the mispricing metric are provided in Section 2.1. *p*-Values are listed below the correlation estimates.

Panel A: Descriptive statistics, 1994–2012

Variable	*N*	Mean	Median	Standard deviation	Minimum	10th percentile	25th percentile	75th percentile	90th percentile	Maximum
MFFLOW	228	0.0033	0.0030	0.005	−0.015	−0.003	0.000	0.006	0.010	0.022
HFFLOW	228	0.0076	0.0095	0.019	−0.102	−0.011	0.000	0.019	0.025	0.074
LONG	228	0.0138	0.0142	0.053	−0.189	−0.055	−0.015	0.050	0.074	0.176
SHORT	228	−0.0060	0.0015	0.074	−0.273	−0.100	−0.053	0.042	0.085	0.185
L−S	228	0.0198	0.0171	0.040	−0.102	−0.023	−0.003	0.037	0.068	0.157
RMRF	228	0.0052	0.0119	0.046	−0.172	−0.058	−0.022	0.035	0.061	0.113
AGGILLIQ	228	0.0427	0.0405	0.024	0.008	0.016	0.022	0.058	0.078	0.128
AGGTURN	228	0.1460	0.1242	0.076	0.049	0.067	0.080	0.199	0.258	0.403
HML	228	0.0045	0.0025	0.033	−0.098	−0.031	−0.014	0.019	0.042	0.138
SMB	228	−0.0012	−0.0016	0.035	−0.220	−0.037	−0.021	0.022	0.037	0.077

Panel B: Pairwise correlations, 1994–2012

Variable	MFFLOW	HFFLOW	LONG	SHORT	L-S	RMRF	AGGILLIQ	AGGTURN	HML
HFFLOW	0.173								
	0.01								
LONG	0.309	0.072							
	0.00	0.28							
SHORT	0.318	−0.015	0.854						
	0.00	0.82	0.00						
L-S	−0.177	0.125	−0.247	−0.715					
	0.01	0.06	0.00	0.00					
RMRF	0.345	0.040	0.836	0.855	−0.470				
	0.00	0.55	0.00	0.00	0.00				
AGGILLIQ	0.479	−0.171	0.018	−0.056	0.130	0.032			
	0.00	0.01	0.78	0.40	0.05	0.63			
AGGTURN	−0.544	−0.179	−0.134	−0.012	−0.158	−0.117	−0.602		
	0.00	0.01	0.04	0.86	0.02	0.08	0.00		
HML	0.024	0.185	−0.119	−0.259	0.323	−0.204	−0.001	−0.104	
	0.71	0.01	0.07	0.00	0.00	0.00	0.99	0.12	
SMB	0.037	−0.029	0.392	0.433	−0.280	0.190	−0.145	0.080	−0.353
	0.58	0.67	0.00	0.00	0.00	0.00	0.03	0.23	0.00

Table 2
Mispricing metric: returns to a long–short strategy that uses cross-sectional return predictors, 1994–2012.

Shown below, in Panel A, are the mean excess returns, Fama and French three-factor alphas, and Fama and French four-factor alphas (including the momentum factor) of a cross-sectional trading strategy that is used as a proxy for cross-sectional mispricing. Results are reported for the long (LONG), short (SHORT), and long minus short (L − S) legs of the strategy. Details on the construction of the mispricing metric are provided in Section 2.1. Panel B presents the intercepts from regressions of the returns of the 11 individual cross-sectional anomalies used to construct the mispricing metric on the Fama and French three-factor model (FF3 alphas) and the Fama and French four-factor model (FF4 alphas). Coefficient estimates on the factors in Panel B have been suppressed for brevity. Details of the individual anomalies are provided in the Appendix. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

Panel A: Mispricing metric returns

	Mean excess returns			Fama and French three-factor alphas			Fama and French four-factor alphas		
Variable	L − S	LONG	SHORT	L − S	LONG	SHORT	L − S	LONG	SHORT
Alpha	0.0198	0.0138	−0.0060	0.0204	0.0084	−0.0120	0.0187	0.0060	−0.0127
	6.26	3.88	−1.15	8.06	4.73	−5.20	8.28	3.49	−5.94
RMRF				−0.3506	0.9429	1.2935	−0.3471	0.9478	1.2949
				−5.48	14.30	19.89	−5.63	14.88	19.50
HML				0.2312	0.2409	0.0097	0.2217	0.2276	0.0059
				1.64	2.90	0.07	1.63	2.77	0.04
SMB				−0.1515	0.4402	0.5916	−0.1564	0.4333	0.5897
				−2.27	2.67	3.01	−2.33	2.81	3.00
UMD							0.1191	0.1662	0.0472
							1.50	1.97	0.61
N	228	228	228	228	228	228	228	228	228
Adj. R^2				0.281	0.771	0.805	0.303	0.797	0.805

Panel B: Individual anomaly returns

	Return on assets (1)	Ohlson O-score (2)	Failure probability (3)	Gross profitability (4)	Net stock issues (5)	Total accrual (6)	Composite equity issues (7)	Investment- to-assets (8)	Net operating assets (9)	Asset growth (10)	Momentum (11)
FF3 alphas	0.0060	0.0066	0.0227	0.0094	0.0105	0.0042	0.0143	0.0035	0.0098	0.0089	0.0163
	1.82	2.88	8.97	3.14	4.78	2.53	4.69	2.33	2.75	3.61	4.37
Adj. R^2	0.110	0.054	0.241	0.079	0.386	0.021	0.081	0.035	0.092	0.173	0.050
FF4 alphas	0.0092	0.0094	0.0249	0.0116	0.0115	0.0030	0.0092	0.0034	0.0069	0.0064	0.0104
	2.45	2.92	9.43	3.72	5.04	1.89	3.18	2.12	1.86	2.60	2.08
Adj. R^2	0.151	0.109	0.268	0.097	0.391	0.052	0.233	0.031	0.133	0.234	0.135

Table 3
Aggregate mutual fund flows, hedge fund flows, and individual anomaly returns, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L–S) return series of the 11 individual cross-sectional anomalies used to construct the mispricing metric. The independent variables of interest, measured contemporaneously with the dependent variable, are aggregate mutual fund flow (MFFLOW) and aggregate hedge fund flow (HFFLOW). Control variables are 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.

Individual anomaly returns											
Variable	Return on assets (1)	Ohlson O-score (2)	Failure probability (3)	Gross profitability (4)	Net stock issues (5)	Total accruals (6)	Composite equity issues (7)	Investment-to-assets (8)	Net operating assets (9)	Asset growth (10)	Momentum (11)
MFFLOW	−1.840	−1.755	−2.111	−1.833	−1.311	−0.988	−0.192	0.546	0.764	0.188	−1.846
	−1.91	−2.53	−2.90	−2.00	−2.11	−2.88	−0.26	1.58	0.88	0.47	−1.91
HFFLOW	0.151	0.216	0.147	0.226	0.010	0.103	−0.074	0.077	0.012	−0.066	0.295
	0.91	1.31	1.11	2.03	0.09	1.54	−0.45	0.86	0.07	−0.54	0.87
RMRF	−0.335	−0.184	−0.439	−0.185	−0.386	0.079	0.344	−0.138	0.106	−0.130	−0.194
	−3.76	−2.67	−5.33	−2.13	−5.54	1.52	3.45	−2.75	1.06	−1.83	−1.07
AGGILLIQ	0.303	0.477	−0.048	0.267	0.229	0.097	0.284	0.074	0.136	0.027	−0.309
	2.19	4.88	−0.39	2.08	2.45	1.47	1.88	0.88	0.88	0.32	−1.36
AGGTURN	−0.016	0.032	−0.152	−0.034	−0.018	−0.053	0.026	0.027	−0.003	−0.048	−0.287
	−0.35	0.99	−3.17	−0.89	−0.55	−2.30	0.41	0.69	−0.04	−1.46	−3.68
HML	0.182	−0.026	−0.052	0.258	0.484	0.038	0.081	−0.056	−0.403	0.391	0.007
	1.43	−0.33	−0.36	1.96	4.92	0.61	0.56	−0.93	−3.23	3.76	0.03
SMB	0.062	0.032	−0.139	0.380	−0.012	−0.081	−0.014	0.037	−0.374	−0.037	−0.360
	0.25	0.20	−1.04	2.00	−0.09	−1.01	−0.09	0.91	−1.66	−0.35	−1.82
Intercept	0.000	−0.015	0.053	0.007	0.007	0.010	−0.001	−0.006	0.002	0.015	0.076
	−0.01	−1.73	4.47	0.68	0.87	1.64	−0.04	−0.69	0.13	1.63	3.68
N	228	228	228	228	228	228	228	228	228	228	228
Adj. R^2	0.114	0.085	0.281	0.086	0.394	0.049	0.080	0.036	0.091	0.175	0.096

predictors of Stambaugh, Yu, and Yuan (2012) serving as a measure of cross-sectional mispricing.

4. The effect of flows on individual anomalies

We now turn to the tests of our main hypothesis. We begin by examining the contemporaneous relation between fund flows and the long–short returns obtained with each individual anomaly. This allows us to understand the channels through which fund flows affect aggregate mispricing and to determine which stock characteristics are viewed as attractive by mutual fund and hedge fund managers.

We create 11 dependent variables, each corresponding to the long–short returns of a portfolio that is based on a single anomaly. For example, for the return on assets (ROA) anomaly, we form a long–short portfolio that buys stocks with high ROA and shorts stocks with low ROA. We repeat this process for the remaining 10 anomalies of Stambaugh, Yu, and Yuan. We then regress the long–short returns to each individual anomaly on aggregate mutual fund flows, hedge fund flows, and the set of control variables discussed in Section 3. The results are presented in Table 3.

The first line in Table 3 shows that the relation between aggregate mutual fund flows and returns to the individual anomalies is negative and significant for seven factors: ROA, Ohlson O-score, failure probability, gross profitability, net stock issues, accruals (columns 1–6) and Momentum (column 11). However, this relation is insignificant for the composite equity issues, Investment-to-Asset, Net Operating Assets, and Asset Growth anomalies in columns (7), (8), (9), and (10), respectively.

Upon closer examination, a commonality emerges among the last four factors: Those are characteristics of stocks that usually belong to the real investment predictability category.[9] The real investment phenomenon refers to tendencies of firms with high past investment in real assets to under-perform in the cross section of stock returns. This is an interesting finding because each of these four factors is shown in Table 2 to independently predict the cross section of stock returns. The fact that these factors are unrelated to mutual fund flows suggests that mutual fund investors do not appear to chase stocks that score highly on real investment–type predictability. In other words, the real investment channel is not one through which mutual fund flows exacerbate cross-sectional mispricing. In contrast, mutual fund flows appear to be disproportionately directed to stocks that are overvalued with respect to factors other than real investment.

In examining the anomalies in which mutual fund flows seem to exacerbate mispricing, anomalies in Columns 1–5 in Table 3 appear to share a common

[9] Composite equity issues in column 7 measures the issuance of equity for all purposes, including for stock-only acquisitions, capturing the real investment component through acquisition activities. The investment-to-assets factor in Column 8 measures the amount of past real investment in relation to the firms' total assets. The net operating assets factor in Column 9 captures the real investment aspect in that firms with high past investments are likely to have higher net operating assets. Finally, the asset growth factor in Column 10 measures the previous change in total assets, a change that clearly captures the real investment component. We thank Lu Zhang for this observation.

characteristic: they relate to the well-known value-growth dimension.[10] The value-growth concept is embodied in these five factors as follows: (1) The O-score factor measures the probability of bankruptcy, and the most important component of the O-score is leverage (see Table 4 in Ohlson, 1980). Mutual fund flows directed to stocks with higher O-scores are also directed to stocks with higher leverage and thus, to stocks with higher market-to-book multiples, a well-known characteristic of growth stocks. (2) The failure probability of Campbell, Hilscher, and Szilagyi (2008) is strongly related to leverage (see Table III in their paper). Thus, mutual fund flows directed to stocks with high failure probability are essentially directed to stocks with high market-to-book multiples, or growth stocks. (3) The gross profitability factor captures the firm's profit margin. Mutual fund flows directed to firms with lower profit margins that are likely to be directed to stocks with higher multiples, or growth stocks. (4) The return on assets (ROA) factor captures firm profitability in relation to total assets. Mutual fund flows directed to firms with low ROA are essentially directed to firms with high assets relative to earnings, another dimension of the growth concept. (5) The net stock issues factor measures seasoned equity offers (SEO) activity. Mutual fund flows directed to firms with high SEO are also likely directed to firms with low current cash flows because such firms must rely on the SEO market to raise the capital needed for future growth. Thus, stocks that are overvalued with respect to anomalies in Columns 1 through 5 appear to be primarily growth stocks.[11]

A characteristic common to growth stocks is the unusually high rate of earning growth preceding the portfolio formation period. Thus, investor tendency to extrapolate growth rates in past earnings (Lakoniskok, Shleifer, and Vishny, 1994) could cause them to chase stocks with growth characteristics, precisely the stocks that should be avoided. This magnet toward overvalued stocks could be magnified by the fact that many mutual funds explicitly market themselves as growth funds, implying that growth is an acceptable investment philosophy, contrary to established evidence (see, e.g., Frazzini and Lamont, 2008).

