Dumb money: Mutual fund flows and the cross-section of stock returns

Andrea Frazzini
Department of Economics, Yale University

Owen A. Lamont
Yale School of Management and NBER

PRELIMINARY AND INCOMPLETE

This draft: March 27, 2005
First draft: February 23, 2005

JEL Classification: G14, G23, G32
Key words: Mutual funds, individual investors

We thank Nicholas Barberis and Judith Chevalier for helpful comments. We thank Breno Schmidt for research assistance.

ABSTRACT

We use mutual fund flows as a measure for individual investor sentiment for different stocks, and find that high sentiment predicts low future returns. Fund flows are dumb money – by reallocating across different mutual funds, retail investors reduce their wealth in the long run. This dumb money effect is strongly positively related to the value effect. High sentiment also is associated high corporate issuance, interpretable as companies increasing the supply of shares in response to investor demand.

Individual retail investors actively reallocate their money across different mutual funds. Individuals tend to transfer money from low performing funds to high performing funds. In addition to looking at past returns of funds, individuals also may consider economic themes or investment styles in reallocating funds. Collectively, one can measure individual sentiment by looking at which funds receive inflows and which receive outflows, and can relate this sentiment to different stocks by examining the holdings of mutual funds. This paper tests whether sentiment affects stock prices, and specifically whether one can predict future stock returns using a flow-based measure of sentiment. If sentiment pushes stock prices above fundamental value, high sentiment stocks should have low future returns.

For example, in 1999 investors sent $36 billion to Janus funds but only $20 billion to Fidelity funds, despite the fact that Fidelity had more than three times the assets under management at the beginning of the year. Thus in 1999 retail investors as a group made an active allocation decision to give greater weight to Janus funds, and in doing so they increased their portfolio weight in tech stocks held by Janus. By 2001, investors had changed their minds about their allocations, and pulled about $12 billion out of Janus while adding $10 billion to Fidelity. In this instance, the reallocation caused wealth destruction to mutual fund investors as Janus and tech stocks performed horribly after 1999.

According to the "smart money" hypothesis of Gruber (1996) and Zheng (1999), some fund managers have skill and some individual investors can detect that skill, and send their money to skilled managers. Thus (in contrast to the Janus example) flows should be positively correlated with future returns. Gruber (1996) and Zheng (1999) show that the short term performance of funds that experience inflows is significantly better than those that experience outflows, suggesting that mutual fund investors have selection ability.[1]

Our focus is on stocks, not on funds. We are interested in how investor sentiment affects stocks prices, and see fund flows as a convenient (and economically important) measure of sentiment. To test whether investor sentiment causes mispricing, one must test whether high sentiment today predicts low return in the future, and we focus on cross-sectional stock return predictability over periods of months and years. We ask the question of whether, over the long-term, investors are earning higher returns as a result of their reallocation across funds.

For each stock, we calculate the mutual fund ownership of the stock that is due to reallocation decisions reflected in fund flows. For example, in December 1999, 17% of the shares outstanding of Cisco were owned by the mutual fund sector (using our sample of funds), of which 2.5% was attributable to disproportionately high inflows over the previous 3 years. That is, under certain assumptions, if flows had occurred proportionately to asset value (instead of disproportionately to funds like Janus), the level of mutual fund ownership would have been only 14.5%. This 2.5% difference is our measure of investor sentiment. We then test whether this measure predicts differential returns on stocks.

Our main results are as follows. First, as suggested the example of Janus and Cisco in 1999, on average from 1980 to 2003, retail investors direct their money to funds which invest in stocks that have low future returns. To achieve high returns, it is best to do the opposite of these investors. We calculate that mutual fund investors experience total returns that are significantly lower due to their reallocations. Therefore, mutual fund investors are dumb money in the sense that their reallocations reduce their wealth on average. We call this predictability the "dumb money" effect. This dumb money effect poses a challenge to rational theories of fund flows.

Second, the dumb money effect is highly related to the value effect. The returns on portfolios constructed using our flow-based measure of sentiment are quite positively correlated with the returns on portfolios constructed using market-book ratio. Money flows into mutual funds that own growth stocks, and flows out of mutual funds that own value stocks. This pattern poses a challenge to risk-based theories of the value effect, which would need to explain why one class of investors (individuals) is engaged in a complex dynamic trading strategy of selling “high risk” value stocks and buying “low risk” growth stocks.

Third, demand by individuals and supply from firms are highly related. When individuals indirectly buy more stock of a specific company (via mutual fund inflows), we also observe that company increasing the number of shares outstanding (for example, through seasoned equity offerings, stock-financed mergers, and other issuance mechanisms). This pattern is consistent with the interpretation that individual investors are dumb, and smart firms are opportunistically exploiting their demand for shares.

These results give a different perspective on the issue of individuals vs. institutions. A large literature explores whether institutions have better average performance than individuals. In the case of mutual funds, for example, Daniel, Grinblatt, Titman, and Wermers (1997) show that stocks held by mutual funds have higher returns, and Chen, Jegadeesh, and Wermers (2000) show that stocks bought by mutual funds outperform stocks sold by mutual funds. Both results suggest mutual fund managers have stock-picking skill.

Unfortunately, since individuals ultimately control fund managers, it can be difficult to infer the views of fund managers by looking only at their holdings. For example, when the manager of tech fund experiences large inflows, his job is to buy more technology stocks, even if he thinks the tech sector is overvalued. So if we observe the mutual fund sector as a whole holding technology stocks, that does not imply that mutual managers as a whole believe tech stocks will outperform. It is hard for a fund manager to be smarter than his clients. Mutual fund holdings are driven by both managerial choices in picking stocks and retail investor choices in picking managers. We provide some estimates of the relative importance of these two effects.

This paper is organized as follows. Section I reviews the literature. Section II discusses the basic measure of sentiment and describes the data. Section III presents regression results on the determinants of sentiment and the relation between sentiment and future returns. Section IV uses calendar time portfolios to put the results in economic context, showing the magnitude of wealth destruction caused by flows, comparing the sentiment measure with other well-known strategies, and providing evidence on whether mutual fund managers have stock-picking skill. Section V presents conclusions.

I. Background and literature review

A. Determinants of fund flows

A series of papers have documented a strong positive relation between mutual fund past performance and subsequent fund inflows (see, for example, Ippolito (1992), Chevalier and Ellison (1997), and Sirri and Tufano (1998)). In addition, retail investors appear to allocate their wealth to funds that have caught their attention either thought marketing or advertising (see Jain and Wu (2000), and Barber, Odean and Zheng (2004)). Benartzi and Thaler (2001) report evidence that retail investors employ simple rule-of–thumbs in allocating across different types of mutual funds.

For individual stocks, the picture looks different. Odean (1999), and Barber and Odean (2000, 2001, 2004) present extensive evidence that individual investors suffer from biased-self attribution, and tend be overconfident, thus engaging in (wealth-destroying) excessive trading. But in contrast to their return-chasing behavior in mutual funds, a variety of recent evidence suggests that individual investors act as contrarians when trading individual stocks (see Grinblatt and Keloharju (2000), Goetzmann and Massa (2002)).

While this apparent contradiction between return-chasing and contrarianism is interesting, the hypothesis we wish to test does not depend on resolving this issue. We are interested in testing whether individual investor sentiment predicts future returns, so our hypothesis is not contingent on measuring whether investors are ultimately return-chasing or not. For example, if individual investor sentiment causes prices to be wrong and prices eventually revert to fundamental value, then sentiment should negatively predict future returns no matter what – whether individuals over-react or under-react, whether they return-chase or not. As it turns out, in the data we study, mutual fund flows are indeed return-chasing, and flows tend to go to stocks that have gone up recently.

B. Causal effects of flows on prices

There is evidence that fund flows have positively contemporaneous correlations with stock returns (see, for example, Brown et al (2002)). Although it is difficult to infer causality from correlation, one interpretation of this fact is that inflows drive up stock prices. We do not attempt to test this hypothesis with our data, for three reasons. First, we are primarily interested in whether sentiment causes long-term mispricing, not the short term dynamics of precisely how trading affects prices. Second, we observe flows and holdings and fairly low frequency (quarterly), so our data is not well suited to studying short-term price dynamics. Third, although the fund flows we consider are certainly economically large, we view them as an imperfect measure of sentiment since individual investor sentiment can be manifested in many other ways. While individuals were sending mutual fund money to tech funds in 1999, and thus indirectly purchasing tech stocks, they may have also been buying tech stocks directly in their brokerage accounts, or investing in hedge funds that bought tech stocks.

Thus the hypothesis we wish to test is that stocks owned by funds with big inflows are overpriced. These stocks could be overpriced because inflows force mutual funds to buy more shares and thus push stock prices higher, or they could be overpriced because overall demand (not just from mutual fund inflows) pushes stock prices higher. In either case, inflows reflect the types of stocks with high investor demand.

C. Styles

A paper closely related to ours is Teo and Woo (2001), who also find evidence for a dumb money effect. Following Barberis and Shleifer (2003), Teo and Woo (2001) consider categorical thinking by mutual fund investors along the dimensions of large/small or value/growth. They show that when a particular category has large inflows, stocks in that category subsequently underperform. Like us, they relate mutual fund flows to stock returns, but unlike us they look only at style returns, not individual stock returns.

While Teo and Woo (2001) provide valuable and convincing evidence, our approach is more general. The benefit is that we do not have to define specific styles or categories, such as value/growth. While categorical thinking and style classification are undoubtedly important in determining fund flows, from a practical point of view it is difficult for the researcher to identify all relevant categories used by investors over time. For example, the growth/value category was not widely used in 1980. Instead, we impose no categorical structure on the data and just follow the flows. Most strikingly, we are able to document that the fund flow effect is highly related to the value effect, a finding that could not have been discovered using the method of Teo and Woo (2001).

II. Constructing the flow variable

Previous research has focused on different ownership levels, such as mutual fund ownership as a fraction of shares outstanding (for example, Chen, Jegadeesh, and Wermers, 2000). We want to devise a measure that is similar, but is based on flows. Specifically, we want to take mutual fund ownership and decompose it into the portion due to flows and the portion not due to flows. By "flows," we mean flows from one fund to another fund (not flows in and out of the entire mutual fund sector).

Our central variable is FLOW, the percent of the shares of a given stock owned by mutual funds that are attributable to fund flows. This variable is defined as the actual ownership by mutual funds minus the ownership that would have occurred if every fund had received identical proportional inflows (instead of experiencing different inflows and outflows), every fund manager chose the same portfolio weights in different stocks as he actually did, and stock prices were the same as they actually were. We define the precise formula later, but the following example shows the basic idea.

Suppose at quarter 0, the entire mutual fund sector consists of two funds: a technology fund with \$20 B in assets and a value fund with \$80 B. Suppose at quarter 1, the technology fund has an inflow of \$11 B and has capital gains of \$9 B (bringing its total assets to \$40 B), while the value fund has an outflow of \$1 B and capital gains of \$1 B (so that its assets remain constant). Suppose that in quarter 1 we observe the technology fund has 10% of its assets in Cisco, while the value fund has no shares of Cisco. Thus in quarter 1, the mutual fund sector as a whole owns \$4 B in Cisco. If Cisco has \$16 B in market capitalization in quarter 1, the entire mutual fund sector owns 25% of Cisco.

