

Investing in mutual funds: the determinants of implied and actual net cash flows

Diego Víctor de Mingo-López^a and Juan Carlos Matallín-Sáez^{a*}

^a*Department of Finance and Accounting, Universitat Jaume I, Castellón, Spain*

Abstract: Estimating the fund investors' demand plays an important role in the mutual fund management. In this line, mutual fund demand can be measured as the total net cash flows experienced by the fund during a period. Due to a lack of the data for inflows and outflows in some countries and databases, many authors estimate the net cash flows using fund size and return information. This rough measure, although being a good approximation, implicitly assumes an error in its calculation. For a sample of 2,985 US open-end funds, we find evidence that estimating this implied fund flows, the error generated is higher for smaller funds, funds with higher returns, and for those experiencing higher levels of inflows or outflows. This lack of precision leads to a distortion in the estimation of the effect of some determinants on the mutual fund demand, especially when longer periods are considered when constructing the net cash flows.

Keywords: mutual fund, net cash flow, implied flow, fund demand

JEL classification: G11, G23

* We are grateful to an anonymous referee and the editor for their comments, which have contributed to improve the overall quality of the article. Juan Carlos Matallín-Sáez and Diego Víctor de Mingo-López acknowledge the financial support provided by the Spanish Ministerio de Economía y Competitividad (ECO2014-55221-P) and Universitat Jaume I (PREDOC-2014-04). Address for correspondence: Diego Víctor de Mingo-López, Department of Finance and Accounting, Universitat Jaume I, E-12071 Castellón, Spain. Tel: +34 964 728 569, Fax: +34 964 728 565. E-mail: demingo@uji.es.

1. Introduction

Understanding mutual fund investors' behaviour has attracted much attention from both professionals and researchers. Many authors have studied the effect of characteristics such as performance or expenses on the demand for mutual funds. And measuring the flows into and out of the funds is a reasonable way to estimate this demand.

In this context, the standard definition of net cash flows or "new money" in a fund during a given period is equal to the fund size in the same period minus the appreciation of the fund size in the previous period; that is, the growth of the fund with respect to the growth that would have happened with no flows, and with all the dividends reinvested in the fund. This rough definition has been used in several studies (e.g., Barber *et al.*, 2005; Cooper *et al.*, 2005; Gruber, 1996; Guercio and Reuter, 2014; Huang *et al.*, 2007; Jayaraman *et al.*, 2002; Zhao, 2005), mainly due to the lack of more specific data.

However, it is worth noting that this estimate of net cash flow entails certain implicit assumptions. For example, net cash flows occur in the last moment of each period, so they incur neither return nor related costs during that period. Aware of this fact, some authors have also considered that such flows occur at the beginning of the period, and conclude that using one or the other method does not lead to significant differences in their results (Bhattacharya *et al.*, 2013; Feng *et al.*, 2014; Friesen and Sapp, 2007; Sirri and Tufano, 1998; Zheng, 1999). Nevertheless, this is also a rough definition for estimating mutual fund flows, and does not provide the precise amount of investors' flows going into and out of the fund.

Other authors, however, have emphasized the importance of defining mutual funds' net cash flows using specific information on inflows and outflows (Andreu and Sarto, 2016; Christoffersen *et al.*, 2013; Ivković and Weisbenner, 2009; Keswani and Stolin, 2008). That is, the amount of net cash flows experienced by the fund during a period is equal to the total inflows minus the total outflows generated in that period. Thus, this method estimates net cash flows accurately, while previous methods provide an approximate picture of the real flows into and out of a mutual fund.

Despite this, the fund data required to estimate fund flows precisely are not always available for some countries and databases.

Accordingly, an error may be introduced by using a rough method as the definition of mutual fund flows, which would lead to differences in the results of the determinants of net flows, or even in their persistence. To our knowledge, only Keswani and Stolin (2008) make a brief comparison between rough and accurate methods for a sample of UK mutual funds, comparing the regressions of these measures on some variables, such as the lagged flow or the performance experienced by the fund. They attribute the differences in the slopes of each regression to the inherent noise created when estimating implied flows.