The anomaly in Column 6 from Table 3 captures the accrual predictability factor, the non-cash component of earnings. Stocks that are overvalued with respect to the accrual anomaly are also likely to be glamor stocks, with earnings partially inflated by the use of accruals. This relation has been shown by Desai, Rajgopal, and Venkatachalam (2004), who argue that the accruals anomaly is related to the glamour stock phenomenon.

Overall, we conclude that new mutual fund flows are directed primarily to the purchase of stocks with growth and high accrual characteristics, precisely the type of stocks that should be avoided. In contrast, mutual fund flows do not appear to chase characteristics that are related to the real investment dimension of cross-sectional predictability. Thus, in addition to the Lou (2012) momentum channel, the tendency of mutual fund investors to chase stocks with growth and accrual characteristic appears to be an important channel through which mutual fund flows exacerbate cross-sectional mispricing.

Turning now to the second row in Table 3, we examine the effect of hedge fund flows on individual anomalies. The coefficients of HFFLOW carry mixed signs and do not attain significance in most cases. The general lack of significance could reflect a higher noise in the aggregate hedge fund flows data. It could also result from the notion that hedge funds do not normally trade on single anomalies because the precision associated with individual anomalies is lower than that associated with composite measures of mispricing.

Unlike mutual fund flows' effect on mispricing, which is most likely unintentional, the effect of hedge fund flows on mispricing is likely intentional. That is, if hedge fund flows are smart money, hedge fund managers deliberately target stocks whose return predictability is known with high precision (such as high Sharpe ratios of long–short strategies). Therefore, in the case of hedge funds, the use of the higher-precision signal becomes critical to align the empirical tests with the actual behavior of managers. For this reason, we would expect hedge fund managers to trade on composite signals based on several anomalies instead of on signals obtained from individual anomalies.

To test this conjecture, we group the individual anomalies into two complementary sets, based upon the year when the anomaly was first discovered. We expect hedge funds to trade on anomalies that had been known for a longer period of time. We also regroup the same anomalies into two different complementary sets, according to the nature of the anomaly: growth, momentum, and accruals on one side versus real investments on the other. For each of these four sets, we compute a composite mispricing measure that is based on the anomalies only in that particular set. We then repeat the analysis in Table 3 using the four composite mispricing measures instead of the

[10] Broadly speaking, we view the value-growth dimension as one in which stocks are sorted based on the affordability of their market price in relation to their earning potential. A value stock is one that is particularly cheap in relation to its assets in place or to its earning potential. These could be stocks trading for low multiples of cash flows or stock with low market-to-book ratios. By contrast, a growth stock is one that is particularly expensive. The value versus growth predictability factor refers to the tendency of value stocks to outperform growth stocks in the cross section of returns.

[11] We conduct two tests to confirm our interpretation of anomalies in Columns 1 through 5 as value versus growth. First, we conduct a principal component analysis on the 11 mispricing factors. The first five are the only ones that load positively onto the first principal component. (The first principal component explains 49% of the variation among the 11 factors). Second, we construct an alternative value-growth factor using the ratio of enterprise value to free cash flows. This implementation of the value-growth dimension is common in the active asset management industry. We approximate enterprise value as the book value of assets, plus the market value of equity, minus the book value of equity. We approximate free cash flows as earnings plus depreciation. Firms whose enterprise value is a high multiple of cash flows are deemed growth firms and their stocks are expected to under-perform. Likewise, firms with low multiples are value firms and should outperform. We then build the alternative value-growth factor by taking long positions on low-multiple firms and short positions on high-multiple firm. We then measure the correlation between the first principal component and the long–short return of this alternative growth factor. We find a very high correlation (ρ=0.9002, $p<0.0001$), which corroborates our conclusion that the anomalies in Columns 1 through 5 of Table 3 capture the value versus growth dimension.

Table 4

Aggregate mutual fund flows, aggregate hedge fund flows, and cross-sectional mispricing: combinations of individual anomalies, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L−S) return series of four combinations of the 11 anomalies using to construct the mispricing metric. The four combinations are non-investment-based anomalies, investment- based anomalies, anomalies documented prior to 1997, and anomalies from 1997 forward. The seven non-investment-based anomalies are return on assets, Ohlson's O-score, failure probability, gross profitability, net stock issues, total accruals, and momentum. The four investment-based anomalies are investment-to- assets, asset growth, net operating assets, and composite equity issues. The four anomalies documented in the literature prior to 1997 (and year of publication) are Ohlson O-score (1980), net stock issues (1991), momentum (1993), and total accruals (1996). The seven anomalies from 1997 forward are failure probability (2008), composite equity issues (2006), net operating assets (2004), gross profitability premium (2010), asset growth (2008), return on assets (2006), and investment-to-assets (2004). The publication dates are from papers cited in Stambaugh, Yu, and Yuan (2012). The independent variables, measured contemporaneously with the dependent 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 2.1. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

Variable	Non-investment anomaly group	Investment anomaly group	Four anomalies documented prior to 1997	Seven anomalies from 1997 forward
MFFLOW	−2.927	0.297	−2.464	−0.990
	−3.33	0.46	−3.81	−1.96
HFFLOW	0.367	0.033	0.356	0.153
	3.29	0.21	2.92	1.78
RMRF	−0.478	0.109	−0.333	−0.221
	−6.92	1.19	−5.04	−4.03
AGGILLIQ	0.280	0.225	0.200	0.300
	2.63	1.75	1.97	2.63
AGGTURN	−0.123	0.013	−0.132	−0.049
	−3.25	0.26	−3.69	−1.11
HML	0.228	0.071	0.327	−0.005
	1.60	0.70	2.22	−0.05
SMB	−0.100	−0.085	−0.211	−0.109
	−0.68	−0.56	−2.75	−1.97
Intercept	0.030	−0.003	0.030	0.015
	3.38	−0.26	3.38	1.48
N	228	228	228	228
Adj. R^2	0.360	0.007	0.445	0.210

long–short returns to individual anomalies. The results are presented in Table 4.

The first column in Table 4 shows the results obtained with a composite mispricing metric that excludes the real investment factors. As expected, the loading on mutual fund flows is negative, consistent with our findings from Table 3. The loading on hedge fund flows is now significantly positive, despite the fact that individual loadings in Table 3 were generally insignificant. The second column in Table 4 shows the results obtained with a composite mispricing metric formed exclusively from real investment-based anomalies. Neither mutual fund flows nor hedge fund flows have any significant effect on this composite mispricing metric.

The results in the first two columns of Table 4 corroborate our earlier conjecture that mutual funds flows mainly affect mispricing related to the growth, momentum, and accrual dimensions. In contrast, mutual fund flows do not appear to affect mispricing related to the real investment dimension of cross-sectional predictability. The hedge fund results are also interesting. Hedge fund flows appear to be invested in a manner that corrects mispricing along the same dimensions in which mispricing is exacerbated by mutual fund flows. The results in the first two columns present an interesting symmetry in that hedge fund flows and mutual fund flows are both significant in the first column (albeit of opposite signs) and are both insignificant in the second column. This symmetry allows us to refine our concept of smartness in the case of hedge fund flows. These flows are particularly smart, in that they trade precisely on the type of return predictability induced by mutual fund flows but tend to stay away from trading on predictability that is not related to mutual fund flows.[12]

The last two columns of Table 4 examine the relation between fund flows and the two composite metrics obtained by segregating anomalies into old versus new. We examine the date of publication for each of the anomalies of Stambaugh, Yu, and Yuan and note that four of these anomalies were discovered prior to the year 1997: Ohlson O-score (1980), net stock issues (1991), momentum (1993), and total accruals (1996). The remaining anomalies were discovered after 2004. Given a typical three-year lag required for a working paper to go through the review process and appear in print, it is reasonable to assume that all anomalies published before 1997 were known (or at least knowable) to hedge fund investors as early as 1994, the year when our sample period begins. Accordingly, we group the Stambaugh-Yu-Yuan predictors into two groups that correspond to the time period when each anomaly was discovered, using the year 1997 as the cutoff year.

We then repeat the analysis from the first two columns with the two new composite mispricing metrics: an old anomaly metric that aggregates the four anomalies known prior to 1997 and a new anomaly metric that aggregates the remaining seven anomalies. The results are presented in third and fourth columns of Table 4.

Corroborating our conjecture that hedge fund flows are more effective at correcting mispricing induced by anomalies that are known to hedge fund managers, the coefficient on hedge fund flows is $+0.356$ ($t=+2.92$) in the case of old anomalies (third column), compared with a lower $+0.153$ ($t=+1.78$) in the case of new anomalies (fourth column). The difference between the two coefficients is statistically significant (p-value$=0.0629$). This is consistent with hedge funds' trading being more closely aligned with the less noisy, four anomaly strategy.

Our results in Section 4 also carry a potentially important implication for the asset management industry. Our

[12] This could be because hedge fund investors tend to be fairly sophisticated and usually require an economic reason for why an anomaly shown in a backtest should be expected to work out of sample. That is, hedge fund investors are more likely to invest money into an anomaly such as growth, if there is a good economic explanation for the anomaly. Price pressure from mutual funds (dumb money) provides one such good explanation for some anomalies.

finding that hedge funds are less likely to allocate capital to real investment anomalies could be explained in the context of *q*-theory (see, e.g., Hou, Xue, and Zhang, 2015). The theory holds that real investment predicts returns, in part, because of differences among firms in terms of cost of equity. If the Fama-French factors in our paper do not completely span the pricing kernel, some of the return predictability that in Table 2 we find for real investment anomalies could be driven by unobserved differences among firms in terms of cost of equity.

Thus, our results imply that hedge fund investors are particularly sophisticated in that they are less likely to allocate funds to anomalies whose return predictability could reflect an unobserved risk factor and are more likely to allocate capital to anomalies whose predictability is fueled by temporary price pressure caused by the less sophisticated mutual fund investors.

5. The effects of flows on mispricing

Having analyzed the effect of flows on each individual anomaly, we now shift our focus to an aggregate metric of mispricing that is based on all 11 anomalies. Following Stambaugh, Yu, and Yuan (2012, forthcoming), we combine the anomalies into a single mispricing metric and form a long–short portfolio based on this metric. Stambaugh, Yu, and Yuan (forthcoming) argue that by aggregating the anomalies in this manner, a more precise measure of mispricing is obtained. The higher precision of this aggregate metric allows us to conduct a number of robustness and corroborative tests, which strengthen our conclusion related to the dumb and smart characterization of mutual funds and of hedge funds, respectively.

5.1. Contemporaneous relation between fund flows and mispricing

We begin by repeating the analysis in Tables 3 and 4 with a single, aggregate measure of cross-sectional mispricing. The results are shown in Table 5. In the first column of the table, our focus is on the long–short portfolio. In the second and third columns, we examine the LONG and SHORT components separately. Doing so allows us to determine if the effect of flows on aggregate mispricing operates through the purchase of overvalued stocks, through the sale of undervalued stocks, or through both types of transactions.

As expected, mutual fund flows are negatively and significantly related to the returns of the L–S portfolio formed based on the aggregate metric (-1.941, $t=-3.62$). This implies that mutual fund flows, in the aggregate, exacerbate mispricing in the cross section of US stocks. Looking at the LONG and SHORT components separately, mutual fund flows are unrelated to the LONG component (-0.247, $t=-0.60$) but are positively and significantly related to the SHORT component of the aggregate metric ($+1.693$, $t=+2.53$). This positive coefficient 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.