We now construct a world where investors simply allocate flows in proportion to initial fund asset value. Since in quarter 0 the total mutual fund sector has \$100 B in assets and the total inflow is \$10 B, the counterfactual assumption is that all funds get an inflow equal to 10%

of their initial asset value. To simplify, we assume that the flows all occur at the end of the quarter (thus the capital gains earned by the funds are not affected by these inflows). Thus in the counterfactual world the technology fund would receive (.20)*(10) = $2 B (giving it total assets of $31 B), while the value fund would receive (.80)*(10) = $8 B (giving it total assets of $89). In the counterfactual world the total investment in CISCO is given by (.1)*(31) = $3.1, which is 19.4% of its market capitalization. Hence, the FLOW for CISCO, the percent ownership of Cisco due to the non-proportional allocation of flows to mutual funds, is 25 – 19.4 = 5.6%.

FLOW is an indicator of what types of stocks are owned by funds experiencing big inflows. It is a number that can be positive, as in this example, or negative (if the stock is owned by funds experiencing outflows or lower-than-average inflows). It reflects the active reallocation decisions by investors. What FLOW does not measure is the amount of stock that is purchased with inflows; one cannot infer from this example that the technology fund necessarily used its inflows to buy Cisco. To the contrary, our assumption in constructing the counterfactual is that mutual fund managers choose their percent allocation to different stocks in a way that is independent of inflows and outflows.

Is it reasonable to assume that managers choose their portfolio weights across stocks without regard to inflows? Obviously, there are many frictions (for example, taxes and transaction costs) that would cause mutual funds to change their stock portfolio weights in different stocks in response to different inflows. Thus, we view FLOW as an imperfect measure of demand for stocks due to retail sentiment.

In equilibrium, of course, a world with different flows would also be a world with different stock prices, so once cannot interpret the counterfactual world as an implementable alternative for the aggregate mutual fund sector. Later, when we discuss the effects of flows on investor wealth, we consider an individual investor (who is too small to affect prices by himself) who behaves like the aggregate investor. We test whether this individual representative investor benefits from the active reallocation decision implicit in fund flows. For individual investors, refraining from active reallocation is an implementable strategy.

A. Flows

We calculate mutual fund flows using the CRSP US Mutual Fund Database. The universe of mutual funds we study includes all domestic equity funds that exists at any date between 1980 and 2003 for which quarterly net asset values (NAV) are available and for which we can match CRSP data with the common stock holdings data from Thomson Financial (described in the next subsection). Since we do not observe flows directly, we infer flows from fund return and net asset value (NAV) as reported by CRSP. Let N_t^i be the total NAV of a fund i and let R_t^i be its return between quarter $t-1$ and quarter t. Following the standard practice in the literature (e.g. Zheng (1999), Sapp and Tiwari (2004)), we compute flows for fund i in quarter t, F_t^i, as the dollar value of net new issues and redemptions using

$$F_t^i = N_t^i - (1+R_t^i)\cdot N_{t-1}^i - MGN_t^i \qquad (1)$$

where MGN is the increase in total net assets due to mergers during quarter t. Note that (1) implicitly assumes that inflows and outflows occur at the end of the quarter, and that existing investors reinvest dividends and other distributions in the fund. We assume that investors in the merged funds place their money in the surviving fund. Funds that are born have inflows equal to their initial NAV, while funds that die have outflows equal to their terminal NAV.

Counterfactual flows are computed under the assumption that each fund receives a pro rata share of the total dollar flows to the mutual fund sector between date $t-k$ and date t, with the proportion depending on NAV as of quarter t-k. More precisely, in order to compute the

FLOW ownership at date t, we start by looking at the net asset value of the fund at date $t-k$. Then, for every date s we track the evolution of the fund's counterfactual NAV using:

$$\hat{\mathrm{F}}_{\mathrm{s}}^{\mathrm{i}} = \frac{N_{t-k}^{i}}{N_{t-k}^{Agg}} \mathrm{F}_{\mathrm{s}}^{\mathrm{Agg}} \qquad (2)$$

$$\hat{\mathrm{N}}_{\mathrm{s}}^{\mathrm{i}} = (1+R_t^i)\hat{\mathrm{N}}_{\mathrm{s-1}}^{\mathrm{i}} + \hat{\mathrm{F}}_{\mathrm{s}}^{\mathrm{i}} \qquad (3)$$

$$t-k \le s \le t$$

where $\hat{\mathrm{F}}^{\mathrm{i}}$ and $\hat{\mathrm{N}}^{\mathrm{i}}$ are counterfactual flows and NAV's. $\mathrm{F}^{\mathrm{Agg}}$ is the actual aggregate flows for the entire mutual fund sector, while $\mathrm{N}_{\mathrm{t-k}}^{\mathrm{Agg}}$ is the actual aggregate NAV at date $t-k$. Equations (2) and (3) describe the dynamics of funds that exist both in quarter *t-k* and in quarter *t*. For funds that were newly created in the past *k* quarters, $\hat{\mathrm{N}}^{\mathrm{i}}$ is automatically zero – all new funds by definition represent new flows. The resulting counterfactual net asset value $\hat{\mathrm{N}}_{\mathrm{t}}^{\mathrm{i}}$ at date t represents the fund size in a world with proportional flows in the last k quarters.

For a detailed numerical example of our counterfactual calculations, see the appendix (which also discusses adjustments to equations (2) and (3) in the case of funds that die). We obtain a quarterly time series of counterfactual net asset values for every fund by repeating the counterfactual exercise every quarter t, and storing the resulting $\hat{\mathrm{N}}^{\mathrm{i}}{}_t$ at the end of each rolling window.

Consider a representative investor who represents a tiny fraction, call it *q*, of the mutual fund sector. Suppose this investor behaves exactly like the aggregate of mutual investors, sending flows in and out of different funds at different times. The counterfactual strategy described above is an alternative strategy for this investor, and is implementable using the same information and approximately the same amount of trading by the investor. To implement this strategy, this investor only needs to know lagged fund NAV's and aggregate flows. For this investor, $q\hat{N}_t^i$ is his dollar holding in any particular fund.

In designing this strategy, our aim is to create a neutral alternative to active reallocation, which matches the total flows to the mutual fund sector. One could describe this strategy as a more passive, value-weighting alternative to the active reallocation strategy pursued by the aggregate investor. It is similar in spirit to the techniques of Daniel, Grinblatt, Titman, and Wermers (1999) and Odean (1999) in that it compares the alternative of active trading to a more passive strategy based on lagged asset holdings. A feature of our counterfactual calculations is that they do not mechanically depend on the actual performance of the funds. A simpler strategy would have been to simply hold funds in proportion to their lagged NAV. The problem with this strategy is that it mechanically tends to sell funds with high returns and buy funds with low returns. Since we wanted to devise a strategy that reflected only flow decisions by investors (not return patterns in stocks), we did not used this simpler strategy.

Let x_{it} be the net asset value of fund i in month t as a percentage of total asset of the mutual fund sector:

$$x_{it} = \frac{N_t^i}{N_t^{Agg}} \qquad (4)$$

The counterfactual under proportional flows is:

$$\hat{x}_{it} = \frac{\hat{N}_t^i}{N_t^{Agg}} \qquad (5)$$

The difference between x_{it} and $\hat{x}_{it}$ reflects the active decisions of investors to reallocate money from one manager to another over the past k quarters in a way that is not proportional to the NAV of the funds. This difference reflects any deviation from value weighting by the NAV of the fund in marking new contributions. In theory, this difference could reflect rebalancing away from high performing funds and into poorly performing funds, in order to maintain some fixed weights (instead of market weights). In practice, investors tend to unbalance (not rebalance), sending money from poorly performing funds to high performing funds.

B. Holdings

Thomson Financial provides the CDA/Spectrum mutual funds database, which includes all registered domestic mutual funds filing with the SEC. The data show holdings of individual funds collected via fund prospectuses and SEC N30D filings. The holdings constitute almost all the equity holdings of the fund (see Appendix for a few small exceptions). The holdings data in this study run from January 1980 to December 2003.

Most funds report their holdings quarterly, although the SEC requires mutual funds to disclose their holdings on a semi-annual basis. Approximately 60% of the funds report quarterly holdings, with the rest semiannual. Although reports may be made on any day, the last day of the quarter is most commonly the report day. A typical fund-quarter-stock observation would be as follows: as of March 30th, 1998, Fidelity Magellan owned 20,000 shares of IBM. The holdings data are notably error-ridden, with obvious typographical errors (sometimes involving transposed digits and misplaced decimal points). Furthermore, some reports are missing from the database. We use a series of filters to eliminate data errors and to handle missing reports (see appendix).

In matching the holdings data to the CRSP mutual fund database, we utilized fund tickers, fund names and total net asset values. Our matching system works better in the latter part of the sample: coverage of the dollar assets of the total CRSP universe of funds rises from about 64% in 1980 to 96% in 2003 (in future version of this paper, we hope to obtain more accurate matching data from WRDS). For each fund and each quarter, we calculate w_{ij} as the portfolio weight of fund i in stock j based on the latest available holdings data.[2] Hence the portfolios weights w_{ij} reflect fluctuations of the market price of the security held.

Let z be the actual percent of the shares outstanding held by the mutual fund sector,

$$z_j = \left(\sum_i x_i \cdot w_{ij} \cdot N_t^{Agg} \right) / MKTCAP_j \qquad (6)$$

where $\mathrm{MKTCAP_j}$ is the market capitalization of firm j. The ownership that would have occurred with proportional flows into all funds and unchanged fund stock allocation and stock prices would be

$$\hat{z}_j = \left(\sum_i \hat{x}_i \cdot w_{ij} \cdot N_t^{Agg} \right) / MKTCAP_j \qquad (7)$$

For each stock, we calculate our central variable, FLOW, as the percent of the shares outstanding with mutual fund ownership attributable to flows. The flow of security j is given by

$$FLOW_{j,t} = z_{j,t} - \hat{z}_{j,t} = \left\{ \sum_i \left[x_{i,t} - \hat{x}_{i,t} \right] \cdot w_{ij} \cdot N_t^{Agg} \right\} / MKTCAP_{j,t} \qquad (8)$$

This flow has the following interpretation. If each portfolio manager had made exactly the same decisions in terms of percent allocation of his total assets to different stocks, and if stock prices were unchanged, but the dollars had flown to each portfolio manager in proportion to their NAV for the last k periods, then mutual fund ownership in stock j would be lower by FLOW. Stocks with high FLOW are stocks that are owned by mutual funds that have experienced high inflows.[3]

In this paper, we focus on a three year horizon in calculating FLOW. Since our goal is to understand the long-term effects on investor wealth, the longer the horizon, the better. Since our sample is less that 25 years long, three years is approaching the longest horizon that is appropriate given data limitations. Three years is also the approximate frequency of the value effect or reversal effect in stock returns.