Then, the main interest of this paper is to analyse the effect of the determinants of the mutual fund demand, showing that the use of a rough measure can lead to a noise in the estimate of the flows experienced by the fund. For this purpose, we analyse a large sample of US equity mutual funds. Our results indicate that depending on the methodology applied, there are important differences in the net cash flow estimates. These differences are higher for smaller funds, funds with higher inflows and outflows, and funds experiencing higher returns in absolute terms. Furthermore, the use of the implied flows in the regressions causes an error in estimating the effect of determinants that explain the variability of mutual fund flows, especially during bullish periods. This lack of precision is higher particularly when fund flows are estimated for longer periods.

The rest of this paper is structured as follows. The next section describes the data and the applied methodology. Section 3 reports the results. Finally, the main conclusions are presented.

2. Data and methodology

The period analysed runs from December 1999 to July 2015. The sample initially comprises 17,773 US domestic equity share-class funds. We aggregate multiple share classes of the same fund, a common practice in the literature. We then remove all the funds from the sample with no available data for size, return, sales and redemptions. This information is required to construct

both measures of fund flows. Our sample finally consists of 2,985 US domestic equity mutual funds. There is no survivorship bias since the sample includes both disappeared and new funds during the sample period.

For each fund we obtain from Morningstar the fund's name, fund Id, inception date, and fund objective. Since we want to estimate different net cash flows and to show their relation to other variables, we also download monthly information on total net asset (TNA), return, sales, and redemptions. Finally, we also obtain information on the annual expense ratio of each fund.

As commented above, in some previous studies, net cash flows have been indirectly estimated, as shown in equation (1).

$$Implied\ Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} * (1 + R_{i,t})}{TNA_{i,t-1}} \quad (1)$$

Where $Implied\ Flow_{i,t}$ is the estimated monthly net cash flow in relative terms that fund i experiences during period t . $R_{i,t}$ is the monthly return of that fund during period t , and $TNA_{i,t}$ refers to the total net assets of the same fund during period t . Two important assumptions are made in this method: the generated dividends are entirely reinvested in the fund, and cash flows occur in the last moment of the period.

We now calculate net cash flows directly. Thus, as shown in (2), we define the fund net cash flow as the total inflows minus the total outflows that occur in a mutual fund in the same period.

$$Net\ Cash\ Flow_{i,t} = \frac{Sales_{i,t} - Red_{i,t}}{TNA_{i,t-1}} \quad (2)$$

Where $Net\ Cash\ Flow_{i,t}$ is the monthly net cash flow in relative terms that fund i experiences during period t , $Sales_{i,t}$ is the total inflows made by investors of fund i during period t , and $Red_{i,t}$ refers to the total redemptions made by investors of fund i during period t .

Following Cashman *et al.* (2012), we eliminate from the sample observations that appear to contain data errors. Specifically, we remove observations in which the net flow, sales or redemptions exceed 70% of the size of the fund in the previous period. Additionally, in order to ensure a consistent comparison, we require the funds to present information on both fund flows' measures during each period.

Table 1
Characteristics of the mutual fund sample

	Mean	Standard deviation
TNA (\$millions)	754.78	3,178.88
Sales (\$millions)	334.93	353.81
Redemptions (\$millions)	334.53	359.64
Net Cash Flow (\$millions)	1.62	154.29
Implied Flow (\$millions)	31.74	79.33
Net Cash Flow (%)	2.99%	12.61%
Implied Flow (%)	8.66%	6.90%
Annualized Return %	6.11%	1.62%
Net Expense Ratio (%)	1.32%	0.72%

This table shows some descriptive statistics of our sample of 2,985 US equity mutual funds for the period December 1999-July 2015. *TNA* represents the assets of the fund under management and *Annualized Return* is the annualized monthly return of the fund. *Net Expense Ratio* is the annual net expense ratio borne by the fund. *Sales* and *Redemptions* describe the flows going into and out of the fund, respectively. *Net Cash Flow* and *Implied Flow* are the accurate and approximate net fund flow measures, respectively. The units of these characteristics are in parentheses.