Our hypothesis is also supported in the case of hedge fund flows. This coefficient is significantly positive in the L–S regression. From the separate LONG and SHORT results, the hedge fund flow results are also driven exclusively by the SHORT leg, suggesting that flows from the hedge fund sector are invested primarily in the form of short positions in overvalued stocks.[13]

In the fourth, fifth, and sixth columns of Table 5, we repeat the analysis from the first three columns by focusing only on the subset of anomalies that do not belong to the real investment category. The fourth column shows the results obtained with the L–S portfolio. These are the same as those presented in the first column of Table 4 and are repeated here for completeness. In the fifth and sixth columns we examine the LONG and SHORT components for anomalies in the non-real investment category. Consistent with the results from the second and third columns, we find that the effect of flows on this subset of anomalies operates primarily through the SHORT leg. That is, mutual funds appear to take long positions, and hedge funds appear to take short positions, on stocks that are overvalued with respect to these particular anomalies.

Overall, the results in Table 5 suggest that aggregate mutual fund flows exacerbate cross-sectional mispricing, particularly so in the case of anomalies that do not belong to the real investment group. Moreover, this effect operates primarily through stocks in the SHORT leg of the mispricing metric, those that are the most overvalued. In other words, mutual fund flows contribute to cross-sectional mispricing through the purchase of overvalued stocks, not the sale of undervalued stocks.[14] In contrast, aggregate hedge fund flows appear to correct cross sectional mispricing mostly through short positions in overvalued stocks.[15] These results corroborate our previous conclusion that aggregate flows to mutual funds can qualify as dumb money, while aggregate flows to hedge funds appear to be smart money.[16]

[13] In untabulated results we include the betting against beta (BAB) factor of Frazzini and Pedersen (2014) as an additional control variable. Using the L–S portfolio as dependent variable, we find that the loadings on the MFFLOW and HFFLOW variables are quantitatively similar to those presented in Table 5. The results of Tables 3 and 4 are also not altered by the inclusion of the BAB factor.

[14] One important distinction between mutual funds and other institutions is that mutual funds are generally prohibited from shorting stocks. The evidence presented in Table 5 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.

[15] An asymmetry is noted in Table 2 in the exposure of the LONG and SHORT legs to the HML factor. We perform additional analyses to confirm that the results in Table 5 are not driven by this asymmetry. We examine if the inclusion of HML introduces a bias. We do so by repeating the tests in Table 5 without the HML factor. The results are qualitatively similar to those obtained in Table 5. For example, for the L-S specification, the coefficient estimates obtained are -1.840 ($t=-3.37$) for mutual fund flows and $+0.338$ ($t=+3.45$) for hedge fund flows. The coefficients in the Long and Short specifications are also similar to those obtained in Table 5.

[16] The large intercepts in Table 5 are due to the inclusion of aggregate turnover in our set of control variables. While most of our control

Table 5
Aggregate mutual fund flows, hedge fund flows, and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long (LONG), short (SHORT), or long minus short (L – S) return series of the cross-sectional mispricing metric and a combination of the individual anomalies that excludes the investment-based anomalies. The seven non-investment anomalies are return on assets, Ohlson's O-score, failure probability, gross profitability, net stock issues, total accruals, and momentum. The independent variables, measured contemporaneously with the dependent variable, 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 2.1. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	Mispricing metric			Non-investment anomalies		
Variable	L – S	LONG	SHORT	L – S	LONG	SHORT
MFFLOW	– 1.941	– 0.247	1.693	– 2.927	0.037	2.964
	– 3.62	– 0.60	2.53	– 3.33	0.10	2.96
HFFLOW	0.291	0.061	– 0.230	0.367	0.114	– 0.253
	2.72	0.59	– 2.02	3.29	1.69	– 2.00
RMRF	– 0.311	0.942	1.253	– 0.478	0.875	1.354
	– 5.07	16.21	21.43	– 6.92	18.41	17.56
AGGILLIQ	0.261	0.056	– 0.205	0.280	0.085	– 0.195
	2.76	0.67	– 2.80	2.63	0.95	– 1.85
AGGTURN	– 0.102	– 0.030	0.072	– 0.123	– 0.031	0.092
	– 2.84	– 1.17	1.85	– 3.25	– 1.46	2.48
HML	0.210	0.232	0.022	0.228	0.215	– 0.013
	1.68	2.98	0.18	1.60	2.98	– 0.09
SMB	– 0.109	0.450	0.560	– 0.100	0.430	0.530
	– 1.66	2.62	2.90	– 0.68	2.80	1.88
Intercept	0.028	0.011	– 0.017	0.030	0.010	– 0.020
	3.13	1.60	– 2.10	3.38	1.54	– 2.38
N	228	228	228	228	228	228
Adj. R^2	0.348	0.770	0.815	0.360	0.814	0.711

Though not critical to our main hypothesis, an interesting relation emerges between MFFLOW and HFFLOW in Table 5: The effects of mutual fund flows on cross-sectional mispricing appear to dominate those of hedge fund flows. In separate tests, we find that a one standard deviation move in mutual fund flows affects the L–S portfolios return by – 1.16% and a one standard deviation move in hedge fund flows affects the L–S return by +0.55%. We attribute this difference in magnitude to several sources. When compared with hedge funds, the assets under management (AUM) for mutual funds are approximately 15 times larger. Because flow measures generally represent a percentage change in net assets, we expect mutual fund flows to have a larger impact on prices. Further, while the mutual fund data set contains the universe AUM of that industry, the hedge fund data set is self-reported. If the manager's choice to include or exclude a hedge fund from the data set is correlated with the funds recent performance, we would expect our proxy for aggregate hedge fund flows to be noisier than the true aggregate measure.

Given the significant profitability of the L–S portfolio shown in Table 2, a relevant question is why smart money does not immediately take advantage of dumb money.[17] If it did, the predictability would be arbitraged away, and the alphas reported in Table 2 would be zero. The limits-to-arbitrage literature provides one possible explanation: market frictions could prevent hedge funds from completely eliminating mispricing. De Long, Shleifer, Summers, and Waldmann (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 can lead investors to withdraw capital and that arbitrageurs may reduce arbitrage intensity in anticipation of this possibility. Abreu and Brunnermeier (2002) show that arbitrage can be delayed as arbitrageurs reduce their capital intensity until a critical mass of capital is directed toward the mispricing. Together, this literature suggests that hedge funds might not be able to instantaneously erase the price effects of dumb money flows.

5.2. Separating aggregate inflows and outflows

We now separately examine aggregate mutual fund inflows and outflows (obtained from the N-SAR filings in the Security and Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system database. to determine if mispricing is primarily exacerbated by purchases or by sales transactions performed by mutual funds. To disaggregate the flows we start by matching N-

(*footnote continued*)
variables have means near zero, the mean of aggregate turnover is significantly positive. The inclusion of aggregate turnover simply shifts the intercept upward to account for the larger mean value and does not affect the interpretation of our empirical tests.

[17] This relation can 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 might not fully characterize aggregate flows.

Table 6

Aggregate mutual fund flows (N-SAR), hedge fund flows, and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long (LONG), short (SHORT), or long minus short (L−S) return series of the cross-sectional mispricing metric (Panel A), the 11 individual anomalies (Panel B), and combinations of the individual anomalies based on investment and non-investment groupings (Panel C). The seven non-investment anomalies are return on assets, Ohlson's o-score, failure probability, gross profitability, net stock issues, total accruals, and momentum. The four investment-based anomalies are investment-to-assets, asset growth, net operating assets, and composite equity issues. The independent variables, measured contemporaneous with the dependent variables, are four disaggregated measures of monthly mutual fund flows obtained from fund-level N-SAR filings (NEW INVESTMENT, REINVESTMENT, OTHER INFLOWS, and REDEMPTIONS), 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). The N-SAR flow data were downloaded from the Security and Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system database. We construct the inflow (NEW INVESTMENT, REINVESTMENT, and OTHER INFLOWS) and outflow (REDEMPTIONS) measures by matching, when available, monthly N-SAR flows with the set of CRSP mutual funds used to construct the measure of aggregate mutual fund flows (MFFLOW). Monthly dollar N-SAR flows are scaled by lagged total net assets (obtained from CRSP) of the matched funds. The coefficients on the lead variables are suppressed to conserve space. Details on the construction of the mispricing metric are provided in Section 2.1. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

Panel A: Mispricing metric

Variable	L−S	LONG	SHORT
NEW INVESTMENT	−1.140	0.656	1.796
	−2.57	0.85	1.94
REINVESTMENT	0.123	−0.064	−0.187
	0.42	−0.41	−0.71
OTHER INFLOWS	−4.374	27.163	31.537
	−0.32	2.23	2.70
REDEMPTIONS	−0.753	0.330	1.083
	−1.97	0.75	2.23
HFFLOW	0.259	0.018	−0.241
	2.27	0.18	−1.98
RMRF	−0.348	0.930	1.278
	−5.40	17.32	22.63
AGGILLIQ	0.229	−0.026	−0.256
	2.11	−0.23	−2.17
AGGTURN	−0.072	−0.022	0.050
	−1.96	−0.74	1.45
HML	0.206	0.221	0.015
	1.53	3.10	0.11
SMB	−0.132	0.447	0.579
	−1.92	2.85	3.43
Intercept	0.033	0.000	−0.034
	2.91	−0.02	−2.12
N	228	228	228
Adj. R^2	0.316	0.773	0.818

Panel B: Individual anomalies

	Return on assets (1)	Ohlson O-score (2)	Failure probability (3)	Gross profitability (4)	Net stock issues (5)	Total accruals (6)	Composite equity issues (7)	Investment to-assets (8)	Net operation assets (9)	Asset growth (10)	Momentum (11)
NEW INVESTMENT	−2.936	−2.082	−2.053	−0.804	−1.137	0.026	1.405	−0.023	0.567	1.027	−0.273
	−2.08	−2.26	−2.99	−1.18	−1.80	0.06	1.62	−0.09	0.89	1.62	−0.25
REINVESTMENT	−0.082	−0.070	0.092	−0.071	0.025	0.134	0.035	−0.079	0.186	0.121	1.034
	−0.31	−0.27	0.49	−0.58	0.18	0.76	0.11	−0.98	1.46	0.55	2.23
OTHER INFLOWS	1.632	−7.303	−4.341	21.643	−2.412	4.728	23.815	14.832	−17.756	−6.663	12.396
	0.09	−0.41	−0.25	1.68	−0.21	0.53	1.57	1.59	−1.42	−0.50	0.31
REDEMPTIONS	−1.365	−1.268	−1.015	−0.251	−0.355	−0.137	0.239	0.162	0.206	0.418	0.089
	−1.86	−2.13	−2.08	−0.62	−0.95	−0.46	0.39	0.83	0.47	1.16	0.10
HFFLOW	0.159	0.195	0.194	0.110	−0.038	0.062	−0.131	0.082	0.017	−0.038	0.463
	1.30	1.30	1.70	1.08	−0.46	0.98	−0.67	1.09	0.14	−0.39	1.37
RMRF	−0.381	−0.184	−0.378	−0.011	−0.370	0.039	0.407	−0.069	−0.191	−0.101	−0.317
	−4.99	−2.78	−5.16	−0.19	−5.49	0.94	4.28	−1.70	−3.32	−1.67	−1.96
AGGILLIQ	0.567	0.547	0.064	0.288	0.272	0.061	0.098	0.063	−0.012	−0.070	−0.305
	2.86	4.06	0.35	2.20	2.81	1.07	0.61	0.83	−0.10	−0.77	−1.37
AGGTURN	0.024	0.051	−0.088	0.064	−0.008	−0.024	0.050	−0.011	−0.077	−0.048	−0.219
	0.44	1.20	−1.67	1.98	−0.22	−1.13	0.72	−0.37	−1.70	−1.45	−2.56
HML	0.227	−0.032	−0.037	0.014	0.395	0.004	0.091	0.034	−0.096	0.346	0.021
	1.89	−0.44	−0.23	0.16	4.02	0.06	0.60	0.73	−1.18	3.26	0.07
SMB	−0.059	−0.023	−0.208	0.258	−0.073	−0.034	0.043	0.041	−0.326	−0.031	−0.393
	−0.31	−0.17	−2.55	3.33	−0.71	−0.65	0.32	1.04	−3.70	−0.44	−2.18
Intercept	0.032	0.005	0.065	0.002	0.026	−0.001	−0.038	0.005	0.008	−0.001	0.064
	1.45	0.33	4.11	0.21	2.13	−0.07	−2.42	0.56	0.54	−0.12	2.56