We first describe the data for funds. Table I shows the top and bottom funds in December 1999, ranked on the difference between actual fraction of the fund universe (*x*) and counterfactual fraction ($\hat{x}_{it}$). In 1999, the Magellan fund has assets that constituted 3.5% of our sample mutual fund universe, but had been receiving below average inflows over the past three years. Had Magellan received flows in proportion to its size over the previous three years, it would have been 4.8% of the universe instead of 3.5%. The table shows that in 1999, the funds receiving big inflows tended to be technology and growth funds.

Figure 1 shows the experience of some of the largest funds in our sample. We show the actual fraction of the mutual fund universe, and the counterfactual fraction using a three year horizon. As expected, the counterfactual percent ownership tends to mirror the actual ownership with a three year lag. In interpreting these graphs, bear in mind that the actual level (*x*) of the fund reflects two things: the fund's performance, and its inflows and outflows. The variable we are interested in is the difference between the actual level and the counterfactual level.

Figure 1 shows that Magellan was a highly successful fund attracting inflows through the 1980's, then subsequently faded and did not receive proportional inflows. The Vanguard 500 fund steadily attracted more inflows as indexing grew more popular over the sample period. Apparently index funds experienced a sharp drop in popularity around the year 2000. Note that index funds are appropriate for the purposes of our study. If investor sentiment favors index funds, then one would expect stocks in the index to be overpriced relative to other stocks, and there is some evidence from the index inclusion literature to support this idea. The Janus 20 fund shows a striking oscillation over time, with the tech stock mania period of 1999/2000 reflecting very high inflows.

Table II shows some results for individual firms as of December 1999. The table shows the top and bottom ten firms ranked on total dollar flows over the past three years (in the analysis, we focus on flows as a percent of market value, but here we rank on dollar flows in order to generate familiar names). The effect of flows on mutual fund ownership can be fairly sizeable, with flows raising the total ownership of Sun Microsystems from 16% to 20%. Stocks with the biggest inflows tend to be technology stocks, while stocks with the biggest outflows tend to be boring financial or manufacturing firms, closely correspond to our perceptions of investor sentiment in the three year period ending 1999.

In interpreting the flow variable, it is important to remember that flow is a relative concept driven only by differences in flows and holdings across different funds holding different stocks. Flow is not intended to capture is any notion of the absolute popularity of stock. For example, consider Alcoa. The fact the flow variable is large and negative in Table II does not mean that Alcoa was unpopular with mutual funds, nor does it mean that mutual funds are selling Alcoa. It could be that every mutual fund loved Alcoa, held a lot of it, and bought more of it in 1999. What the negative flow means is that the funds which overweighted Alcoa in 1999 received lower-than-average inflows in 1999. Individual investors favored funds which tilted toward stocks like Cisco more than funds which tilted towards stocks like Alcoa.

Table III shows regression evidence on the determinants of the flow variable for individual stocks. We use a six month horizon in defining flow for these regressions and use non-overlapping six month periods to avoid complications in calculating standard errors. The six month horizon is natural to use for these regressions since funds are required to report holdings at least semiannually (so that our variables are updated at least every six months).

Table III follows the basic format of the regressions in this paper. In many of the regressions in this paper, we transform all variables into percentiles for each month. Thus each variable is a number between 0 and 1, representing the rank of the stock on that dimension compared to all other CRSP stocks in that time period. We do this to avoid outliers, to put all variables into the same units for comparison purposes, and to make the results more interpretable in light of the standard portfolio practices of forming quintiles. The regressions are pooled OLS regressions including fixed effects for each month. The independent variables are always lagged at least one month. The standard errors have been adjusted for time clustering, as in Rogers (1993). Thus the regressions are quite similar to Fama-Macbeth. The variables are constructed using the standard CRSP and Compustat sources.

The first two columns of Table III show univariate regressions of six month flows regressed against stock returns over the past year and the past three years. As expected given the previous literature, these regressions show positive and significant coefficients. Flows tend to go to funds that have high past returns, and since funds returns are driven by the stocks that they own, flows tend to go to stocks that have high past returns. The coefficient of 0.10 in the first column means as one goes from the stocks with the most lowest past returns to the stocks with the highest past returns, the percentile ranking of flows goes up 10% (say, from the 40th percentile to the 50th percentile).

The next two columns show two other variable which will be important in understanding flows. The first is the market-book ratio. The market-book ratio follows the definition of Fama and French (1993). Their method of constructing the variable induces substantial staleness (of 6-18 months) in the market-book ratio. The second variable measures corporate issuance. In contrast to the trading by individuals, reflecting uninformed and possibly irrational demand, the actions of firms represents informed and probably more rational supply. A substantial body of research studies whether firms opportunistically take advantage of mispricing by issuing equity when it is overpriced and buying it back when it is underpriced (for example Loughran and Ritter, 1995). Corporate managers certainly say they are trying to time the market (Graham and Harvey, 2001).

We measure firm behavior using the composite share issuance measure of Daniel and Titman (2004). Our version of this variable is 1 minus the firm's (split-adjusted) ratio of the number of shares outstanding firm three years ago to the number of shares outstanding today. For example, if the company has 100 shares and has a seasoned equity issue of an additional 50 shares, the composite issuance measure is 33%, meaning that 33% of the existing shares today were issued recently. We define the variable in this way to make it comparable to the flow measure (both are expressed as a fraction of market value of the company, and are variables bounded above by 100% and unbounded below). The measure can be negative (reflecting for example repurchases) or positive (reflecting for example options given to executives, seasoned equity offerings, stock-financed mergers). Issuance and value are strongly related: growth firms tend to issue stock, value firms tend to repurchase stock. Past research, such as Fama and French (1993) and Daniel and Titman (2004), shows that when either issuance or market-book is high today, returns are low over the next year. Table III shows that for six month flows, neither the market-book ratio nor corporate issuance appear to be important determinants of flows.

Figure 2 shows the history of the top and bottom firms in Table II. In interpreting these graphs, bear in mind that both issuance and flows are rolling backward-looking three year concepts, while market-book is an annual snapshot updated in July of each year (following Fama and French (1993)). Looking at Alcoa, it starts the sample as an extreme value firm, and market-book slowly climbs until it is around average at the end of the sample period. For Cisco, valuations are high throughout the sample period, and Cisco is always a growth firm as measured by market-book. In contrast to market-book, flows seem more variable for these two firms, with Alcoa coming in to favor around 1995 while Cisco falls out of favor at the same time, then the two reverse in the later 1990's. Looking at the figures, there is some sense that the three different variables (market-book, issuance and flows) are positively correlated, but clearly the three variables also contain some information independent of each other.

III. Regression results: Flows and returns

A. Univariate relation between returns and flows

Table IV shows univariate regressions predicting monthly returns with lagged variables. We show the predictive power of flows, and for comparison we show several other variables that are related to flows, and which may have their own previously documented predictive power for returns. The dependent variable is monthly returns in percentage points, while the independent variables are the latest available percentilized independent variables, variously updated at the annual (market-book), monthly (for momentum, reversals, and issuance), or semi-annual/quarterly (flows) frequency.

We first discuss the results for flows. The table shows flows over horizons stretching from three months (one quarter, the shortest interval we have for calculating flows) to three years. Looking at the first column, it is striking that for every horizon but three months, high flows today predict low future stock returns. This relation is statistically significant at the one year and three year horizon. If one is interested in the long-term effects of investor reallocation (whether over time investors benefit from reallocating money across different funds), longer horizons are the appropriate measure. This dumb money effect is the central result of this paper. Focusing on the three year results, the coefficient of -0.90 means as one goes from the stocks with the most extreme outflows to the stocks with the most extreme inflows, average monthly returns fall by 90 basis points.

Perhaps surprisingly, we find no solid evidence for the smart money effect in raw returns, even at the horizons of six to twelve months where one might expect price momentum to dominate. This difference from previous results may be due to two factors. First, by focusing on stock returns instead of fund returns, we avoid many complications involving expense ratios, trading costs, and cash holdings by funds. Second, our measure of flows is quite different than standard because we focus on net flows into individual stocks, not net flows into individual funds. Gruber (1996) and Zheng (1999) focus on funds that have disproportionately high inflows, while we focus on stocks that are disproportionately owned by fund with inflows (as measured by dollar flows compared to market capitalization of the stock). For example, if Cisco is owned by 100 large funds, all of which have slightly higher than average inflows, our measure would classify Cisco as a high sentiment stock. In contrast, the papers cited above would look at individual funds, perhaps small funds that had very high inflows in the past.

The second column shows regressions where returns have been adjusted to control for value, size, and momentum. Following Daniel, Grinblatt, Titman, and Wermers (1997), it subtracts from each stock return the return on a portfolio of firms matched on market equity, market-book, and prior one-year return quintiles (a total of 125 matching portfolios).[4] Here the dumb money effect is substantially reduced, with the coefficient falling from -0.90 to -0.34 for three year flows, still significantly negative but less than half as large. As we shall see, this partially reflects the fact that high sentiment stocks tend to be stocks with high market-book. Thus using a three year horizon, the dumb money effect is statistically distinct from the value effect, but obviously highly related.

One might ask whether the dumb money effect is an implementable strategy for outside investors using information available in real time. Our methodology involves substantial built-in staleness of flows largely reflecting the way that Thomson Financial has structured the data.[5] So the variable in Table IV is certainly in the information set of any investor who has access to all the regulatory filings and reports from mutual funds, as they are filed. Currently, filings appear on the SEC EDGAR system on the next business days following a filing, but information lags were probably longer at the beginning of the sample period.

To address this issue, Table IV shows the six month and three year flow variables, both generously lagged another six months. Even lagged six months, the three year flow variable remains a statistically significant predictor of returns. Thus the dumb money effect is not primarily about short-term information contained in flows, it is about long-term mispricing. Indeed, lagging the six month flow variable causes a substantial improvement in predictive power. This improvement probably reflects that by skipping the most recent six months, we avoid the positive correlation of short-term momentum and short-term flows.

For comparison, Table IV also shows univariate relations for other variables which are both known to predict returns and which are related to flows. The positive relation between lagged annual returns and future returns reflects the momentum effect of Jegadeesh and Titman (1993). The negative coefficient on lagged three year returns reflects the reversal effect of De Bondt and Thaler (1985). The negative coefficient on market-book reflects the value effect of Fama and French (1993). The negative coefficient on issuance reflects the issuance effect of Daniel and Titman (2004), combining a variety of previously documented effects involving repurchases, mergers, and seasoned equity issues. The three year flow effect seems to be roughly the same order of magnitude to these other effects. We also show predictive power of actual mutual fund ownership, and counterfactual three year ownership. Neither variable comes in significant; we return to this relation later.