Table 1 reports the descriptive statistics for the sample. The average fund experiences similar amounts of sales (334.93 millions of dollars) and redemptions (334.53 millions of dollars), but net flows seem to differ according to the methodology applied. Net cash flows estimated by equation (1) are greater (31.74 million dollars) than the funds' actual net cash flows (1.62 million dollars). In relative terms, these differences in the net flows also hold (8.66% for implied flows, and 2.99% for net cash flows). This information reveals important differences between the indirect and the direct estimate of net cash flows. Consequently, the results of the analysis of the

determinants of investors' demand for funds could differ when using different cash flows measures.

3. Results

3.1 Differences between net and implicit cash flow

In this section, we analyse whether there are significant differences between the two flow measures during the sample period. For robustness purposes, the same analysis is also run for four sub-periods related to different market conditions: two bear regimes (from December 1999 to December 2003, and from January 2008 to December 2009) and two bull regimes (from January 2004 to December 2007, and from 2010 to the end of the sample period).

We apply ordinary least squares to the time-series regressions for each fund in the sample. Previous studies (Edelen, 1999; Peng *et al.*, 2011; among others) also employ this approach. The advantage of using this methodology is that the estimates of the regressions are allowed to differ across funds, so we allow flows in each fund to respond differently to the explanatory variables. If the coefficients from the regressions were mainly negative (or positive) in most of the regressions, then the mean of these coefficients would be negative (or positive) and significantly different from zero.

Accordingly, we regress the estimated flows (*Implied Flow_{i,t}*) on the actual net cash flows (*Net Cash Flow_{i,t}*) as follows:

$$Implied\ Flow_{i,t} = \beta_{i,0} + \beta_{i,1}Net\ Cash\ Flow_{i,t} + \varepsilon_{i,t} \quad (3)$$

If there are no significant differences, we should obtain a close to zero intercept, a slope close to unity, and a very high coefficient of determination. However, if these are not the cases, results would show that there are important differences estimating fund flows, and a noise would be considered when calculating implied flows. Table 2 presents the results of this analysis.

Table 2
Regression of Implied Flow on Net Cash Flow

	December 1999- July 2015	December 1999- December 2003	January 2004- December 2007	January 2008- December 2009	January 2010- July 2015
	Mean p-value	Mean p-value	Mean p-value	Mean p-value	Mean p-value
Intercept	0.055 (0.000)	0.174 (0.000)	0.023 (0.261)	-0.002 (0.939)	0.037 (0.070)
Net Cash Flow	0.847 (0.000)	0.812 (0.000)	0.845 (0.000)	0.890 (0.000)	0.875 (0.000)
Number of funds	2,985	1,644	2,153	1,536	2,023
Adjusted R ²	0.737	0.693	0.743	0.831	0.801

This table reports the average of the coefficient estimates across the OLS time-series regressions for each fund in the sample. The dependent variable is the *Implied Flow* of the fund, defined as the estimated monthly net cash flow in relative terms. The explanatory variable is *Net Cash Flow*, measured as the net percentage fund flow using flows into and out of the fund in the same period. Results are reported for the whole period and the sub-periods considered. P-values (in parentheses), the number of funds and the average adjusted R² for each period are also reported.

Results show significant differences between the two measures of estimated cash flows. Regarding the main period, the adjusted R² is quite high (0.737), which suggests that implied flows are a good estimate of net cash flows. In contrast, the mean coefficient on the net cash flows (0.847) is positive but significantly lower than the unity, which is in line with the lower variability of implied flows observed in Table 1. The evidence is very similar for all of the sub-periods considered. In short, implied flows seem to be a good estimate of actual cash flows into and out of the fund, but also entail an implicit error in their calculation.

We next analyse the variables that can potentially create these differences. That is, the components that lead to an increase in the deviation of the two flow measures. Specifically, the variables that are involved in both fund flows methodologies: return, size, sales, and redemptions.