Table 6 (*continued*)

Panel B: Individual anomalies

	Return on assets (1)	Ohlson O-score (2)	Failure probability (3)	Gross profitability (4)	Net stock issues (5)	Total accruals (6)	Composite equity issues (7)	Investment to-assets (8)	Net operation assets (9)	Asset growth (10)	Momentum (11)
N	228	228	228	228	228	228	228	228	228	228	228
Adj. *R*2	0.212	0.094	0.278	0.066	0.498	0.004	0.159	0.027	0.180	0.169	0.135

Panel C: Anomaly combinations

	Non-investment anomalies			Investment anomalies		
Variable	L−S	LONG	SHORT	L−S	LONG	SHORT
NEW INVESTMENT	−2.834	0.703	3.538	1.204	2.003	0.800
	−2.75	1.18	2.39	1.34	1.56	1.38
REINVESTMENT	0.083	−0.035	−0.118	0.198	0.171	−0.028
	0.33	−0.24	−0.40	0.73	0.83	−0.19
OTHER INFLOWS	−0.037	27.484	27.521	13.016	29.680	16.664
	0.00	2.97	1.75	0.67	1.55	1.79
REDEMPTIONS	−1.713	0.213	1.926	0.719	1.145	0.426
	−3.12	0.66	2.64	1.13	1.54	1.31
HFFLOW	0.339	0.092	−0.247	0.027	−0.010	−0.037
	2.84	1.33	−1.85	0.18	−0.07	−0.31
RMRF	−0.501	0.874	1.375	0.091	1.180	1.089
	−7.00	20.18	18.88	1.14	12.62	21.47
AGGILLIQ	0.339	0.013	−0.327	0.151	−0.071	−0.222
	2.18	0.13	−1.85	1.02	−0.42	−2.22
AGGTURN	−0.084	−0.033	0.052	0.024	0.019	−0.005
	−1.81	−1.48	1.20	0.45	0.43	−0.15
HML	0.231	0.209	−0.022	0.057	0.292	0.234
	1.55	3.20	−0.15	0.52	3.16	2.20
SMB	−0.140	0.436	0.576	−0.086	0.429	0.515
	−1.11	3.17	2.42	−0.58	1.71	3.69
Intercept	0.053	−0.005	−0.057	−0.020	−0.029	−0.009
	3.27	−0.43	−2.51	−1.13	−1.42	−0.69
N	228	228	228	228	228	228
Adj. R^2	0.351	0.821	0.718	0.008	0.635	0.845

SAR filings with the set of equity mutual funds used in our aggregate mutual fund flow measure. We compute for each month the aggregate dollar inflow and outflows by summing the fund-level flows across funds in the matched sample. We scale the dollar flows by the prior month's aggregate net assets for the matched sample. We then repeat the tests in Table 5, replacing the aggregate mutual fund flow variable with four deconstructed flow measures: three separate inflows for new investment, reinvestment, and other inflows, and one outflow measure for redemptions (signed as negative to preserve comparison with our main results). Our focus is on the coefficients on new investment and redemptions. The results are presented in Table 6.

In Panel A of Table 6, we focus on aggregate mispricing. The panel presents the relation between the flow measures and the return to the L–S strategy. The two coefficients of interest are those of new investment (−1.140, $t=-2.57$) and redemptions (−0.753, $t=-1.97$). When examining the LONG and SHORT legs separately, the results are driven by the SHORT leg of the mispricing strategy. The positive coefficient on new investments on the SHORT leg suggests that new money is invested into overvalued stocks, on the wrong side of cross-sectional mispricing, consistent with the theme in our paper. However, the positive coefficient on redemptions on the same SHORT leg suggests that, when facing redemptions, mutual fund managers tend to sell stocks on the right side of cross-sectional mispricing, those that are most overvalued. This does not necessarily imply that mutual fund managers are smart when it comes to sales. Instead, the effect could be caused by managers being restricted to selling the stocks held in their portfolio, which the inflow results suggest are overvalued. Comparing the magnitudes of the two coefficients, the mispricing induced by new investments is stronger than the correction induced by redemptions. This observation, along with the fact that the dollar volume of new investments is higher (in absolute value) than that of redemptions, explains why mutual fund flows, in the aggregate, have an exacerbating effect on mispricing: The new investment effect dominates the redemption effect, both in terms of volume and intensity.

In Panel B of Table 6, we repeat the analysis from Panel A using the long–short returns obtained with individual anomalies, and in Panel C we group anomalies into those that belong to the real investment category and all others. In both panels, our focus continues to be on the coefficients of new investments and redemptions. We conjecture from our previous results that the effect of mutual fund flows on aggregate mispricing operates primarily through new money invested into stocks that are overvalued in relation to growth and accrual. Thus, we expect to find negative coefficients on new investments for anomalies that belong to the growth and accrual categories. For these same anomalies, we also expect negative coefficients on redemptions (but of lower magnitude) and positive coefficients on hedge fund flows. The results in Panel B broadly support this conjecture (taking into account the higher noise of individual anomalies). The coefficients on new investment are negative for all five growth anomalies (Columns 1 through 5) and are significant in four of the five cases. As expected, the coefficients on redemptions are also negative and of lower magnitude than those of new investments. And, the coefficients of hedge fund flows are positive in four of the five cases.

The contrast between anomalies is more clearly illustrated in Panel C. The first three columns in Panel C combine anomalies related to growth, accrual, and momentum, and the last three columns combine the four real investment anomalies. As expected, for the first group of anomalies the coefficient on new investment is significantly negative (-2.834, $t=-2.75$) and the coefficient on redemptions is also significantly negative, but of lower magnitude (-1.713, $t=-3.12$). And, the coefficient on hedge fund flows remains significantly positive ($+0.339$, $t=+2.84$). Examining the LONG and SHORT legs separately, the effect of flows on mispricing continues to operate through the SHORT leg, the one that likely contains overvalued stocks. Turning now to the last three columns of Panel C, we find no relation between flows and the returns of anomalies that belong to the real investment category, consistent with the results presented in Table 4.

From the results of Table 6 we conclude that the deconstructed new investment component of mutual fund exacerbates cross-sectional mispricing. The most likely channel is one in which these new inflows are used to purchase stocks that are overvalued on the growth and accrual dimensions. The effect of hedge fund flows is opposite: These flows appear to attenuate mispricing through short positions in stocks that are overvalued on these same dimensions. Neither mutual fund flows nor hedge fund flows are related to the returns of anomalies that belong to the real investment category.

6. Robustness checks and corroborating evidence

We conduct a number of tests to assess the robustness of the relation between fund flows and aggregate mispricing. And, because our main hypothesis, that mutual fund flows tend to exacerbate mispricing while hedge funds flow tend to correct it, carries a number of additional implications that cannot be directly tested by the contemporaneous regression approach, we also perform a number of corroborative tests to examine the internal validity of our study. For expositional simplicity, we present these results with the aggregate measure of cross-sectional mispricing. However, we have verified that our results in this section are substantially similar for the group of non-investment anomalies.

6.1. Forward predictability and subsequent return reversal

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. If flows to mutual funds are dumb money, 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 future returns to the L–S strategy. In the case of hedge fund flows, if these are smart money that reduce mispricing, then we expect no relation between current flows and future L–S returns. If the price effect of hedge fund flows represents a correction toward fundamental value, it should not reverse in subsequent months. This contrasts with the price effect of mutual fund flows, which is expected to reverse when stock prices ultimately converge to fundamental value.

We seek to determine if the effects in Tables 5 and 6 reverse during the subsequent months. We do this by examining the relation between flows measured at month t and future returns of the same stocks that at month t were included in the mispricing metric. The future returns are measured, alternatively, over the subsequent one- and three-month periods.

The results are presented in Table 7. The specifications also include the control variables from Tables 3–6, which are now measured contemporaneously to the return measurement period, either over month $t+1$ or from month $t+1$ to $t+3$.

Panel A of Table 7 shows results obtained with the aggregate mutual fund flows, and Panel B shows results when mutual fund flows have been deconstructed into the several flow types using the N-SAR data. The first line of Panel A shows the relation between current MFFLOW and the future return of the mispricing metric. If mutual funds exacerbate cross-sectional mispricing and if the mispriced stocks experience a return reversal in the subsequent period, we expect a positive relation between current MFFLOW and future returns to the long–short strategy. The results support this conjecture. The coefficient on MFFLOW is significantly positive in the L–S regression, over both the one- and three-month future horizons.

Of particular interest are the results obtained separately with the LONG and SHORT legs of the strategy. We find mutual fund flows being disproportionately invested in overvalued stocks, creating temporary upward price pressure for these stocks, but we find no corresponding results for undervalued stocks. If so, we would expect to see a reversal in the price of overvalued stocks, but not in the price of undervalued stocks. Therefore, stocks in the short leg should revert in price during the period from $t+1$ to $t+3$, but no such reversal should occur for stocks in the long leg. Again, the results in Table 7 confirm this conjecture. As expected, we find no significant relation between current MFFLOWS and future returns to the LONG component of the mispricing metric. However, we find a significantly negative relation for the SHORT component. Because stocks in the SHORT component were likely overvalued when purchased by mutual funds at time t, the negative relation suggests that these stocks are now converging in price toward intrinsic value.

Table 7
Aggregate mutual fund flows, hedge fund flows, and future cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the future return series of the stocks included in the cross-sectional mispricing metric measured over the forward one-month ([$t+1$]) and three-month ([$t+1$, $t+3$]) windows. The independent variables in Panel A 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). Panel B repeats the analysis in Panel A replacing the aggregate mutual fund flow variable with four disaggregated measures of monthly mutual fund flows obtained from fund-level N-SAR filings (NEW INVESTMENT, REINVESTMENT, OTHER INFLOWS, and REDEMPTIONS). The N-SAR flow data were downloaded from the Security and Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system database. We construct the inflow (NEW INVESTMENT, REINVESTMENT, and OTHER INFLOWS) and outflow (REDEMPTIONS) measures by matching, when available, monthly N-SAR flows with the set of CRSP mutual funds used to construct the measure of aggregate mutual fund flows (MFFLOW). Monthly dollar N- SAR flows are scaled by lagged total net assets (obtained from CRSP) of the matched funds. The coefficients on the lead variables are suppressed to conserve space. Details on the construction of the mispricing metric are provided in Section 2.1. The t-statistics shown below the coefficient estimates are based on Newey- West standard errors.