Table V shows multiple regressions predicting raw monthly returns. We focus on the three year flow horizon. We put measures of value, momentum, and issuance on the right-hand side as an alternative method of controlling for these known effects. Controlling for these variables does not make the dumb money effect go away. Although these additional variables reduce the magnitude of the flow coefficient, it remains significantly negative. We show several robustness tests. First, we shows the results for regressions that are restricted to stocks that are above the median market cap for all CRSP stocks. This restriction modestly decreases the coefficient on flows. We also try splitting the sample in two halves, an important check because of the extreme events of the late 1990's. The table shows that flows remain significant in both halves of the sample, although the dumb money effect is stronger in the second half of the sample period. In this version of the paper, our flow data is noisier in the first half of the sample (due to difficulty matching the holdings data with the CRSP database). In addition, the universe of mutual funds itself was much smaller in the beginning of the sample. Both these facts make the early part of the sample less reliable from a data standpoint.

In summary, three year mutual fund flows strongly negatively predict future stock returns, and there is no horizon at which flows reliable positively predict returns. The dumb money effect is present controlling for value and momentum, present in both large and small cap stocks, and present in different time periods. In terms of statistical significance, sign, and absolute magnitude, it is similar to the value, reversal, and the issuance effects.

IV. Calendar time portfolios, economic significance, and manager skill

In this section, we move from cross-sectional regression evidence to examining monthly returns on calendar time portfolios. We start by forming standard long/short portfolio returns consisting of the top quintile and bottom quintile of various variables. Table VI shows the results, forming portfolios sorted on three year flows, lagged returns, market-book, and corporate issuance. These portfolios are rebalanced monthly with the latest available values. All the portfolios are formed in the same way. We first show equal weight portfolios, while in the next subsection, we look at portfolios that are value weighted (where the weights come from the aggregate holdings of the entire mutual fund sector).

Panel A of Table VI shows summary statistics for these calendar time portfolios, and tells a story similar to tables IV and V. Stocks with high flows have returns that are significantly below stocks with low flows. Looking at mean returns, the difference between high sentiment and low sentiment stocks is 59 basis points per month. This mean differential is somewhat smaller than the other four differentials shown. Looking at t-statistics, however, the dumb money effect is comparatively strong. It is second only to the value effect in statistical significance. Turning now to the monthly return correlations, it is clear from Table V that three year flows produce returns that are highly correlated with issuance and value. Despite the fact that (as shown in Tables IV and V), the dumb money effect holds controlling for value and issuance, the high correlations show that these three effects are very related.

Panel B of Table VI shows yet another way of testing whether the dumb money effect is independent of the value effect. It shows spanning tests, whether the dumb money factor is priced by the value factor. We first regress the equal weight flow portfolio on the equal weight market-book portfolio (using the same monthly return series shown in Panel A). In this formulation, the flow portfolio loads positively on the market-book portfolio, while the alpha is insignificantly different from zero. So this column says that the low returns associated with flows are explained by market-book. On the other hand, the next column shows an alternative measure of the value effect, the HML factor of Fama and French (1993). Here, a significant intercept term remains. The next column shows the full Fama and French (1993) three factor model, which again fails to fully explain the returns on the flow portfolio. The last three columns of panel B give more evidence on subsample stability. Again, the dumb money effect is significant in both halves of the sample, but much stronger in the second half.

In summary, using calendar time portfolio returns shows that the dumb money effect is a statistically strong effect. The evidence on whether the dumb money effect is fully explained by value is mixed at best. The dumb money effect is certainly highly correlated with the value effect.

A. The magnitude of wealth destruction

So far we have shown that stocks owned by funds with large inflows have poor subsequent returns. What is the economic significance of this fact? In this section, we measure the wealth consequences of active reallocation across funds, for the average investor. We abstract from the important issues of fund expenses and trading costs, and look only at the effect on mean returns earned by investors. These expenses and trading costs are another real source of wealth destruction for individual investors, but they have been amply documented elsewhere. We assess the economic significance by measuring the average pre-cost return earned by a representative investor, and comparing it to the pre-cost return he could have earned by simply refraining from engaging in non-proportional flows.

Define R^{ACTUAL} as the return earned by a representative mutual investor who owns a tiny fraction of each existing mutual fund. The returns would reflect a portfolio of stocks where the portfolio weights reflect the portfolio weights of the aggregate mutual fund sector:

$$R_t^{ACTUAL} = \sum_i x_{i,t} \left[\sum_j w_{ij,t} R_t^j \right] \qquad (9)$$

where R^J is the return on stock j. The return from a strategy of refraining from non-proportional flows, R^{NOFLOW}, is

$$R_t^{NOFLOW} = \sum_i \hat{x}_{i,t} \left[\sum_j w_{ij,t} R_t^j \right] \qquad (10)$$

We use three year flows in these calculations. Table VII shows excess returns on these two portfolios, and for comparison shows the value weighted market return as well. Since the two mutual fund portfolios use weights based on dollar holdings, they are of course quite similar to each other and to the market portfolio.

Although very similar, these portfolios are not identical. Table VII shows investor flows cause a significant reduction in both average returns and Sharpe ratios earned by mutual fund investors. A representative investor who is currently behaving like the aggregate mutual fund sector could increase his Sharpe ratio 9% (from a monthly Sharpe ratio of 0.139 to 0.152) by refraining from active reallocation and just directing his flows proportionally.[6]

One can assess the significance of this difference in mean returns by looking at the returns on the long-short portfolio R^{ACTUAL} - R^{NOFLOW}. This return is similar to the long-short portfolio studied in Table V, except that here all stocks owned by the mutual fund sector are included, and the weights are proportional to the dollar value of the holdings. The difference is negative and highly significant.

Thus investor flows cause wealth destruction. This conclusion is, of course, a partial equilibrium statement. If all investors switched to proportional flows, presumably stock prices would change to reflect that. But for one individual investor, it appears that fund flows are harmful to wealth. One component we do not attempt to measure is the transactions cost of switching from one fund to another for individual investors, which would increase the wealth destruction implied by current behavior.

How large is the wealth destruction caused by flows, compared to other costs associated with mutual funds? As a ballpark estimate, it appears that this wealth destruction is perhaps one fourth or one half as big (in terms of mean returns) by the effects of expenses and trading costs. The annual reduction in returns shown in Table VI is about 0.60%. Since expense ratios on actively managed funds are in the neighborhood of 1% per year and turnover is in the neighborhood of 100% per year, the total incremental cost (compared to the alternative of a low turnover, low cost index fund) of active management is around 1% to 2%. Thus the other sins of active management probably outweigh the deleterious effects of fund flows.

B. Better identification of mutual fund manager skill

Table VII helps disentangle the effect of flows from the effect of manager stock picking. We start by considering the average of $R^{ACTUAL} – R^{M}$, which measures the net return benefit of owning the aggregate fund holdings instead of holding the market (ignoring trading costs and expenses). The average of this difference, 0.02, consists of two components. The first, $R^{ACTUAL} - R^{NOFLOW}$, is the net benefit of reallocations. We already have seen that this dumb money effect is negative. The second, $R^{NOFLOW} – R^{M}$, measures the ability of the mutual fund managers to pick stocks which outperform the market (using value weights for managers). As shown in the table, using raw returns, this stock picking effect is 0.08 per month, with a t-statistic of 1.86.

Thus there is some modest evidence that mutual fund managers do have the ability to pick stocks that outperform the market, once one controls for their clients' tendencies of switching money from one fund to another. As shown in the table, this modest skill is obscured (when looking only at actual holdings) by the clients anti-skill at picking funds.

A different question is whether the fund managers have the ability to outperform more specific benchmarks. The bottom half of Table VII closely follow the approach of Daniel, Grinblatt, Titman, and Wermers (1997) and replaces raw returns in the calculations with characteristic adjusted returns, thus showing whether mutual fund managers can pick stocks that outperform their benchmarks defined by size, value, and momentum. Adjusting for these characteristics, the net benefit of owning mutual funds is now zero. The dumb money effect remains negative and significant (though it is smaller controlling for value, as usual). The stock picking effect is still positive at 0.02, but insignificant. Thus we don't find much evidence that mutual fund managers can beat passive benchmarks, even controlling for their clients behavior.

To help gauge the economic magnitude of the dumb money effect shown in Table VII, we show at the bottom the statistics over the same sample period for HML, the value factor as constructed by Fama and French (1993). Consider the raw dumb money effect of -0.05% per month with associated Sharpe ratio of -0.167. This portfolio is a value weighted concept in that it reflects the difference between a value weighted actual portfolio and a "value weighted" counterfactual portfolio. It does not involve taking large positions in small or illiquid stocks. An investor seeking to exploit the dumb money effect would go long the stocks with outflows, short the stocks with inflows, and would earn a Sharpe ratio of 0.167 per month. Now consider the HML effect, which uses a combination of size-stratification and value weighting to reduce the influence of small stocks. An investor exploiting the value effect in this way could earn a Sharpe ratio of 0.122.

Thus using this metric, the dumb money effect looks stronger than the value effect, and if one regresses HML on this measure of the dumb money effect, it turns out that the intercept is insignificantly different from zero. Thus this particular weighting scheme has the dumb money effect subsuming the value effect. We have now looked at many different methods of whether the dumb money effect is subsumed by the value effect, and gotten three different answers: that the dumb money effect is not subsumed by value (Tables IV and V and parts of Table VI), that it is subsumed by value (parts of Table VI), and that it subsumes value. Thus it appears that the two effects are so related that they are difficult to disentangle. It seems clear that the dumb money effect and the value effect are stemming from the same underlying phenomenon.

V. Conclusion

In this paper, we have shown that individual investors have a striking ability to do the wrong thing. They send their money to mutual funds which own stocks that do poorly over the subsequent months and years. Individual investors are dumb money, and one can use their mutual fund reallocation decisions to predict future stock returns. The dumb money effect is robust to a variety of different control variables, not due to one particular time period, and implementable using real-time information. By doing the opposite of individuals, one can construct a portfolio with high returns. Individuals hurt themselves by their decisions, and we calculate that aggregate mutual fund investor could raise his Sharpe ratio by 9% simply by refraining from destructive behavior. These facts pose a challenge to rational theories of fund flows. Of course, rational theories of mutual fund investor behavior already face many formidable challenges, such as explaining why investors consistently invest in active managers when lower cost, better forming index funds are available.

The evidence on mutual fund flows and stock returns closely mirrors two central patterns in stock prices: value (or reversals) and momentum. There is a tension between these two patterns, since momentum says stocks that have gone up continue to go up, while value/reversals says that stocks that have gone up subsequently go down; the only difference is the time horizon. In the case of the dumb money effect, this tension is less severe: the only reliable pattern is that stocks with high inflows have low subsequent returns. However, the competing momentum and value/reversal effects are clearly present in the data. In the short term, mutual fund flows are highly related to momentum: individual investors send money to funds which have recently gone up. At the short term horizon, the positive momentum effect and the negative dumb money effect seem to cancel each other out. At the long-term, the dumb money effect behaves a lot like the value effect. Although the dumb money effect is by some measures statistically distinct from the value effect, it is clear these two effects are highly related.