To find out how the different characteristics of a fund affect these differences in net cash flow estimates, we create a new variable, *Implied Excess Flow*, which we define as the absolute value of the difference between the implied flows and the actual net cash flows. Consequently, the higher the value of this variable, the higher the deviation generated through equation (1). In other words, the higher the error assumed when using implied flows.

Therefore, we regress the *Implied Excess Flow* of each fund on the aforementioned variables, as described in Equation (4):

$$Implied\ Excess\ Flow_{i,t} = \beta_0 + \beta_1 R_{i,t} + \beta_2 R_{i,t}^2 + \beta_3 LogTNA_{i,t} + \beta_4 Sales_{i,t} + \beta_5 Red_{i,t} + \varepsilon_{i,t} \quad (4)$$

Where $LogTNA_{i,t}$ refers to the size of fund i during period t , measured as the logarithm of the total net assets under the fund management.

On the one hand, we hypothesize that higher returns, in absolute terms, should increase the deviation between the two flow measures. Because of that, we also consider the square of the return as an explanatory variable. Thus, we expect β_1 not to be statistically significant (the greater effect of a positive return on the estimated flows would be diminished by the higher effect of a negative return). Nonetheless, β_2 may be positive and significantly different from zero. On the other hand, fund size should negatively affect the *Implied Excess Flows*, because, given a certain level of net cash flow, the more assets the fund manages, the lower this level of relative cash flows will be. Finally, sales (redemptions) may positively affect the differences in estimates of cash flows since the appreciation experienced (not experienced) during the period will be considered as an inflow (outflow) when using implied flows.

Table 3
Explaining the Implied Excess Flow

	Mean	p-value
Intercept	2.184	(0.001)
Return	0.002	(0.235)
Return ²	0.001	(0.005)
Logarithm TNA	-0.238	(0.005)
Sales	0.060	(0.000)
Redemptions	0.101	(0.000)
Number of funds	2,909	
Adjusted R ²	0.269	

This table reports the results of the coefficient estimates across the OLS time-series regressions for each fund in the sample. The dependent variable is the *Implied Excess Flow*, measured as the absolute value of the difference between the implied flows and the actual net cash flows. The table includes the mean and p-value (in parentheses) of the coefficients of the explanatory variables, namely, the return and the square of the return experienced by the fund in the period, the log of the total assets under management, and the sales and redemptions going into and out of the fund in the same period. The number of funds and the average adjusted R² are also reported.

Table 3 reports the results of this analysis. As we expected, results show that the coefficient on fund return is not significant (p-value of 0.235), so it seems that it does not contribute to the deviation in the two net cash flow estimates. However, both high positive and negative returns generate larger differences in the flow measures since β_2 is positive and statistically significant. Regarding the effect of the fund size on the *Implied Excess Flows*, results show that the coefficient of this variable is negative (-0.238) and statistically significant. It means that given a certain level of cash flow, the deviation among both measures is smaller when considering funds managing more assets, and so lower differences are generated. Finally, coefficients on the fund sales (0.060) and fund redemptions (0.101) are also significant, implying that greater levels of these variables lead to larger differences in the two fund flow estimates.

In short, results in Table 3 reflect that implied flows defined in (1) does not accurately estimate the real net flows experienced by a fund during a period, and it generates an error which is greater in smaller funds, and in the presence of higher levels of sales, redemptions, or return achieved by the fund.

3.2 Analysis of the determinants of fund cash flow

Consequently, and in view of the results of the previous section, the lack of precision in the calculation of net cash flows may also create an error in estimating the determinants of investors' fund flows. Therefore, we regress both flow measures on the variables that can affect the fund investors' demand, according to the previous literature: the return experienced by the fund during the previous period (*Return*), the risk borne by the fund portfolio (*Risk*), and the growth rate of net flow for all funds in the same objective as the fund in the previous period (*Lagged Objective Flow*). We also consider the *Lagged Cash Flows* in the analysis, in order to observe the persistence of this variable over time. Finally, we also consider some control variables, such as the fund size (*Log Lagged TNA*), the level of expenses (*Expense Ratio*), and the age of the fund since inception (*Age*).