Panel A: Future mispricing metric returns

	One-month forward return [$t+1$]			Three-month forward return [$t+1$, $t+3$]		
Variable	L−S	LONG	SHORT	L−S	LONG	SHORT
MFFLOW	2.105	0.540	−1.565	3.606	−0.329	−3.501
	3.28	1.53	−2.50	2.70	−0.31	−2.46
HFFLOW	−0.047	−0.243	−0.196	−0.265	−0.439	−0.209
	−0.42	−1.78	−1.31	−1.10	−1.77	−1.01
RMRF	−0.277	−0.124	0.153	−0.279	−0.234	0.066
	−5.52	−3.35	2.49	−2.37	−2.73	0.54
AGGILLIQ	0.619	0.378	−0.241	0.899	0.690	−0.416
	2.05	1.97	−0.92	1.44	2.14	−0.74
AGGTURN	0.019	0.037	0.019	−0.091	−0.103	−0.040
	0.34	0.81	0.36	−0.62	−0.90	−0.31
HML	−0.008	−0.040	−0.032	−0.100	−0.159	−0.062
	−0.12	−0.75	−0.47	−0.56	−1.22	−0.38
SMB	0.090	0.050	−0.040	0.230	0.099	−0.105
	1.08	0.85	−0.44	1.50	0.91	−0.69
Intercept	0.007	0.009	0.002	0.057	0.024	−0.036
	0.59	1.04	0.16	1.54	0.91	−1.23
N	227	227	227	225	225	225
Adj. R^2	0.370	0.794	0.808	0.402	0.799	0.813

Panel B: Future mispricing metric returns and N-SAR mutual fund flows

Variable	One-month forward return [$t+1$]			Three-month forward return [$t+1$, $t+3$]		
	L−S	LONG	SHORT	L−S	LONG	SHORT
NEW INVESTMENT	0.967	0.378	−0.589	3.657	−0.645	−3.994
	1.38	0.80	−0.93	2.55	−0.62	−3.42
REINVESTMENT	−0.551	−0.356	0.195	−0.530	−0.564	−0.227
	−2.44	−2.18	0.76	−1.18	−2.46	−0.70
OTHER INFLOWS	44.187	2.873	−41.313	63.086	11.483	−44.165
	2.12	0.29	−1.89	1.23	0.51	−0.91
REDEMPTIONS	0.612	0.072	−0.540	2.329	−0.325	−2.433
	1.15	0.19	−1.21	1.86	−0.42	−2.26
HFFLOW	−0.073	−0.241	−0.168	−0.388	−0.441	−0.129
	−0.66	−1.78	−1.27	−1.32	−1.73	−0.68
RMRF	−0.209	−0.107	0.102	−0.226	−0.224	0.035
	−4.18	−2.88	1.69	−1.74	−2.74	0.32
AGGILLIQ	0.617	0.302	−0.314	0.644	0.648	−0.285
	2.13	1.79	−1.16	0.90	2.03	−0.47
AGGTURN	−0.138	0.027	0.165	−0.206	0.004	0.138
	−1.46	0.52	1.46	−1.17	0.03	0.86
HML	0.001	−0.026	−0.028	0.008	−0.168	−0.141
	0.02	−0.48	−0.44	0.04	−1.40	−0.88
SMB	0.114	0.062	−0.052	0.255	0.066	−0.167
	1.28	0.83	−0.52	1.60	0.64	−1.21
Intercept	0.006	−0.015	−0.020	0.055	−0.055	−0.104
	0.48	−0.94	−1.06	1.08	−1.20	−1.74
N	227	227	227	225	225	225
Adj. R^2	0.363	0.801	0.814	0.408	0.809	0.826

Turning now to the second line of Panel A, we find no relation between current hedge fund flows and future returns to the mispricing metric. This corroborates our previous conclusion that the price effect induced by hedge fund flows at time t is one that corrects, not exacerbates, mispricing. Once mispricing is corrected at the end of

Table 8
Relation between aggregate fund flows and cross-sectional mispricing controlling for time variation in fund market share, 1994–2012.
Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L−S) return series of the mispricing metric. The independent variables are aggregate mutual fund flow, aggregate hedge fund flow, de-trended mutual fund market share (DTSHARE-MF), de-trended hedge fund market share (DTSHARE-HF), a time trend (TIME), and the indicated interaction variables. TIME is a linear time trend (monthly) scaled to be equal to zero in the first month and equal to one in the last month. Mutual (hedge) fund market share is computed by scaling beginning fund net assets by the market value of all stocks in the CRSP value-weighted index. DTSHARE-MF (DTSHARE-HF) represents the residuals from a regression of mutual (hedge) fund market share on the time trend. DTSHARE is computed as the residuals from a regression of combined mutual fund market share and hedge fund market share on the time trend variable. Control variables are 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.

Variable	L−S (1)	L−S (2)	L−S (3)	L−S (4)	L−S (5)
MFFLOW	−2.278	−1.811	−2.535	−1.608	−2.193
	−3.05	−3.75	−3.26	−3.16	−3.11
MFFLOW*TIME	−0.393		0.605		0.476
	−0.24		0.41		0.31
MFFLOW*DTSHARE−MF		−88.774	−99.713		
		−2.34	−2.53		
MFFLOW*DTSHARE				−89.236	−83.768
				−2.78	−2.59
HFFLOW	0.415	0.341	0.310	0.265	0.343
	1.65	2.71	1.17	2.13	1.35
HFFLOW*TIME	−0.243		−0.103		−0.191
	−0.85		−0.32		−0.62
HFFLOW*DTSHARE−HF		−0.707	−14.366		
		−0.03	−0.54		
HFFLOW*DTSHARE				11.001	13.055
				2.01	2.08
TIME	−0.043		−0.070		−0.041
	−2.09		−2.08		−2.24
DTSHARE−MF		0.738	0.513		
		2.11	1.30		
DTSHARE−HF		1.020	−0.603		
		1.41	−0.53		
DTSHARE				0.729	0.667
				2.78	2.35
RMRF	−0.274	−0.290	−0.256	−0.307	−0.280
	−4.20	−4.87	−3.75	−5.11	−4.23
AGGILLIQ	0.041	0.113	−0.215	0.101	−0.065
	0.24	1.11	−0.97	1.10	−0.43
AGGTURN	−0.010	−0.150	0.014	−0.128	−0.035
	−0.17	−3.85	0.18	−3.31	−0.66
HML	0.228	0.227	0.239	0.213	0.219
	1.77	1.87	1.95	1.73	1.70
SMB	−0.109	−0.109	−0.117	−0.101	−0.104
	−1.66	−1.64	−1.69	−1.51	−1.56
Intercept	0.047	0.040	0.067	0.038	0.054
	3.56	4.10	3.74	4.26	4.38
N	228	228	228	228	228
Adj. R^2	0.354	0.362	0.367	0.366	0.369

month t, we find no further predictable price action during months $t+1$ to $t+3$.

In Panel B of Table 7, we again disaggregate mutual fund flows into several flow types using the N-SAR data and the methodology from Table 6. Our initial focus is on the coefficients of new investment. Our main hypothesis implies that the mispricing induced by new investments at time t will correct itself in the subsequent one to three months. Hence, we expect positive coefficients on new investment in the two L–S regressions of Panel B. Moreover, if the mispricing induced at time t operates through the purchase of overvalued stocks, but not through the sale of undervalued stocks, we expect the coefficient on new investments to be negative in the SHORT regressions and insignificant in the LONG regressions.

Our results support this conjecture. The coefficient of new investments for the three month holding period is significantly positive in L–S regression (+3.657, $t=+2.55$) and significantly negative in the SHORT regression (−3.994, $t=-3.42$). For the one-month holding period, the results are similar, but weaker. The sign is preserved, but the results do not attain statistical significance, suggesting that the price convergence toward intrinsic value occurs throughout the longer, three-month period. The positive coefficient on redemptions suggests that part of the redemptions effect from Table 6 is related to temporary price pressure that reverses over the subsequent three-month period.

Overall, the results of Table 7 corroborate our previous conclusion that mutual fund flows are primarily dumb, in

Please cite this article as: Akbas, F., et al., Smart money, dumb money, and capital market anomalies. Journal of Financial Economics (2015), http://dx.doi.org/10.1016/j.jfineco.2015.07.003

Table 9
Relation between aggregate index and non-index mutual fund flows and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L − S) return series of the cross-sectional mispricing metric. The independent variables are aggregate mutual fund flow to index and non-index funds, aggregate hedge fund flow, de- trended mutual fund market share to index funds (DTSHARE-INDEX) and non-index funds (DTSHARE-NON-INDEX), de-trended ratio of index fund total net assets to non-index fund total net assets (DTASSETS-INDEX/NON-INDEX), a time trend (TIME), and the indicated interaction variables. TIME is a linear time trend (monthly) scaled to be equal to zero in the first month and equal to one in the last month. Mutual fund market share, computed separately for index and non-index funds, is determined by scaling the respective beginning fund net assets by the market value of all stocks in the CRSP value-weighted index. Control variables are 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.

Variable	L-S (1)	L-S (2)	L-S (3)	L-S (4)	L-S (5)
NON-INDEX MFFLOW	−1.670	−2.331	−1.782	−2.589	−2.370
	−2.91	−3.05	−3.65	−2.64	−2.33
NON-INDEX MFFLOW*TIME		0.288		1.171	0.855
		0.16		0.71	0.51
NON-INDEX MFFLOW*DTSHARE-NON-INDEX			−139.928	−125.115	−132.582
			−3.26	−3.10	−2.93
INDEX MFFLOW	−0.169	0.121	0.187	0.129	0.134
	−1.11	0.42	0.80	0.45	0.47
INDEX MFFLOW*TIME		−0.932		−3.149	−3.368
		−1.84		−0.94	−1.01
INDEX MFFLOW*DTSHARE-INDEX			−9.007	33.745	36.281
			−1.52	0.81	0.87
TIME		−0.044		−0.050	−0.176
		−2.03		−1.83	−1.12
DTSHARE-INDEX			−0.376	0.202	−0.836
			−0.69	0.27	−0.55
DTSHARE-NON-INDEX			0.827	0.743	1.047
			3.05	2.61	2.24
DTASSETS-INDEX/NON-INDEX					0.307
					0.79
HFFLOW	0.294	0.270	0.298	0.287	0.290
	2.68	2.47	2.47	2.40	2.42
RMRF	−0.316	−0.276	−0.280	−0.260	−0.261
	−5.18	−4.15	−4.74	−4.04	−4.03
AGGILLIQ	0.263	0.028	0.089	−0.178	−0.173
	2.70	0.17	0.65	−0.77	−0.75
AGGTURN	−0.113	−0.009	−0.122	−0.029	−0.024
	−3.29	−0.15	−2.98	−0.46	−0.38
HML	0.206	0.239	0.244	0.249	0.244
	1.61	1.87	2.07	2.13	2.09
SMB	−0.114	−0.105	−0.106	−0.119	−0.126
	−1.72	−1.62	−1.62	−1.76	−1.91
Intercept	0.028	0.047	0.035	0.061	0.055
	3.33	3.47	3.44	3.32	2.91
N	228	228	228	228	228
Adj. R^2	0.341	0.350	0.372	0.375	0.374

that new money is invested on the wrong side of anomalies at time t, leading to a contemporaneous increase in the price of overvalued stocks and a subsequent price reversal for these same stocks during $t+1$, $t+3$. The results also support our conclusion that hedge fund flows are primarily smart money. The absence of price action during the $t+1$, $t+3$ period confirms that the price effect of hedge fund flows at time t is one of correction, not exacerbation, of cross-sectional mispricing.

6.2. Changes in aggregate mispricing rank

One implication of our findings in Table 5 is that aggregated mispricing scores should change during month t in a manner that is consistent with exacerbation (for mutual funds) and correction (for hedge funds) of cross-sectional mispricing. However, while the momentum, net issuance, and composite equity factors change on a monthly basis, the other factors change only at an annual frequency. This said, not all firms have the same fiscal year end, and differences in the fiscal year-end across sample firms could also introduce some slight month-to-month variation in these other eight factors, in the aggregate.

Keeping these limitations in mind, we perform a simple test to see if contemporaneous changes in mispricing scores are consistent with the main inference from Table 5. We compute the correlation between fund flows and changes in the mispricing scores of underlying stocks forming the long and short legs of the mispricing metric. The results (untabulated for brevity) corroborate our previous findings. By convention, overvalued stocks (those in the short portfolio) have low scores and undervalued stocks have high scores. The correlations show that when mutual fund flows are high, stocks in the undervalued long portfolio become even more undervalued ($\rho=+0.1811$, p-value$=0.0032$) and stocks in the overvalued short portfolio become more overvalued ($\rho=-0.1729$, p-value$=0.0049$). As before, the opposite result is obtained with hedge funds. When hedge fund flows are high, stocks in the undervalued portfolio become less undervalued ($\rho=-0.1201$, p-value$=0.0710$) and stocks in the overvalued portfolio become less overvalued ($\rho=+0.1080$, p-value$=0.1046$).

6.3. *The role of funds' share in equity markets*

Over the past few decades, the size of both mutual and hedge fund industries has grown significantly and the role of fund flows in financial markets has become increasingly important. If the main theme in our paper is valid, we should observe that fund flows have a larger effect on mispricing (both exacerbating and corrective) when funds represent a larger share of the market. To test this conjecture, we compute the industry market share for funds as the aggregate assets under management of mutual funds and hedge funds, scaled by the prior month-end aggregate market capitalization of stocks in the CRSP universe. We expect higher industry market shares to be associated with a stronger influence of funds in the equity market.

Also, a strong time trend is present in our measures of funds' market share. To differentiate between the effects of time and changes in market share, we de-trend our market share variables using a monthly linear time trend variable (TIME) standardized to take values between zero and one. The residuals obtained from the first stage regressions, DTSHARE-MF, DTSHARE-HF, and DTSHARE, represent the inter-temporal evolution of mutual fund market share, hedge fund market share, and combined market share, orthogonal to the time trend.[18] The results are presented in Table 8.