The evidence on issuers and flows presents a somewhat nonstandard portrait of capital markets. Past papers have looked at institutions vs. individuals, and tried to test if institutions take advantage of individuals. Here, the story is different. Individuals do trade poorly, but these trades are executed through their dynamic allocation across mutual funds, that is, via financial institutions. As far as we can tell, it is not financial institutions that exploit the individuals, but rather the non-financial institutions that issue stock and repurchase stock. We find some modest evidence that mutual fund managers have stock picking skill, but that any skill is swamped by the actions of retail investors in switching their money across funds. In our data, financial institutions seem more like passive intermediaries who facilitate trade between the dumb money, individuals, and the smart money, firms.

It is clear that any satisfactory theory of the value effect will need to explain three facts. First, value stocks have higher than average returns than growth stocks. Second, using issuance and repurchases, the corporate sector tends to sell growth stocks and buy value stocks. Third, individuals, using mutual funds, tend to buy growth stocks and sell value stocks. One coherent explanation of these three facts is that investor sentiment causes some stocks to be overvalued relative to other stocks, and that firms exploit this mispricing.

ENDNOTES

[1] Sapp and Tiwari (2004) challenge this view, arguing plausibly that the effect merely reflects the price momentum effect in stock returns, not selection ability. Another hypothesis, explored by Wemers (2004), is that (rather than any manager selection ability) mutual fund inflows actually push prices higher.

[2] We handle missing reports as follows: whenever a fund has a missing report between two valid report dates, we assume that the fund did not change its holdings with respect to the previous report.

[3] Another way of describing FLOW is that it is the actual percent ownership by the mutual fund sector, minus the counterfactual percent ownership. Since the actual percent ownership is bounded above by 100%, FLOW is bounded above by 100%. In the counterfactual case, there is no accounting identity enforcing that the dollar value of fund holdings is less than the market capitalization of the stock. Thus FLOW is unbounded below. Values of FLOW less than -100% are very rare, occurring less than 0.01% of the time for three year flows.

[4] These 125 portfolios are reformed every month based on the market equity, M/B ratio, and prior year return from the previous month. Following Fama and French (1993), the M/B ratio is only updated annually in July, based on the value as of the previous December. The portfolios are equal weighted and the quintiles are defined with respect to the entire universe in that month.

[5] The data shows holdings for points in time that reflect both a "vintage" file date (FDATE) and a report date. Neither of the two dates correspond to the actual filing date with the SEC. The report date is the calendar day when a snapshot of the portfolio is recorded, while Thomson Financial always assigns file dates to the corresponding quarter ends of the filings. The report date coincides with the file date about 60% of the time , but in some cases dates back as much as

6 months prior to the file date, as fund manager have discretion about when to take a snapshot of their portfolio to be filed at a subsequent date. These holdings eventually become public information. For accuracy, we always use the end of quarter file date assigned by Thomson Financial. This quarterly interval introduces a source of staleness into the holdings data.

[6] Lamont (2002) finds similar results for the policy of refraining from buying new issues.

References

Barber, B. and Odean T., 2000, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors" with Brad Barber, *Journal of Finance*, Vol. LV, No. 2, 773-806.

Barber, B. and Odean T., 2001, "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment", *Quarterly Journal of Economics*, February 2001, Vol. 116, No. 1, 261-292

Barber, B. and Odean T., 2004, “All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, working paper

Barber, B., Odean T. and Zheng L. (2004) Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows, *Journal of Business*, forthcoming

Barberis, N, and Shleifer, A., 2004, “Style investing”, *Journal of Financial Economics*, Volume 68, Issue 2, May 2003, Pages 161-199

Brown, Stephen J., Goetzmann, William N., Hiraki, Takato, Shiraishi, Noriyoshi and Watanabe, Masahiro, 2002, "Investor Sentiment in Japanese and U.S. Daily Mutual Fund Flows" (September 2002). Yale ICF Working Paper No. 02-09

Chen, H., Jegadeesh, N., Wermers, R., 2000. The value of active mutual fund management: an examination of the stockholdings and trades of fund managers. Journal of Financial and Quantitative Analysis 35, 343–368.

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

Cohen, Randolph B., Coval, Joshua and Pástor, Lubos, 2003Judging fund managers by the company they keep; Working paper 04-023.; Boston: Division of Research Harvard Business School, 2003.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52, 1035-1058.

Daniel, Kent, and Sheridan Titman, 2004, Market reaction to tangible and intangible information, working paper

E. Kacperczyk, Marcin T., Clemens Sialm, and Lu Zheng (2005): “On the Industry Concentration of Actively Managed Equity Mutual Funds,” Forthcoming: *Journal of Finance*, 2005.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3—56.

Goetzmann, W., and Massa, M., 2002. Daily momentum and contrarian behavior of index fund investors. *Journal of Financial and Quantitative Analysis* 37, 375-390.

Graham, John R., and Campbell Harvey, 2001, The Theory and Practice of Corporate Finance: Evidence from the Field, *Journal of Financial Economics* 60, 187-243.

Grinblatt, M., and S. Titman, and R. Wermers, 1995, Momentum Investment Strategies, Portfolio Performance and Herding, A Study of Mutual Fund Behavior, *American Economic Review*, 85, 1088-1105

Grinblatt, M., and Keloharju, M., 2000. The investment behavior and performance of various investor types: A study of Finland's unique data set. *Journal of Financial Economics* 55, 43-67.

Gruber, Martin, 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783-810

Ippolito, Richard A., 1992, "Consumer reaction to measures of poor quality: evidence from the mutual fund industry," *Journal of Law and Economics* 35 (April 1992), 45-70.

Jain, Prem C., and Wu, Joanna Shuang, 2000, Truth in mutual fund advertising: Evidence on future performance and fund flows, *Journal of Finance* 55, 937-958.

Jegadeesh, N. and S. Titman, 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65-91.

Lamont, Owen A., 2002, Evaluating value weighting: Corporate events and market timing, NBER Working Paper No. 9049.

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

Odean T., 1999, "Do Investors Trade Too Much?", *American Economic Review*, Vol. 89, 1279-1298.

Rogers, W., 1993. Regression standard errors in clustered samples. Stata Technical Bulletin 13, 19-23.

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

Teo, Melvyn and Sung-Jun Woo, 2001, Persistence in Style-Adjusted Mutual Fund Returns, *Journal of Financial Economics* forthcoming

Travis Sapp and Ashish Tiwari, 2004, Does Stock Return Momentum Explain the "Smart Money" Effect?, *Journal of Finance*, forthcoming

Wermers, R., 1999, Mutual Fund Trading and the Impact on Stock Prices, *Journal of Finance*, 54, 581-62

Wermers R., 2000, Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses, *Journal of Finance*, 55, 1655-1695

Wermers, Russell, 2004, Is Money Really "Smart"? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence, Working Paper

Zheng, Lu, 1999, Is money smart? A study of mutual fund investors' fund selection ability, *Journal of Finance* 54, 901-933.

Data Appendix

A. Holdings data and error screens

We obtain data on stocks holding for equity mutual funds investing in the US between 1980 and 2003 from the Thomson Financial CDA/Spectrum Mutual Funds database. Since our focus is on US equity funds, we remove all US-based international funds, fixed-income funds, real estate funds and precious metal funds.

Holdings are identified by CUSIPs, they constitute most of the equities, but are not necessarily the entire equity holdings of the manager or fund. The potential exclusions include: small holdings (typically under 10,000 shares or $200,000), cases where there may be confidentiality issues, reported holdings that could not be matched to a master security file, and cases where two or more managers share control (since the SEC requires only one manager in such a case to include the holdings information in their report).

The statutory requirement for reporting holdings is semi-annual, although some funds file quarterly reports. The data include a report date (RDATE), which is the calendar day when a snapshot of the portfolio is recorded, and a file date (FDATE), which is a vintage date assigned by Thomson. Neither of the two dates corresponds to the actual filing date with the SEC. Thomson always assigns file dates to the corresponding quarter ends of the filings.

Thomson identifies funds using a five-digit number (FUNDNO) but unfortunately numbered identifiers are reused in the data, hence we use a filter to identify new born-funds and generate a unique fund identifier. We start tracking funds as they appear in the database, a fund is then classified as a new-born fund and assigned a new unique identifier whenever there is a gap of more than 1 year between the current report and the last available report. A gap of more than one year between two consecutive reports typically reflects a different and unrelated manager or a major reorganization of the fund.

Holding are adjusted for stock splits, stock distributions, mergers and acquisitions and other corporate events that occur between the report date and the file date. This adjustment relies on the assumption by Thomson that funds report shares held on a pre-adjustment basis.

We merge the holdings with the CRSP/COMPUSTAT data and we use a series of filters to eliminate potential anomalies, probably due to misreporting, errors in data collecting or in computing adjustments. Holdings are set to missing whenever:

1. The report date is subsequent to the file date
2. The number of shares in a fund portfolio exceeds the total amount of shares outstanding at a particular date
3. The total amount of shares outstanding reported by CRSP is zero at a particular date

B. Merging Thomson and CRSP data

The CRSP mutual fund database utilizes a five character alpha-numeric identifier (ICDI). Both database report funds names but they use a different character string with different abbreviations. To match the two datasets we use a matching procedure base on TICKER symbols and fund names, similar in spirit to the technique proposed by Wermers (2000).

Thomson Financial reports funds tickers on a quarterly basis starting from the first quarter of 1999. For fund portfolios offering multiple share classes, multiple ticker symbols are provided. A combination of ticker-date typically uniquely identifies a mutual fund. First, we merge the two databases using a ticker-date match between the first quarter of 1999 and the last quarter of 2003. We generate a list of unique matches between the CRSP fund identifier and the unique identifier in the Thomson data computed above, and extrapolate backwards for the prior years.