On the one hand, we expect that the return and the lagged flows positively affect the net flows, since the return and the previous investments made by other investors can influence the fund investor's choices. On the other hand, we suppose that the risk borne by the portfolio negatively affects the net flows since we assume investors to be risk averse. We also hypothesize that the effect of the independent variables on the fund flows could be distorted when considering implied flows, due to the inherent error that this measure implies.

Table 4
Regression of the monthly fund flows on fund characteristics

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	68.163	(0.000)	57.823	(0.000)	10.340	(0.060)
Log Lagged TNA	-8.482	(0.000)	-7.269	(0.000)	-1.214	(0.016)
Expense Ratio	1.809	(0.310)	1.287	(0.472)	0.522	(0.838)
Age	-0.019	(0.000)	-0.012	(0.016)	-0.008	(0.283)
Return	0.067	(0.000)	0.066	(0.000)	0.001	(0.810)
Risk	-0.268	(0.000)	-0.177	(0.000)	-0.092	(0.044)
Lagged Objective Flow	0.107	(0.000)	0.117	(0.000)	-0.010	(0.775)
Lagged Cash Flow	0.159	(0.000)	0.196	(0.000)	-0.037	(0.000)
Number of funds	2,881		2,881			
Adjusted R ²	0.259		0.291		-0.032	(0.000)

This table shows the average results and their significance (in parentheses) of the coefficient estimates across the OLS time-series regressions for each fund in the sample. The first and second columns report the results of the regression of the *Implied Flows*, defined as the estimated monthly net cash flow in relative terms. The third and fourth columns show the results of the regression of the *Net Cash Flows*, measured as the net percentage fund flow using flows into and out of the fund in the same period. The fifth and sixth columns report the results of the mean differences of each coefficient and their statistical significance. The explanatory variables are the log of the funds under management in the previous period (*Log Lagged TNA*), the net expense ratio borne by the fund during the previous year (*Expense Ratio*), the age of the fund since inception, in months (*Age*), the monthly return of the fund in the previous period (*Return*), the risk of the fund measured as the standard deviation of the last 12 monthly returns (*Risk*), the growth rate of net flow for all funds in the same objective as the fund in the previous period (*Lagged Objective Flow*), and the growth rate of net flow for the fund in the previous period (*Lagged Cash Flow*), measured in the same terms as the dependent variable. The number of funds and the average adjusted R² are also reported.

Table 4 reports the results of this analysis, showing the average coefficients (and their significance) estimated through the regressions for each fund in the sample, as well as the mean differences between the two models. The number of funds and the adjusted coefficient of determination are also reported.

Results are consistent with our expectations. On the one hand, previous returns, past flows related to the fund's objective and those experienced in the portfolio have positive and significant effects on the level of net flows experienced by the fund during the following period, regardless of the net cash flow estimate. In addition, the risk borne by the fund's portfolio impacts negatively on both implied flows (coefficient of -0.268) and net cash flows (-0.177).

On the other hand, the comparison of the mean coefficients of the two models shows that some of the results differ when implied flows (as in (1)) are used. Firstly, there are statistically significant differences in the adjusted coefficient of determination: the adjusted R^2 for the net cash flows (0.291) is significantly higher than the adjusted R^2 for the implied flows (0.259). This implies that the model of actual fund flows is a better fit than the model that considers the approximate flows as the dependent variable. Moreover, there are differences regarding the coefficient of the explanatory variables. Firstly, the coefficient of lagged flows is significantly lower in the model of implied flows (0.159) than in the model of the net cash flows (0.196). This result indicates that implied flows underestimate the effect of the persistence of the fund flows. In addition, the effect of the fund size and the effect of the risk assumed by the portfolio on the fund flows are greater when considering implied flows.