The analysis is similar to Table 5, except that we now add the de-trended market share variables, a time trend variable, and their interactions with mutual fund flows and with hedge fund flows. The first column of Table 8 shows the results without the market share variable. The coefficient estimate on mutual fund flows remains negative and significant, and the coefficient on hedge fund flows is positive and significant. The coefficient on time trend is negative and significant, suggesting that aggregate cross-sectional mispricing decreases over time, consistent with an increase in cross-sectional market efficiency across our sample. The interaction coefficients are insignificant, suggesting that the mere passage of time does not change the effect of flows on mispricing during the main sample period from 1994 to 2012.

In the second column of Table 8, we replace the time trend variable with de-trended mutual fund and hedge fund market shares (DTSHARE-MF and DTSHARE-HF). Once again, the coefficients on mutual and hedge fund flows are significant and carry the expected sign. Our main interest is on the coefficients of the interaction terms. In the case of mutual funds, the interaction coefficient is negative and significant, suggesting a stronger effect on aggregate mispricing in response to higher market shares. In the case of hedge funds the interaction coefficient is not significant. We believe that this lack of significance is due to the self-reporting bias in the hedge fund database, in that the level of aggregate AUM computed from the database might not be representative of the AUM for the entire hedge fund sector. For example, we observe that the trend in market share of hedge funds in our sample after 2008 falls to levels of the mid-1990s.[19] Also of interest is the coefficient on the (non-interacted) DTSHARE-MF variable. This coefficient is positive and significant, suggesting that higher past mutual fund flows (evidenced by larger market share) lead to higher return predictability during the current period. In the third column, we add the time trend variable to the second-column specification. The main inferences are similar to those obtained from the first two columns.

In the fourth and fifth columns, to reduce the influence of the self-reporting bias in the hedge fund database, we combine the total net assets of the mutual fund and hedge fund sample to compute a proxy for the market share of the fund industry, DTSHARE. We then repeat the analysis from the second and third columns using the de-trended industry market share variable (DTSHARE). Once again, the coefficients on mutual fund flows and hedge fund flows are significant and carry the expected signs. Our main focus is on the coefficients of the interaction terms. The coefficient on the interaction between mutual fund flows and total market share is negative and significant. For hedge fund flows, the interaction coefficient is now positive and significant. These results suggest that the effect of fund flows on cross sectional mispricing increases with funds' market share. When the relative size of the fund industry is larger, the exacerbating effect of mutual fund flows and the corrective effects of hedge fund flows are both stronger.

Also of particular interest is the coefficient of the time trend variable, which is significantly negative, and the coefficient of market share (DTSHARE), which is significantly positive. We interpret this asymmetry as follows: The passage of time has two countervailing effects. The first is a secular decrease in cross-sectional mispricing, as evidenced by the negative coefficient on time. The second is an exacerbation of cross-sectional mispricing due to an

[18] We also explore including higher order time trend terms (up to quartic) and found that they did not materially affect the results in Table 8 (and in Table 9). The higher order terms were largely insignificant.

[19] After the 2007–2008 financial crisis, a sizable number of hedge funds in our data sample stopped reporting, resulting in the downward trend in market share we observe at the end of our sample. The number of hedge funds in our sample decreased from approximately 1,800 funds at the end of 2007 to approximately 900 funds at the end of 2012. Self-reporting is less likely to bias our measure of hedge fund flows when compared with the bias it induces in the aggregate AUM measure. This is because our monthly flow measure is computed fund by fund and is scaled by fund-specific AUM measured at the end of the previous month. When a fund stops reporting during a certain month, both the numerator and the denominator of the flow measure are affected. This countervailing effect minimizes the impact of the self-reporting bias. By contrast, there is no countervailing effect when measuring the level of aggregate AUM.

increase in DTSHARE.[20] This could provide an interesting explanation on the persistence of cross-sectional return predictability despite the large number of hedge funds that seek to trade on it. Predictability persists because it is consistently fueled by flows channeled through mutual funds into stocks that are overvalued.

Overall the results in Table 8 show that the exacerbating effect of mutual fund flows on mispricing is more potent when the market share of funds is higher. For hedge funds, the corrective effect of flows on mispricing also increases with the overall market share of funds. Moreover, during our sample period from 1994 to 2012, the two effects are driven by changes in market share instead of by the mere passage of time.

6.4. Extending the sample period using ICI data

Another way to corroborate the industry market share effect is to examine an out-of-sample period when the mutual fund industry played a more subdued role in financial markets. Data on aggregate mutual fund assets dating back to January 1984 are available from The Investment Company Institute (ICI). This allows for a limited extension of our main results to the out-of-sample period from 1984 to 1994.[21] To preserve comparison with our main findings, we first replicate the analysis during the main sample period from 1994 to 2012. We then conduct a separate analysis for the earlier period from 1984 to 1993. The results (untabulated) confirm that during our sample period from 1994 to 2012 the coefficient on ICI mutual fund flow is negative (-7.193) and statistically significant ($t=-2.23$). As expected, we find no relation between mutual fund flows and aggregate mispricing during the earlier (1984 to 1993) period. The coefficient on ICI flows is small ($+0.37$) and statistically insignificant ($t=+0.18$). We attribute this lack of significance to the fact that the mutual fund industry was much smaller during that earlier period. The ICI tests, therefore, corroborate our earlier conclusion that the exacerbating effect of mutual fund flows is a relatively recent phenomenon.

6.5. Index versus non-index mutual fund flows

Thus far, in the paper we have measured aggregate mutual fund flows using all equity funds from the CRSP Survivor-Bias-Free US Mutual Fund Database. If the dumb money nature of mutual fund flows is caused by investors chasing stocks with certain characteristics, we expect to find stronger effects in flows to non-index funds when compared with flows that are invested in passive index funds. Moreover, the impact of non-index funds should be higher when the market share of the mutual fund industry is larger. To test this conjecture we repeat the analysis of Table 8 by disaggregating mutual fund flows into index and non-index funds. The results are presented in Table 9.

The first column in Table 9 contains the main results without any interaction variables. As conjectured, the coefficient of non-index flows is significantly negative ($t=-2.91$) and of relatively large magnitude (-1.670). By contrast, the coefficient on index flows is statistically insignificant ($t=-1.11$) and of smaller magnitude (-0.169).

In the next three columns, we examine the role of time and industry trends on the results obtained in the first column. As in Table 8, we de-trend the market share variables to disentangle the time trend effect from the industry trend effect. We begin, in the second column, by adding the time trend variable and its interactions with the flow variables. As before, the time trend effect is negative. However, time does not seem to affect the relation between flows and aggregate mispricing. (Although the interaction between time and index fund flows is significant, this result is not robust to the specifications in the last two columns.)

In the third column, we include the de-trended market shares of index and non-index funds, as well as their interactions with fund flows. Of main interest are the two interaction coefficients. The coefficient on the interaction between non-index funds share and non-index fund flows is negative and highly significant ($t=-3.26$). This finding corroborates our earlier findings in Table 8 that the impact of mutual fund flows on aggregate mispricing is stronger during periods of higher market share. From Table 9, non-index mutual funds are the main driver of this intertemporal relation. In contrast, the variable that interacts index fund flows with index fund market share carries an insignificant coefficient. Also of interest is the positive and significant relation between non-index mutual fund market share and aggregate cross sectional mispricing. We find a similar relation for the aggregate fund market share in Table 8. The results in Table 9 suggest that this relation is driven primarily by non-index funds. In the fourth column we add a time trend and its interaction with the two flow variables. The results are as expected, and there is no independent effect of time on aggregate mispricing.

Finally, in the last column we augment the specification in the fourth column with the de-trended ratio of total net assets under management of index funds to total net assets of non-index funds (DTASSETS-INDEX/NON-INDEX). The estimated coefficient for this term is insignificant, suggesting that the ratio of index to non-index assets does not independently affect aggregate mispricing.

Overall, our results in Table 9 support our conjecture that flows to non-index mutual funds are the main drivers of exacerbation in cross-sectional mispricing. Moreover, this relation is stronger following periods of high market shares of non-index funds.

6.6. Retail versus institutional mutual fund flows

Frazzini and Lamont (2008) suggest that retail investors are the main sources of dumb money in the mutual funds market. Therefore, we expect fund flows from retail investors to have a stronger effect on aggregate mispricing

[20] Given that the size of the mutual fund industry significantly outweighs that of the hedge fund industry, the net effect of intertemporal changes in the market share of managed funds is one of exacerbation of mispricing. If hedge funds were dominant, the net effect would be one of correction, rather than not exacerbation.

[21] Although the ICI data are available for a longer period, we continue to use the CRSP data in the rest of the paper, not only because CRSP is a more established data source, but also because the ICI data, being too general, cannot be used to conduct some of the more detailed tests we performed.

Table 10

Aggregate mutual fund flows, hedge fund flows, state variables, economic conditions, and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L − S) return series of the cross-sectional mispricing metric. The key independent variables are the flow variables, MFFLOW and HFFLOW, economic state variables, and the interaction of the state variables with the indicated flow variables. Economic state variables are a National Bureau of Economic Research recession dummy (NBER), a financial crisis dummy (FIN CRISIS), a run-up of the Internet bubble dummy (INTERNET), an Internet bubble crash dummy (INTERNET CRASH), change in employment (CHG EMPLOY), change in industrial production (CHG INDPRO), change in consumer durables (CHG CONSDUR), change in consumer nondurables (CHG CONSNON), and change in consumer services (CHG CONSERV). Control variables are monthly excess market returns (RMRF), aggregate illiquidity (AGGILLIQ), and aggregate turnover (AGGTURN). Details on the construction of the mispricing metric are provided in Section 2.1. The *t*-statistics shown below the coefficient estimates are based on Newey-West standard errors.

	State variables								
Variable	NBER	FINANCIAL CRISIS (7/07 to 12/09)	INTERNET BUBBLE (1/99 to 3/00)	INTERNET CRASH (4/00 to 12/03)	CHG EMPLOY	CHG INDPRO	CHG CONSDUR	CHG CONSNON	CHG CONSSERV
MFFLOW	−1.723	−1.987	−1.709	−1.342	−2.194	−1.951	−1.984	−2.200	−1.799
	−2.99	−3.48	−3.54	−2.31	−3.56	−3.42	−3.65	−4.02	−2.45
HFFLOW	0.359	0.372	0.282	0.201	0.305	0.292	0.259	0.273	0.337
	2.77	2.74	2.57	2.37	2.69	2.47	2.23	2.62	3.00
STATE	0.004	0.010	0.034	0.003	−1.436	−0.172	−0.071	−0.251	−0.783
	0.63	1.76	1.77	0.40	−0.81	−0.57	−0.75	−1.11	−1.10
MF*STATE	−2.659	−1.179	−5.936	−2.357	282.215	16.711	15.152	58.573	−15.942
	−1.76	−0.79	−2.11	−1.43	1.65	0.36	1.13	1.65	−0.15
HF*STATE	−0.252	−0.270	−0.117	0.595	21.469	−1.932	4.885	6.229	−9.887
	−1.28	−1.50	−0.21	1.92	0.64	−0.24	0.91	1.30	−0.33
RMRF	−0.296	−0.298	−0.308	−0.296	−0.298	−0.314	−0.307	−0.304	−0.312
	−4.85	−4.91	−5.06	−4.97	−4.91	−4.86	−4.92	−5.18	−4.99
AGGILLIQ	0.214	0.199	0.248	0.222	0.197	0.255	0.243	0.244	0.249
	2.23	2.00	2.61	2.39	1.91	2.62	2.59	2.67	2.61
AGGTURN	−0.119	−0.147	−0.091	−0.084	−0.127	−0.103	−0.109	−0.116	−0.111
	−3.62	−3.50	−2.54	−2.28	−2.97	−2.89	−3.21	−3.67	−2.91
HML	0.228	0.222	0.241	0.208	0.227	0.209	0.196	0.202	0.203
	1.77	1.77	1.90	1.72	1.73	1.65	1.50	1.57	1.63
SMB	−0.095	−0.107	−0.148	−0.124	−0.105	−0.111	−0.117	−0.100	−0.114
	−1.39	−1.60	−1.83	−2.09	−1.65	−1.64	−1.79	−1.43	−1.70
Intercept	0.031	0.035	0.026	0.025	0.034	0.029	0.030	0.032	0.033
	3.66	3.65	2.87	2.69	3.20	3.22	3.40	3.87	3.27
N	228	228	228	228	228	228	228	228	228
Adj. R^2	0.355	0.347	0.350	0.357	0.347	0.339	0.343	0.347	0.342

when compared with fund flows from institutional investors. To test this conjecture we disaggregate mutual fund flows into flows to retail share classes and flows to institutional share classes and repeat the analysis of Table 5. The results (untabulated) show a significantly negative relation between aggregate mispricing and retail mutual fund flows ($t = -3.22$). While institutional flows are also negatively related to the aggregate mispricing, the coefficient estimate is smaller and statistically insignificant ($t = -1.36$). These results suggest that the dumb nature of mutual fund flows is ascribed primarily to flows originating from individual investors. These are the flows most likely to be invested on the wrong side of mispricing anomalies and, therefore, are most likely to exacerbate cross-sectional mispricing.