After this initial merge, we use a "fuzzy" string matching algorithm to match the remaining funds. We use a "SOUNDEX" algorithm to match funds using their name and the corresponding date. The SOUNDEX algorithms were patented by Margaret I. Odell in 1918 and Robert C. Russell in 1922. They are based on an underlying principle of English and other Indo-European languages. That is, most of the words can be reasonably represented by consonants alone. All the names are reduced to a phonetic equivalent character strings which can later be compared. We transform fund names into an alpha-numeric indicator by using the following steps:

1. Retain the first letter of the fund name and discard the letters A E H I O U W Y
2. Assign a numeric value to the following consonant: 1 $\rightarrow$ B F P V, 2 $\rightarrow$ C G J K Q S Z, 3 $\rightarrow$ D T, 4 $\rightarrow$ L, 5 $\rightarrow$ M N, 6 $\rightarrow$ R
3. Discard all duplicate classification values if they are adjacent (that is BB will results in the single value 1)

We use the resulting strings to match the remaining funds at every quarterly date, and we discard every fund for which we could not find a corresponding match. Below we show a portion of the matched file:

date	CDA Fund ID	Thomson name	CRSP ICDI	CRSP name

12/31/2003	204	LORD ABBETT RES LG CAP S	13848	Lord Abbett Large Cap Research Fund/Y
03/31/1995	205	HERITAGE SER TR-VAL EQTY	13596	Heritage Series Trust:Value Equity Fund/A
06/30/1995	205	HERITAGE SER TR-VAL EQTY	13596	Heritage Series Trust:Value Equity Fund/A
06/30/1995	205	HERITAGE SER TR-VAL EQTY	13598	Heritage Series Trust:Value Equity Fund/C
09/30/1995	205	HERITAGE SER TR-VAL EQTY	13596	Heritage Series Trust:Value Equity Fund/A
09/30/1995	205	HERITAGE SER TR-VAL EQTY	13598	Heritage Series Trust:Value Equity Fund/C
12/31/1995	205	HERITAGE SER TR-VAL EQTY	13596	Heritage Series Trust:Value Equity Fund/A
12/31/1995	205	HERITAGE SER TR-VAL EQTY	13598	Heritage Series Trust:Value Equity Fund/C
09/30/2000	252	LIBERTY STRATEGIC BALANC	12722	Liberty Strategic Balanced Fund/B
09/30/2000	252	LIBERTY STRATEGIC BALANC	12724	Liberty Strategic Balanced Fund/C
01/31/1995	253	GOLDMAN S BALANCED FD	13706	Goldman Sachs Tr:Balanced Fund
07/31/1995	253	GOLDMAN S BALANCED FD	13706	Goldman Sachs Tr:Balanced Fund
01/31/1996	253	GOLDMAN S BALANCED FD	13706	Goldman Sachs Tr:Balanced Fund
07/31/1996	253	GOLDMAN S BALANCED FD	13706	Goldman Sachs Tr:Balanced Fund
01/31/1997	253	GOLDMAN S BALANCED FD	13706	Goldman Sachs Equity Port:Balanced Fund/A
07/31/1997	253	GOLDMAN S BALANCED FD	09039	Goldman Sachs Equity Port:Balanced Fund/C

In the CRSP database, if a fund has multiple share classes, each share class is classified as a separate entity. Different share classes have the same portfolio composition and are treated as a single fund in the Thomson database (for example fund # 205 in the table above). Therefore we combine multiple share classes in the CRSP data into a unique "super fund" by aggregating the corresponding net asset values, and computing the weighted average return of the fund using the total net asset value of the different share classes as weights.

As a final step, to ensure matching quality, we compare the net asset values of the matched funds reported by CRSP to the dollar value of their holdings, and discard matches where the total asset value of the fund reported by CRSP differs from the sum of the dollar holdings value by more than 100%.

C. <u>Construction of the counterfactual flows</u>

The purpose of this exercise is to mimic a mechanic alternative allocation of the total flows to the universe of equity funds that ignores returns in the last *k* quarters, and assign to every existing fund a proportional share of the total flows.

Given out definition of flows, funds that are born have inflows equal to their initial NAV, while funds that die have outflows equal to their terminal NAV. We assign a counterfactual net asset value of zero to funds that were newly created in the past *k* months. New funds represent new flows, but in the counterfactual exercise they do not receive assets for the first *k* quarters. The universe of funds we consider when computing the counterfactual flows between date *t-k* and date *t* is funds there were alive at both date *t-k* and *t*.

More specifically, consider at generic date t and let $\mathrm{F}_{\mathrm{s}}^{\mathrm{Agg}}$ be the actual aggregate flows for all funds alive in quarter *t* (including funds who were recently born, but excluding funds that die in month t), for $t-k \le s \le t$. Let $\mathrm{N}_{\mathrm{t\text{-}k}}^{\mathrm{Agg}}$ be the lagged actual aggregate NAV aggregating only over those funds that exist in both month *t-k* and in month *t*. We compute the counterfactual flows by assigning to each fund a share of total as follows:

$$\hat{\mathrm{F}}_{\mathrm{s}}^{\mathrm{i}} = \frac{N_{t-k}^{i}}{N_{t-k}^{Agg}} \mathrm{F}_{\mathrm{s}}^{\mathrm{Agg}} \qquad (1)$$

$$t-k \le s \le t \qquad (2)$$

For funds that die in quarter $s+1$ (so that their last NAV is quarter s), we set $\hat{\mathrm{F}}_{\mathrm{s+1}}^{\mathrm{i}}$ = $-\hat{N}_{s}^{i}$ and $\hat{N}_{s+h}^{i} = 0$ for all $h > 0$.

Table A shows a simplified example where we set *k* = 1 year. Fund # 3 is born in 1981, therefore in 1981 we register a net inflow equal to its initial NAV and set the counterfactual NAV to zero. In 1981 two funds are alive, fund # 1 and fund #2, and in 1980 they represented 2/3 and 1/3 of the total fund sector. Aggregate flows in 1981 were equal to $150, hence in the counterfactual exercise we assign a flow of $100 to fund # 1 (as opposed to the actual realized flow of $50) and a flow of $50 to fund # 2. Given the return of the two funds between 1980 and

1981, we can compute the counterfactual net asset value of fund # 1 and # 2 in 1981. Proceeding in the same manner whenever a fund is alive at date *t-k* and *t*, we track the evolution of the fund's counterfactual NAV using the recursion:

$$\hat{\mathrm{N}}_{\mathrm{t}}^{\mathrm{i}} = (1 + R_t^i)\hat{\mathrm{N}}_{\mathrm{t\text{-}1}}^{\mathrm{i}} + \hat{\mathrm{F}}_{\mathrm{t}}^{\mathrm{i}} \qquad (3)$$

Between 1982 and 1993 fund # 2 dies, hence in the counterfactual world we assign an outflow in 1983 equal to the NAV in 1982 and set the counterfactual NAV to zero thereafter. Note that (2) does not guarantee that counterfactual net asset values are always non-negative in quarters where we have aggregate outflows ($\mathrm{F}_{\mathrm{t}}^{\mathrm{Agg}} < 0$). In this case we override (2) , set $\hat{\mathrm{N}}_{\mathrm{t}}^{\mathrm{i}} = 0$ and redistribute the corresponding counterfactual flows to the remaining funds, to keep the total aggregate dollar outflow the same in both the counterfactual and actual case. Measuring FLOW ownership over 12 quarters, negative counterfactual NAV occur for only 0.12% of the sample.

Finally, we handle mergers as follows: we assume that investors in the merged funds place their money in the surviving fund and keep earning returns on the existing assets. For consistency, when constructing the counterfactual NAV we also merge the lagged NAV of the two funds when we compute the ratio $\frac{N_{t-k}^{i}}{N_{t-k}^{Agg}}$ used to determined the pro-rata share to the total flows.

We obtain a time series of counterfactual NAVs by repeating this exercise at every date with a rolling window of size k .

Figure 1: Individual funds as a percent of the fund universe over time, using actual and counterfactual size

x is the fund's actual percent of dollar value of the total mutual fund universe in our sample. Xhat is counterfactual percent, using a horizon of twelve quarters (three years).

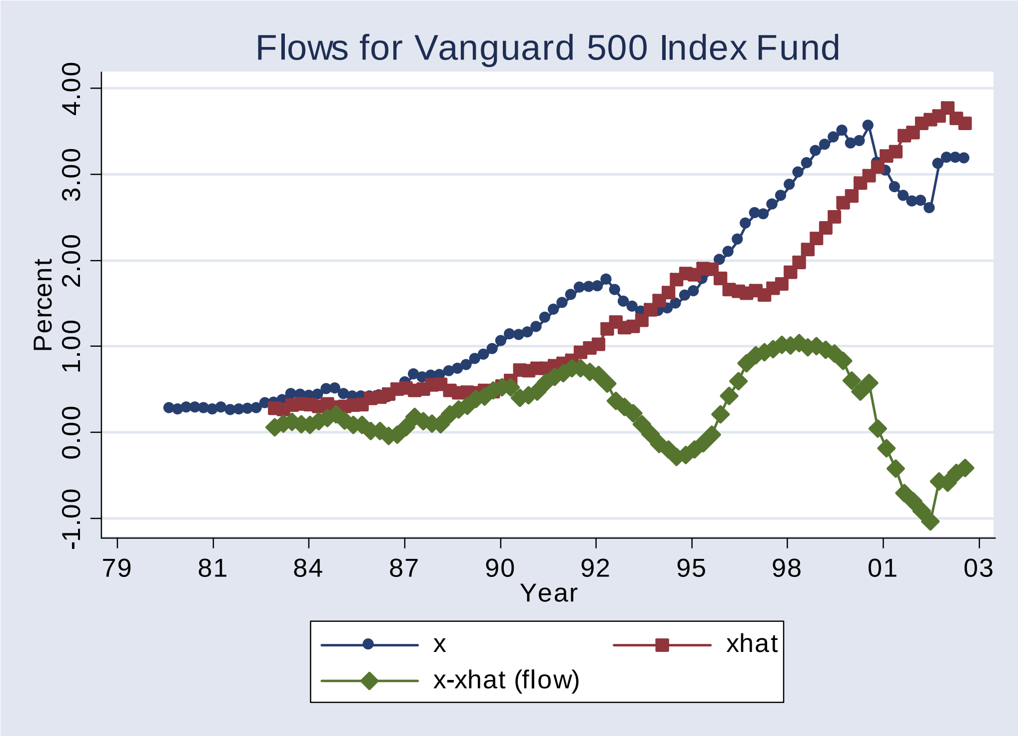
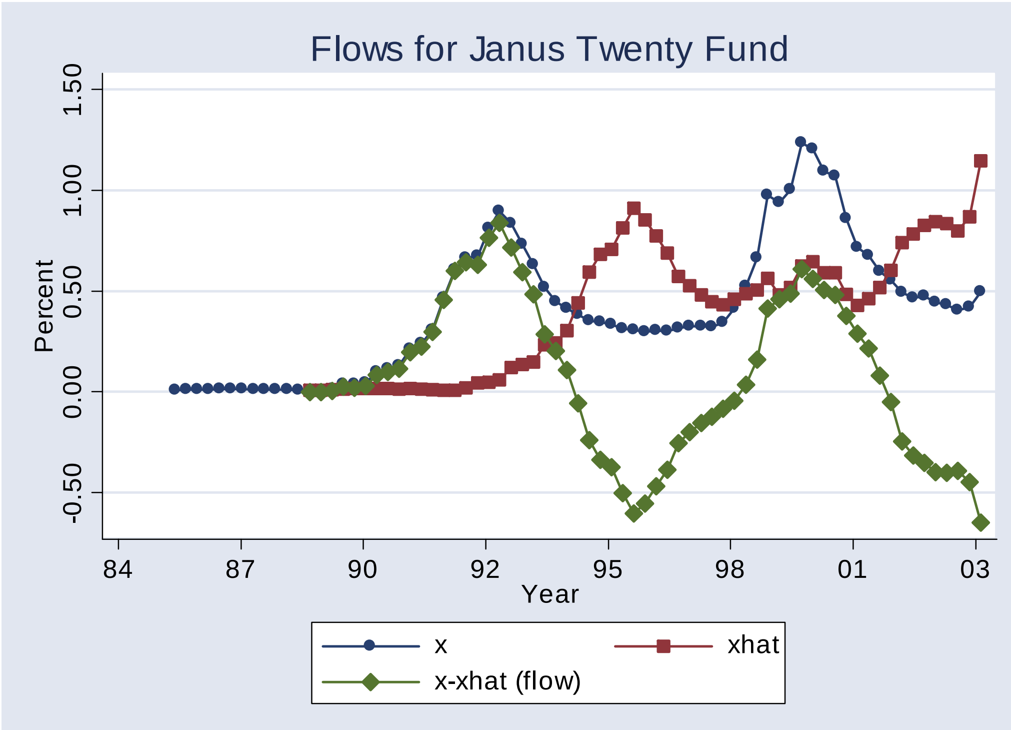


Figure 1, continued



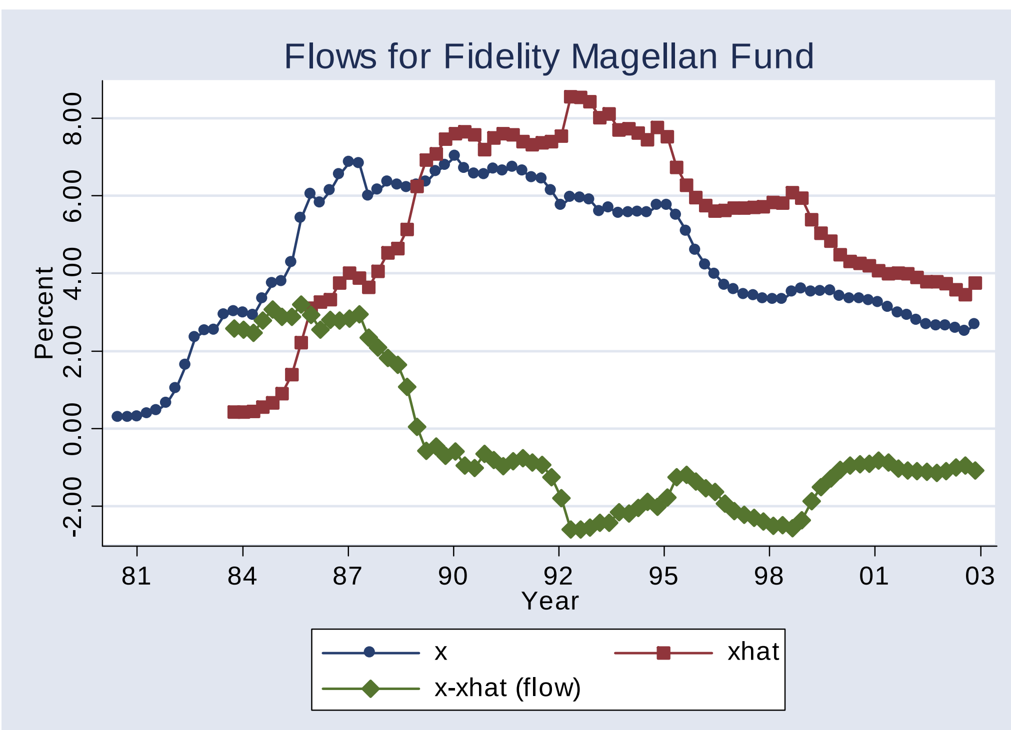


Figure 2
Flows, Market/Book, and Issuance for Individual Firms

Flows use a three year horizon. M/B is defined as in Fama and French and is updated annually. Issuance is the split-adjusted number of shares outstanding three years ago divided by the number of share outstanding today (reflecting shares issued or repurchased by the firm over the past three years). All variables are percentilized, by calculating their rank in the universe of all firms for that month.

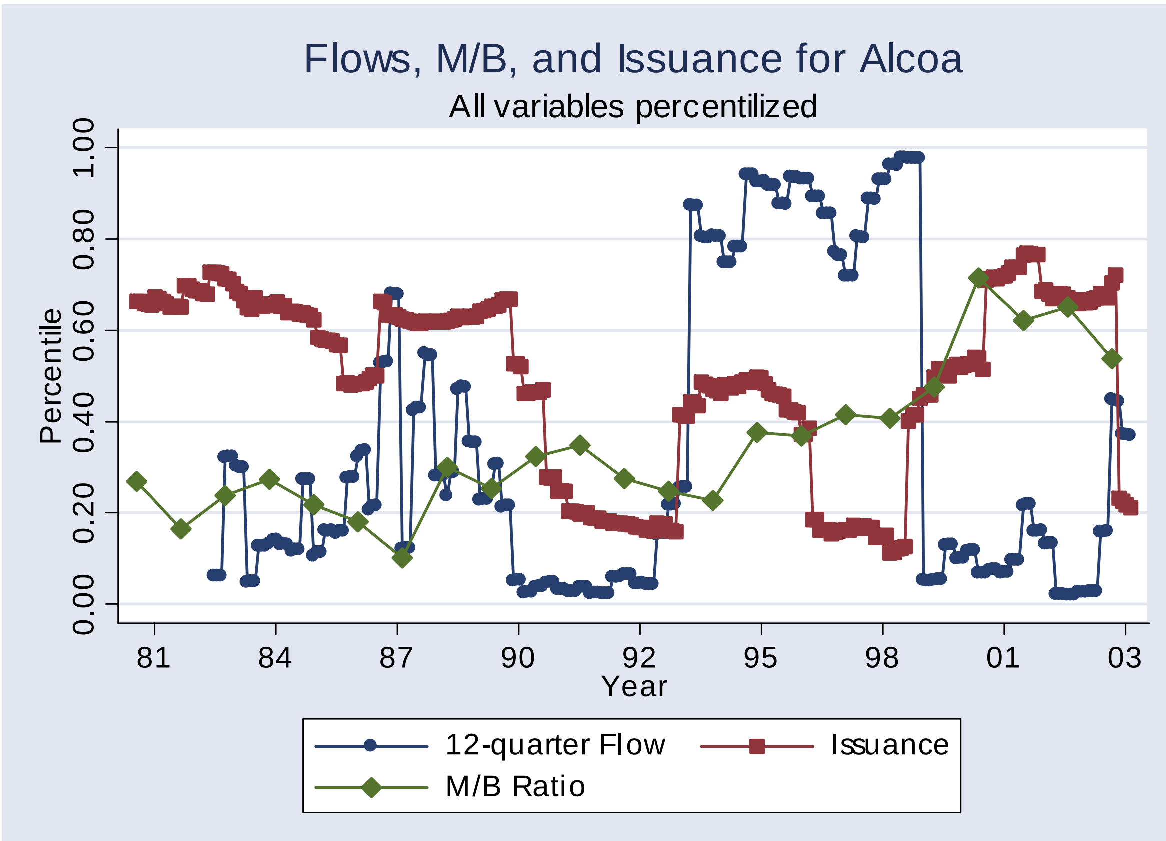


Figure 2, continued

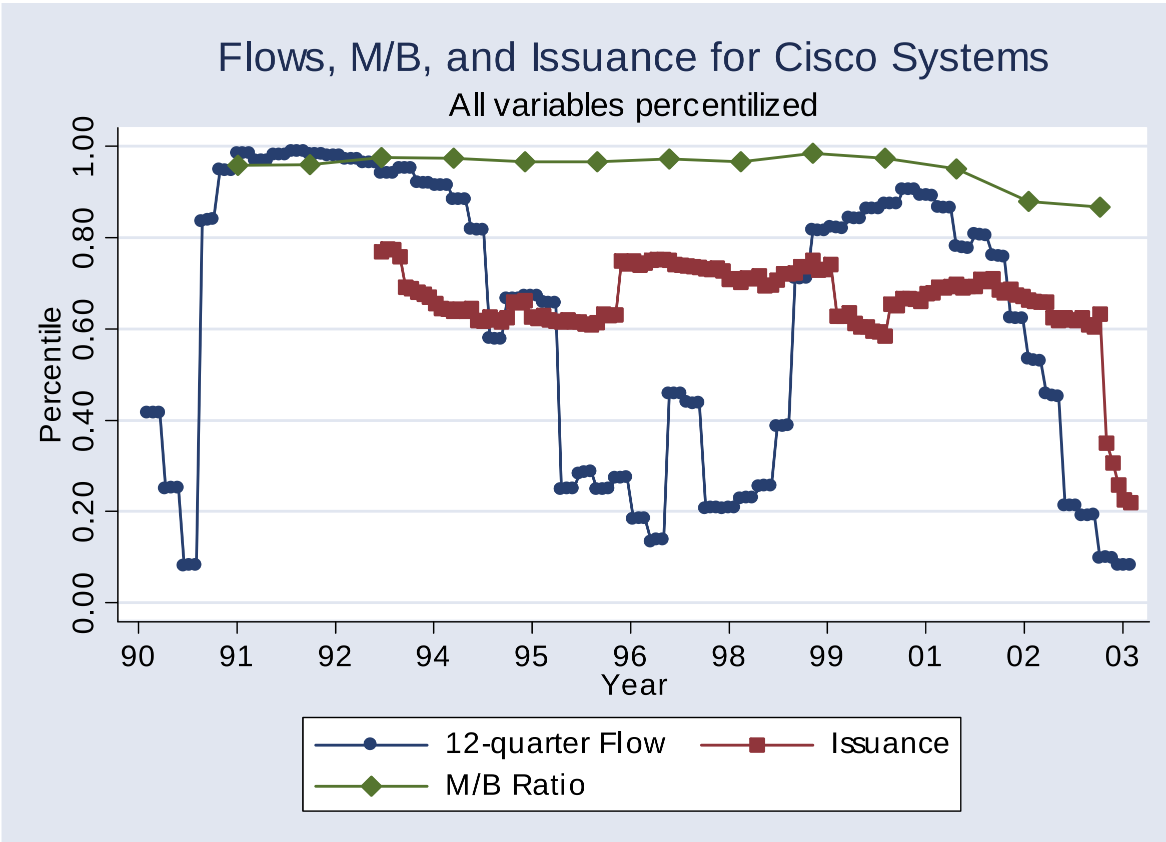


Table I: Flows by fund in December 1999
Using 3 year flows, top and bottom ten ranked on difference between actual and counterfactual