6.7. *Impact of economic conditions*

Because investors could alter their behavior as a result of changes in economic conditions, we extend our analysis to examine the role of time-varying economic conditions on the relation between fund flows and aggregate mispricing. We repeat the analysis of Table 5 by including a number of economic state variables (STATE) and their interactions with our flow measures. The state variables we use in our analysis are an National Bureau of Economic Research recession dummy (NBER), a financial crisis dummy (FIN CRISIS), a dummy that captures the price run-up period leading to the internet bubble (INTERNET), a dummy that captures the market correction following the internet bubble (INTERNET CRASH), change in employment (CHG EMPLOYMENT), change in industrial production (CHG INDPRO), change in consumer durables (CHG CONSDUR), change in consumer nondurables (CHG CONSNON), and change in consumer services (CHG CONSERV). The main variables of interest are the two terms that interact the economic state variable with mutual fund flows and with the hedge fund flows. These terms explain how the effect of flows on aggregate mispricing changes as a result of time-varying economic conditions. The results are presented in Table 10. Two findings are particularly noteworthy.

First, in the third column of Table 10 we observe a significantly negative coefficient on the interaction between mutual fund flows and a dummy for the Internet run-up period (January 1999–March 2000). This coefficient, combined with the baseline coefficient on mutual fund flows (also significantly negative) suggests that while flows to mutual funds exacerbate cross-sectional mispricing throughout our sample period, the Internet price run-up period was one when this effect was particularly potent.

Please cite this article as: Akbas, F., et al., Smart money, dumb money, and capital market anomalies. Journal of Financial Economics (2015), http://dx.doi.org/10.1016/j.jfineco.2015.07.003

Second, in the fourth column of Table 10, we find a significantly positive coefficient on the interaction between hedge fund flows and a dummy for the Internet run-down period (April 2000–December 2003). This coefficient, combined with the baseline coefficient on hedge fund flows (also significantly positive), implies that hedge funds mitigate cross-sectional mispricing throughout the sample period, particularly so during the market correction that followed the Internet run-up period.

Overall, the results in Table 10 demonstrate that the relation between fund flows and aggregate mispricing is not time-varying except for the period surrounding the Internet bubble. Moreover, the results that mutual funds exacerbate mispricing during the Internet run-up while hedge funds correct mispricing during the run-down conform with common intuition and further corroborate the main theme in our paper.

6.8. De-trended and orthogonalized fund flows

One possible concern is that our results are contaminated by the presence of time trends in mutual and hedge fund flows, because the total dollar amount invested in both types of funds has increased during our sample period. Our measure of monthly fund flows imputes new dollar flows as a percentage of net total assets. However, to completely rule out this concern, we repeat our base analysis using de-trended fund flows. To construct the de-trended fund flow variables, we regress aggregate mutual fund flows and, in turn, aggregate hedge fund flows, on Legendre polynomials (up to quartic) of a time trend and retain the residuals.[22]

Another possible concern is that flow variables are highly correlated with market returns, with aggregate illiquidity, or with turnover. To mitigate 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 use these orthogonalized measures in our regressions.

Table 11 presents the results of the de-trended and orthogonalized analyses. The first two columns in Table 11 show that de-trended results are even stronger than those presented in Table 5. The de-trended MFFLOW variable is negatively correlated with the long–short strategy, with t-statistics ranging from -3.77 to -5.37. And, the de-trended HFFLOW variable maintains a positive and significant relation with the L–S strategy, with t-statistics ranging from $+2.88$ to $+3.66$. We conclude that replacing our flow variables with the de-trended series does not materially change our conclusion.

The last two columns in Table 11 show the results of the orthogonalized analysis. We observe a strong negative relation between orthogonalized MFFLOW and contemporaneous returns to the long–short strategy. For orthogonalized HFFLOW, this relation is positive, as expected. Thus, replacing our aggregate flow variables with orthogonalized flows does not materially affect our conclusion.

6.9. Predicted versus residual fund flows

Both mutual fund flows 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 the unexpected component 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 on contemporaneous measures of the following control variables: excess market return, aggregate illiquidity, implied volatility of the S&P 500 index, and the six macroeconomic variables of growth in industrial production, growth in consumer consumption of durables, growth in consumer consumption of nondurables, and growth in consumer consumption of services, growth in employment, and an NBER recession dummy.

The results are presented in Table 12. The table repeats our base analysis from Table 5, replacing the flow variables with their predicted and residual components. Predicted variables are generally insignificant, while the coefficients 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 instead by periods when flows unexpectedly deviate from the trend level.[23]

Fig. 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 L–S returns during months when residual fund flows are high. Panel A presents the results for residual mutual fund flows. Confirming the results of Table 12, the average L–S return is high when residual mutual fund flows are low, with a difference between the Q5 and Q1 portfolios equal to -1.98% per month (p-value$=0.02$). Panel B presents the results for residual hedge fund flows. As in Table 12, the average L–S return is high when residual hedge fund flows are high, with a difference between the Q5 and Q1 portfolios equal to $+2.49\%$ 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 vector autoregression (VAR) using the long–short strategy returns, the residual mutual fund flows, and the residual hedge fund flows. We then obtain the innovations from

[22] In an untabulated analysis, we find that each of the higher order time trend terms (up to quartic) is significantly related to fund flows. Therefore, we include all four time trend terms in our regression specification. Higher order terms are not significant.

[23] We also repeat the analysis in Table 12, augmenting the set of independent variables in the predictive regressions with the investment sentiment measure of Baker and Wurgler (2006), available on Jeffrey Wurgler's website (http://people.stern.nyu.edu/jwurgler). Our results (untabulated) are not materially different from those presented in Table 12.

Please cite this article as: Akbas, F., et al., Smart money, dumb money, and capital market anomalies. Journal of Financial Economics (2015), http://dx.doi.org/10.1016/j.jfineco.2015.07.003

Table 11
De-trended aggregate mutual fund and hedge fund flows and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which the dependent variable is the monthly long minus short (L−S) return series of the cross-sectional mispricing metric. In Panel A, the independent variables of interest are de-trended aggregate mutual fund flow (DT MFFLOW) and de-trended aggregate hedge fund flow (DT HFFLOW). The de-trended flow variables are computed as the residuals obtained by regressing flows on Legendre polynomials (up to quartic) of a time trend. In Panel B, the independent variables of interest are aggregate mutual fund flow orthogonalized against the contemporaneous values of the other independent variables (ORTH MFFLOW) and aggregate hedge fund flow orthogonalized against the contemporaneous values of the other independent variables (ORTH HFFLOW). Control variables are 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.

Variable	De-trended flows		Orthogonalized flows	
	Model 1 L−S	Model 2 L−S	Model 1 L−S	Model 2 L−S
DT MFFLOW	−3.878	−2.434		
	−5.37	−3.77		
DT HFFLOW	0.370	0.275		
	3.66	2.88		
ORTH MFFLOW			−1.703	−1.755
			−2.76	−3.14
ORTH HFFLOW			0.286	0.211
			2.62	1.88
RMRF		−0.295		−0.377
		−4.52		−6.60
AGGILLIQ		0.039		0.050
		0.41		0.55
AGGTURN		−0.083		−0.086
		−2.70		−2.76
HML		0.236		0.210
		1.88		1.68
SMB		−0.105		−0.109
		−1.62		−1.66
Intercept	0.019	0.030	0.020	0.031
	6.41	3.70	6.41	3.79
N	228	228	228	228
Adj. R^2	0.156	0.357	0.045	0.348

Table 12
Predicted versus residual components of aggregate fund flows and cross-sectional mispricing, 1994–2012.

Shown below are coefficient estimates of time series regressions in which 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 MFFLOW and RES MFFLOW) and predicted and residual aggregate hedge fund flow (PRED MFFLOW and RES HFFLOW). 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 Standard & Poor's S&P 500 index (VIX), and six macroeconomic variables including growth in industrial production, growth in consumer consumption of durables, growth in consumer consumption of non-durables, and growth in consumer consumption of services, growth in employment, and a National Bureau of Economic Research recession dummy variable. Control variables are 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.

Variable	Model 1 L−S	Model 2 L−S
PRED MFFLOW	−1.174	−0.568
	−1.83	−0.66
RES MFFLOW	−2.525	−2.817
	−3.31	−4.18
PRED HFFLOW	0.111	−0.020
	0.46	−0.10
RES HFFLOW	0.530	0.434
	2.43	2.87
RMRF		−0.375
		−5.15
AGGILLIQ		0.084
		0.53
AGGTURN		−0.110
		−2.81
HML		0.204
		1.68
SMB		−0.101
		−1.52
Intercept	0.022	0.035
	4.32	2.91
N	216	216
Adj. R^2	0.061	0.367

the VAR model for each series and examine their contemporaneous correlations. Consistent with the results of Table 12, we find that innovations in the long–short return series are negatively correlated with innovations in mutual fund flows (−0.157, p-value=0.0214), and positively correlated with innovations in hedge fund flows (+0.154, p-value=0.0243).

6.10. Implied volatility and sentiment

We now 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 possibility, 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 orthogonalized against a set of macroeconomic variables.[24] Both measures are computed at the time of portfolio formation.

The results are presented in Table 13. We repeat the analysis of Table 5, including VIX and investor sentiment as additional variables. We also include interactions between these two variables and the two flow measures. We find that the coefficient on the interaction of VIX with mutual fund flows is negative and significant. This suggests that the effects of mutual fund flows on aggregate mispricing are higher during high VIX periods. Interestingly, the coefficient on the interaction of sentiment with mutual fund flows is also significantly negative.

The fact that both interaction coefficients (VIX and sentiment) are significantly negative is somewhat puzzling. VIX captures periods of both intense optimism and

[24] Our results are robust to using the alternative University of Michigan measure of consumer sentiment. Results are available upon request.

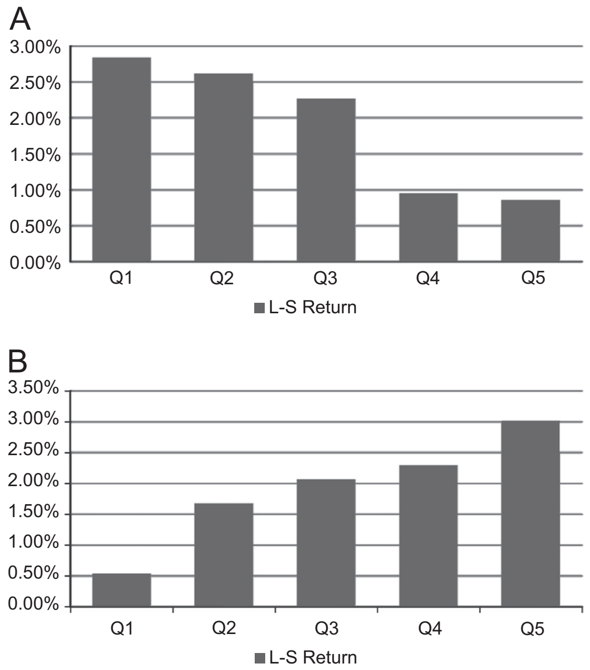

Fig. 1. Residual fund flows and long minus short (L–S) strategy performance. Fig. 1 presents the equal-weighted average long minus short (L–S) return of the cross sectional mispricing metric for quintile portfolios formed using residual fund flows. Q1 represents the average monthly L–S returns when residual 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 (growth in industrial production, growth in consumer consumption of durables, growth in consumer consumption of nondurables, growth in consumer consumption of services, growth in employment, and an NBER recession dummy variable. Panel A: Residual mutual fund flow portfolios and Panel B: Residual hedge fund flow portfolios.

intense pessimism while sentiment is strictly proportional to optimism. Conceptually, VIX can be viewed as an absolute value of sentiment, in the sense that VIX is high when sentiment is either high (such as during the late 1990s) or low (such as late 2008). A simple correlation test (not shown) confirms that the economic relation between VIX and sentiment resembles the hypothesized absolute value pattern.[25] This suggests that the VIX results in Table 13 could be driven primarily by periods of unusually high optimism. These are periods when sentiment and VIX are both high. We perform two additional (untabulated) tests to confirm this conjecture.