	Percent of fund universe, actual	Percent of fund universe, counterfactual	Diff
	x_{it}	$\hat{x}_{it}$	
FIDELITY MAGELLAN FUND	3.54	4.83	-1.29
INVESTMENT COMPANY OF AM	1.87	2.46	-0.58
VANGUARD WINDSOR FUND	1.46	1.97	-0.51
FIDELITY EQUITY INCOME I	0.59	1.07	-0.48
FIDELITY CONTRAFUND	1.61	2.07	-0.46
AIM CONSTELLATION FUND	0.64	1.05	-0.40
AMERICAN CENT ULTRA FUND	1.46	1.84	-0.38
FIDELITY PURITAN FUND	0.81	1.19	-0.37
PBHG GROWTH FUND INCORPO	0.15	0.52	-0.37
FIDELITY ASST MGR	0.44	0.77	-0.33
MUNDER NET NET FUND	0.24	0.00	0.24
FRANKLIN STRAT SML MID C	0.36	0.09	0.27
DAVIS NEW YORK VENTURE F	0.54	0.27	0.27
MFS MA INVESTORS TRUST	0.52	0.23	0.29
MFS MA INVESTORS GWTH ST	0.45	0.15	0.30
VANGUARD TOT STK MKT IND	0.74	0.30	0.44
VANGUARD GROWTH INDEX FU	0.52	0.08	0.44
ALLIANCE PREMIER GROWTH	0.60	0.07	0.53
JANUS TWENTY FUND	1.23	0.62	0.61

Table II: Flows by stock in December 1999
Using three year flows, top and bottom ten ranked on dollar FLOW

	Percent owned by mutual funds		
	Actual	Counterfactual	
	z_{it}	$\hat{z}_{it}$	FLOW
ALCOA INC	34.10	40.74	-6.63
FEDERAL NATIONAL MORTGAGE ASSN	33.58	36.44	-2.86
CENDANT CORP	30.98	39.24	-8.26
VIACOM INC	40.16	44.34	-4.19
FEDERATED DEPT STORES INC DEL	40.48	52.65	-12.17
CHASE MANHATTAN CORP NEW	26.19	28.11	-1.92
CITRIX SYSTEMS INC	33.57	44.40	-10.84
ASSOCIATES FIRST CAPITAL CORP	42.73	48.34	-5.61
GENERAL MOTORS CORP	20.15	22.32	-2.17
EATON CORP	60.24	78.86	-18.62
WAL MART STORES INC	10.30	9.36	0.94
E M C CORP MA	22.14	19.49	2.64
GENERAL ELECTRIC CO	12.22	11.59	0.63
LUCENT TECHNOLOGIES INC	10.51	9.08	1.43
DELL COMPUTER CORP	11.52	8.51	3.01
INTEL CORP	11.86	10.29	1.58
AMERICA ONLINE INC	18.00	15.19	2.81
SUN MICROSYSTEMS INC	20.03	15.80	4.24
MICROSOFT CORP	12.55	11.30	1.25
CISCO SYSTEMS INC	16.96	14.49	2.47

Table III
Determinants of flows

Dependent variable is six month flows. Independent variables are past returns, market-book, and corporate issuance as independent variables. All independent variables are percentilized, by calculating their rank in the universe of all firms for that month. M/B is market-book ratio (market value of equity divided by Compustat book value of equity). The timing of M/B follows Fama and French (1993) and is as of the previous December year-end. Issuance is the split-adjusted number of shares outstanding three years ago divided by the number of share outstanding today (reflecting shares issued or repurchased by the firm over the past three years). Regressions include fixed time period effects. Sample period is 1981 to 2003, and includes observations for June and December of each year. Standard error, clustered by time period, in parenthesis.

Lagged one	0.10				0.08
year return	(0.01)				(0.01)
Lagged three		0.09			0.06
year return		(0.02)			(0.01)
M/B ratio			0.05		0.00
			(0.03)		(0.02)
Corporate issuance				0.03	0.03
				(0.02)	(0.01)
Number of obs, thousands	156K	141K	165K	141K	118K
Number of semiannual periods	45	45	45	45	45

Table IV
Predicting returns, univariate

Monthly returns as dependent variable, and flows, past returns, market-book, and corporate issuance as independent variables. All independent variables are percentilized, by calculating their rank in the universe of all firms for that month. M/B is market-book ratio (market value of equity divided by Compustat book value of equity). The timing of M/B follows Fama and French (1993) and is as of the previous December year-end. Issuance is the split-adjusted number of shares outstanding three years ago divided by the number of share outstanding today (reflecting shares issued or repurchased by the firm over the past three years). Characteristic adjusted are returns minus the returns on an equal weighted portfolio of all CRSP firms in the same size, market-book, and one year momentum quintile. Regressions include fixed time period effects. Sample period is 1981 to 2003. Standard error, clustered by time period, in parenthesis.

	Raw returns		Characteristic adjusted returns	
	Coeff.	(s.e.)	Coeff.	(s.e.)
Three month flow	0.09	(0.32)	0.18	(0.12)
Six month flow	-0.08	(0.33)	0.08	(0.12)
One year flow	-0.50	(0.20)	-0.22	(0.12)
Three year flow	-0.90	(0.29)	-0.34	(0.13)
Actual mutual fund ownership, z	-0.29	(0.40)	0.06	(0.13)
Three year counterfactual mutual fund ownership, $\hat{z}$	-0.10	(0.43)	0.10	(0.14)
Lagged one year return	1.16	(0.47)	0.03	(0.02)
Lagged three year return	-0.83	(0.47)	-0.30	(0.23)
M/B ratio	-1.77	(0.34)	-0.07	(0.02)
Corporate issuance	-0.91	(0.34)	-0.46	(0.19)
Six month flow, lagged six months	-0.58	(0.32)	-0.18	(0.13)
Three year flow, lagged six months	-0.84	(0.26)	-0.36	(0.13)

Table V
Predicting returns, multiple variables

Raw monthly returns as dependent variable, and percentilized flows, past returns, market-book, and corporate issuance as independent variables. "Larger cap stocks" are all stocks with market capitalization above the median of the CRSP universe that month. Regressions include fixed time period effects. Standard error, clustered by time period, in parenthesis.

	All stocks, 1983-2003	Larger cap stocks, 1983-2003	All stocks, 1983-1993	All stocks, 1994-2003
Three year flow	-0.55	-0.41	-0.38	-0.66
	(022)	(0.21)	(0.19)	(0.32)
Lagged one year return	1.63	1.65	1.46	1.65
	(0.61)	(0.62)	(0.43)	(0.86)
Lagged three year return	-1.42	-0.74	-0.04	-2.05
	(0.61)	(0.44)	(0.62)	(0.83)
M/B ratio	-0.52	-0.47	-0.82	-0.38
	(0.41)	(0.40)	(0.36)	(0.46)
Corporate issuance	-0.60	-0.66	-0.64	-0.61
	(0.36)	(0.30)	(0.20)	(0.54)
Number of obs, thousands	657K	460K	240K	417K
Number of months	249	249	129	120

Table VI
Calendar time returns for portfolios constructed using different sorting variables.

Shows the property of monthly calendar time portfolio returns. Portfolios are equal weighted, constructed by going long the top 20% of stocks and short the bottom 20% of stocks. RMRF is returns on the CRSP value weighted portfolio minus T-bill returns. HML is the value factor (the return of low M/B stocks high M/B stocks), SMB is the size factor (the return on small stocks minus big stocks). There are 249 monthly return observations for three year flows from 1983-2003.

Panel A. Summary statistics

				------------- Correlations -------------			
	Mean	Std. Dev	t-stat	Three year flow	Lagged one year return	Lagged three year return	M/B ratio
Sorting variable							
Three year flow	-0.59	2.99	3.11	1.00			
Lagged one year return	0.84	6.19	2.16	-0.12	1.00		
Lagged three year return	-0.70	6.34	1.75	0.11	0.52	1.00	
M/B ratio	-1.42	4.34	5.16	0.66	-0.17	0.00	1.00
Corporate Issuance	-0.84	4.30	3.07	0.58	-0.31	-0.53	0.71

Table VI, continued

Panel B. Multifactor relations and subperiods

	1981-2003				1981-1993	1994-2003
Intercept	0.06	-0.37	-0.42	-0.59	-0.31	-0.89
	(0.15)	(0.15)	(0.16)	(0.19)	(0.16)	(0.35)
Portfolio constructed using M/B ratio	0.45					
	(0.03)					
RMRF			0.03			
			(0.04)			
HML		-0.53	-0.46			
		(0.05)	(0.06)			
SMB			0.11			
			(0.05)			
R2	0.43	0.35	0.36	0.00	0.00	0.00

Table VII
Effects of flows on monthly returns for aggregate mutual fund investor

Shows the property of monthly calendar time portfolio returns. Uses three year flows. R^{ACTUAL} is returns on a mimicking portfolio for the entire mutual fund sector, with portfolio weights the same as the actual weights of the aggregate mutual fund sector. R^{NOFLOW} is returns on a mimicking portfolio for the counterfactual mutual fund sector, with portfolio weights the same as the counterfactual weights of the aggregate mutual fund sector. Tildes indicate characteristic adjusted returns, defined as raw returns minus the returns on a equal weighted portfolio of all CRSP firms in the same size, market-book, and one year momentum quintile.

		Mean	t-stat	SR
Actual excess return on mutual fund holdings	$R^{ACTUAL} - R^{F}$	0.68	2.19	0.139
Counterfactual excess return	$R^{NOFLOW} - R^{F}$	0.73	2.40	0.152
on mutual fund holdings (three year horizon)				
Market excess returns	$R^{M} - R^{F}$	0.62	2.26	0.143
Net benefit of mutual funds	$R^{ACTUAL} - R^{M}$	0.02	0.63	0.040
Dumb money effect	$R^{ACTUAL} - R^{NOFLOW}$	-0.05	2.64	-0.167
Stock picking	$R^{NOFLOW} - R^{M}$	0.08	1.86	0.077
Net benefit of mutual funds	$\tilde{R}^{ACTUAL}$	-0.00	0.04	-0.003
Adjusted for value, size, momentum				
Dumb money effect	$\tilde{R}^{ACTUAL} - \tilde{R}^{NOFLOW}$	-0.02	2.16	-0.137
Adjusted for value, size, momentum				
Stock picking	$\tilde{R}^{NOFLOW}$	0.02	0.58	0.037
Adjusted for value, size, momentum				
The value effect	HML	0.41	1.93	0.122

Table A.1

Hypothetic example showing counterfactual calculation

	Year	1980	1981	1982	1983	1985
ACTUAL DATA FOR INDIVIDUAL FUNDS==						
Returns	Fund 1	10%	10%	5%	10%	5%
	Fund 2	-5%	10%	-10%		
	Fund 3			10%	10%	5%
NAV	Fund 1	100	160	268	395	515
	Fund 2	50	105	144	0	0
	Fund 3		50	45	100	154
FLOWS	Fund 1		50	100	100	100
	Fund 2		50	50	-144	0
	Fund 3		50	-10	50	50
ACTUAL DATA FOR AGGREGATES=======						
NAV	Agg.	150	315	457	494	669
FLOW	Agg.	0	150	140	6	150
NAV, last year, of funds existing this year	Agg.		150	315	313	494
FLOW of non-dying funds	Agg.		150	140	150	150
COUNTERFACTUAL DATA=============						
NAV	Fund 1	100	210	292	449	591
	Fund 2	50	105	141	0	0
	Fund 3			22	46	79
FLOWS	Fund 1		100	71	128	120
	Fund 2		50	47	-141	0
	Fund 3			22	22	30