First, we orthogonalize standardized VIX on standardized sentiment. Conceptually, the residuals from this regression correspond to periods of unusually high pessimism (high VIX and low sentiment numbers). By contrast, the fitted values correspond to periods of unusually high optimism (high sentiment periods). We then interact the residuals and fitted values with the mutual fund flow and with the hedge fund flow variables. If our conjecture is correct, we expect the interaction between fitted VIX and mutual fund flow to be significantly negative. The interaction between residual VIX and mutual fund flows should be closer to zero. That is, we hypothesize that VIX exacerbates the effect of flows on mispricing mainly during periods of intense optimism, captured by the fitted VIX. The results support this conjecture: The coefficient on the interaction between fitted VIX and mutual fund flows is negative and significant ($t=-2.36$). By contrast, the coefficient on the interaction between residual VIX and mutual fund flows is smaller in magnitude and insignificant ($t=-1.62$), suggesting that VIX has no effect on the relation between flows and mispricing during periods of intense pessimism.

Second, we disaggregate the standardized sentiment variable into an optimistic part and a pessimistic part. The optimistic part is defined as zero when standardized sentiment is negative. Otherwise, it is equal to the actual value of standardized sentiment. The pessimistic part is equal to zero when standardized sentiment is positive. Otherwise, it is equal to the absolute value of standardized sentiment. We repeat the analysis in Table 13 replacing sentiment with its optimistic and pessimistic components. The two interaction terms are replaced with four new interaction terms, corresponding to combinations of optimistic versus pessimistic sentiment and mutual fund flows versus hedge fund flows. The results (untabulated) confirm our prior conjecture: The only significant interaction coefficient is the one that corresponds to the interaction between mutual fund flows and the optimistic component of sentiment ($t=-2.82$). Overall, these tests, together with the results in Table 13, confirm that mutual fund flows are more likely to exacerbate cross-sectional mispricing in periods of intense investor optimism.

In the case of hedge fund flows we find no relation between flows and either VIX or investor sentiment. Moreover, none of the interaction terms involving hedge fund flows attain significance in the two additional tests we conducted. Thus, the relation between hedge fund flows and aggregate mispricing seems to be unrelated to investor optimism or pessimism.

6.11. *Longer predictive horizons*

To assess the robustness of the results presented in Table 7, we perform additional (untabulated) analyses to examine the effects of flows on future returns over windows of three, six, nine, and 12 months. Consistent with a reversal phenomenon resulting from the price pressure hypothesis, mutual funds maintain the correct (positive) sign for all long-term horizons examined, although the level of statistical significance varies. The results are significant at the three-month and nine-month

[25] When we correlate a standardized VIX measure with a standardized sentiment measure (both with mean zero), we obtain only a weak positive correlation that is not statistically significant ($\rho=0.114$, p-value$=0.104$). However, when we correlate (standardized) VIX with the absolute value of (standardized) sentiment, the correlation is much stronger and significantly positive ($\rho=0.207$, p-value$=0.003$).

Table 13
Relation between aggregate fund flows and cross-sectional mispricing controlling for implied volatility and investor sentiment, 1994–2012.
Shown below are coefficient estimates of time series regressions in which 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 Standard & Poor's S&P 500 index (VIX), and the Baker and Wurgler (2006) measure of investor sentiment (SENTORTH, available from 1994 to 2010) as well as the corresponding interaction variables. 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 are 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.

Variable	Implied volatility (VIX) L−S			Sentiment (SENTORTH) L−S		
MFFLOW	−1.601	−1.878	−1.839	−1.614	−1.957	−1.645
	−2.57	−3.32	−3.05	−2.90	−3.40	−3.33
MFFLOW*VIX			−11.559			
			−2.74			
MFFLOW*SENTORTH						−2.381
						−2.19
HFFLOW	0.293	0.306	0.315	0.321	0.270	0.263
	2.49	2.70	2.69	2.63	1.83	1.91
HFFLOW*VIX			−0.505			
			−0.64			
HFFLOW*SENTORTH						−0.160
						−0.33
VIX	−0.023	0.016	0.042			
	−0.57	0.56	1.42			
SENTORTH				0.018	0.006	0.015
				2.51	0.86	2.10
RMRF		−0.315	−0.295		−0.292	−0.264
		−5.03	−4.91		−4.45	−4.19
AGGILLIQ		0.239	0.116		0.222	0.217
		2.81	1.15		2.32	2.15
AGGTURN		−0.108	−0.140		−0.098	−0.086
		−3.22	−3.43		−2.36	−2.31
HML		0.212	0.236		0.206	0.182
		1.70	1.86		1.49	1.50
SMB		−0.108	−0.114		−0.105	−0.118
		−1.62	−1.70		−1.55	−1.77
Intercept	0.028	0.026	0.037	0.022	0.029	0.026
	2.59	2.61	3.84	6.01	2.78	2.76
N	228	228	228	205	205	205
Adj. R^2	0.045	0.346	0.354	0.120	0.346	0.364

horizons and are close to significance at the six-month horizon ($t = +1.27$).

The coefficients on hedge fund flows never attain statistical significance when the dependent variable is measured over longer predictive horizons, consistent with our earlier conclusion that hedge fund flows serve primarily to correct contemporaneous mispricing.

7. Summary and conclusion

Using mutual and hedge fund flows as proxies for smart and dumb money, respectively, we show their impacts on cross-sectional equity return anomalies. At the aggregate level, mutual fund flows appear to exacerbate mispricing in the cross section of 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, 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) and of Lou (2012) and that aggregate flows to hedge funds are better suited for the smart money label we introduce in our paper. Our research has not explored implications of our findings for aggregate investor welfare. For example, while there could be a wealth transfer from naïve investors to hedge funds, the economy as a whole could 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 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 (2008) show that stocks with a high probability of failure have lower future returns.

Ohlson 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 under-perform the stocks of nonissuers.

Composite equity issuance: Daniel and Titman (2006) show that firms with higher equity issuance under-perform those with lower measures. Composite Equity issues increase with SEOs and share-based acquisitions and decrease with share repurchases and dividends.

Total Accruals: Sloan (1996) shows that stocks with high accruals under-perform stocks with low accruals.

Net operating assets: Hirshleifer, Hou, Teoh, and Zhang (2004) show that stocks with higher net operating assets under-perform those with lower net operating assets.

Momentum: Jegadeesh and Titman (1993) show that stocks with higher past performance are shown to out-perform stocks with lower past perform.

Gross profitability: Novy-Marx (2013) 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, D., Brunnermeier, M.K., 2002. Synchronization risk and delayed arbitrage. Journal of Financial Economics 66, 341–360.
Akbas, F., Armstrong, W.J., Sorescu, S., Subrahmanyam, A., 2015. Capital market efficiency and arbitrage efficacy. Journal of Financial and Quantitative Analysis. forthcoming.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross section of stock returns. Journal of Finance 61, 1645–1680.
Ben-Rephael, A., Kandel, S., Wohl, A., 2012. Measuring investor sentiment with mutual fund flows. Journal of Financial Economics 104, 363–382.
Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491–527.
Brown, N.C., Wei, K.D., Wermers, R., 2014. Analyst recommendations, mutual fund herding, and overreaction in stock prices. Management Science 60, 1–20.
Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899–2939.
Cao, C., Chen, Y., Liang, B., Lo, A.W., 2013. Can hedge funds time market liquidity? Journal of Financial Economics 109, 493–516.
Cao, J., Han, B., 2010. Idiosyncratic Risk, Costly arbitrage, and cross section of stock returns. Unpublished working paper. The University of Texas at Austin.
Chen, L., Novy-Marx, R., Zhang, L., 2010. An alternative three-factor model. Unpublished working paper.
Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross section of stock returns. Journal of Finance 63, 1609–1652.
Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics 86, 479–512.
Daniel, K.D., Titman, S., 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605–1643.
De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.
DeVault, L., Sias, R., Starks, L., 2014. Who are the sentiment traders? Evidence from the cross section of stock returns and demand. Unpublished working paper. University of Texas, Austin, TX.
Desai, H., Rajgopal, S., Venkatachalam, M., 2004. Value-glamour and accruals mispricing: one anomaly or two? Accounting Review 79 (2), 355–385.
Edelen, R., Ince, O., Kadlec, G.B., 2014. Institutional investors and stock return anomalies. Unpublished working paper. University of California, Davis, CA.
Edelen, R.M., Warner, J.B., 2001. Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. Journal of Financial Economics 59, 195–220.
Ederington, L., Golubeva, E., 2011. The impact of stock market volatility expectations on investor behavior: evidence from aggregate mutual fund flows. Unpublished working paper. University of Oklahoma, Norman, OK.
Edwards, F.R., Zhang, X., 1998. Mutual funds and stock and bond market stability. Journal of Financial Services Research 13, 257–282.
Falkenstein, E.G., 1996. Preference for stock characteristics as revealed by mutual fund portfolio holdings. Journal of Finance 51, 111–135.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 82, 491–518.
Fant, L.F., 1999. Investor behavior of mutual fund shareholders: the evidence from aggregate fund flows. Journal of Financial Markets 2, 391–402.
Frazzini, A., Lamont, O.A., 2008. Dumb money: mutual Fund flows and the cross section of stock returns. Journal of Financial Economics 88, 299–322.
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.
Fung, W., Hsieh, D.A., 2000. Performance characteristics of hedge funds and commodity funds: natural vs. spurious biases. Journal of Financial and Quantitative Analysis 35, 291–307.
Gruber, M.J., 1996. Another puzzle: the growth in actively managed mutual funds. Journal of Finance 51 (3), 783–810.
Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.
Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650–705.
Huang, J., Sialm, C., Zhang, H., 2011. Risk shifting and mutual fund performance. Review of Financial Studies 24, 2575–2616.
Jagannathan, R., Malakhov, A., Novikov, D., 2010. Do hot hands exist among hedge fund managers? An empirical evaluation. Journal of Finance 65, 217–255.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for market efficiency. Journal of Finance 48, 65–91.
Keswani, A., Stolin, D., 2008. Which money is smart? Mutual fund buys and sells of individual and institutional investors. Journal of Finance 63 (1), 85–118.
Kokkonen, J., Suominen, M., 2014. Hedge funds and stock market efficiency. Unpublished working paper. Aalto University, Helsinki, Finland.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.
Lou, D., 2012. A flow-based explanation for return predictability. Review of Financial Studies 25, 3457–3489.
Loughran, T., Ritter, J.R., 1995. The new issues puzzle. Journal of Finance 50, 23–51.
McLean, D., Pontiff, J., 2013. Does academic research destroy stock return predictability? Unpublished working paper. Boston College, Boston, MA.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. Journal of Financial Economics 108, 1–28.
Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18, 109–131.
Ritter, J.R., 1991. The long-run performance of initial public offerings. Journal of Finance 46, 327.
Shive, S., Yun, H., 2013. Are mutual funds sitting ducks? Journal of Financial Economics 107, 220–237.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. Journal of Finance 52 (1), 35–55.
Sirri, E.R., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.
Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Accounting Review 71, 289–315.
Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.

Stambaugh, R.F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. Journal of Finance. 10.1111/jofi.12286, forthcoming.

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

Warther, V.A., 1995. Aggregate mutual fund flows and security returns. Journal of Financial Economics 39, 209–235.

Wermers, R., 2003. Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Unpublished working paper. University of Maryland, College Park, MD.

Zheng, L., 1999. Is money smart? A study of mutual fund investors' fund selection ability. Journal of Finance 54 (3), 901–933.