Next, we wonder if this evidence is robust when we use data with a different frequency. If implied flows entailed an inherent error, we could expect this lack of precision to be higher when using two-quarterly information, for example. We therefore consider different windows to analyse the effect of the previous variables on the cash flows.

Table 5 presents the results for the mean coefficients of the explanatory variables and their significance when using quarterly (Panel A), two-quarterly (Panel B) and annual data (Panel C).

The number of funds and the average adjusted coefficient of determination are also reported.

Table 5
The determinants of quarterly, two-quarterly and annual fund flows.

Panel A: Quarterly flows

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	126.057	(0.000)	111.710	(0.000)	14.347	(0.108)
Log lag TNA	-15.823	(0.000)	-14.088	(0.000)	-1.735	(0.028)
Expense Ratio	6.402	(0.128)	5.899	(0.123)	0.503	(0.929)
Age	-0.004	(0.526)	-0.004	(0.608)	-0.001	(0.937)
Return	0.060	(0.000)	0.050	(0.000)	0.010	(0.156)
Risk	-0.293	(0.000)	-0.224	(0.000)	-0.068	(0.239)
Lagged Objective Flow	0.033	(0.540)	-0.011	(0.841)	0.044	(0.566)
Lagged Cash Flow	0.171	(0.000)	0.200	(0.000)	-0.030	(0.001)
Number of funds	2,213		2,213			
Adjusted R ²	0.330		0.349		-0.019	(0.015)

Panel B: Two-quarterly flows

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	492.488	(0.000)	157.479	(0.000)	335.009	(0.000)
Log lag TNA	-61.183	(0.000)	-19.043	(0.000)	-42.140	(0.000)
Expense Ratio	22.797	(0.001)	4.795	(0.260)	18.002	(0.023)
Age	0.112	(0.000)	0.014	(0.064)	0.098	(0.000)
Return	-0.150	(0.000)	0.016	(0.042)	-0.165	(0.000)
Risk	-0.829	(0.000)	-0.376	(0.000)	-0.454	(0.000)
Lagged Objective Flow	1.163	(0.000)	0.312	(0.002)	0.851	(0.000)
Lagged Cash Flow	0.190	(0.000)	0.206	(0.000)	-0.016	(0.232)
Number of funds	903		903			
Adjusted R ²	0.305		0.335		-0.030	(0.010)

Table 5 (Continued)**Panel C: Annual flows**

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	-1,565.324	(0.000)	-370.469	(0.000)	-1,194.855	(0.000)
Log lag TNA	164.365	(0.000)	39.826	(0.000)	124.539	(0.000)
Expense Ratio	95.254	(0.005)	21.136	(0.160)	74.118	(0.046)
Age	0.143	(0.012)	-0.023	(0.362)	0.166	(0.008)
Return	-0.001	(0.986)	0.052	(0.024)	-0.053	(0.451)
Risk	3.291	(0.000)	0.784	(0.065)	2.507	(0.001)
Lagged Objective Flow	-1.043	(0.115)	-0.885	(0.011)	-0.158	(0.832)
Lagged Cash Flow	-0.381	(0.000)	-0.024	(0.525)	-0.358	(0.000)
Number of funds	139		139			
Adjusted R ²	0.595		0.391		0.204	(0.000)

This table shows the average results and the significance (in parentheses) of the coefficient estimates across the OLS time-series regressions for each fund in the sample. Panel A, Panel B and Panel C refer to fund flows estimated in a quarterly, two-quarterly and annual basis, respectively. The first and second columns report the results of the regression of the *Implied Flows*, defined as the estimated net cash flow in relative terms. The third and fourth columns show the results of the regression of the *Net Cash Flows*, measured as the net percentage fund flow using flows into and out of the fund in the same period. The fifth and sixth columns report the results of the mean differences of each coefficient and their statistical significance. The explanatory variables are the log of the funds under management in the previous period (*Log Lagged TNA*), the net expense ratio borne by the fund during the previous year (*Expense Ratio*), the age of the fund since inception, in months (*Age*), the return of the fund in the previous period (*Return*), the risk of the fund measured as the standard deviation of the last 12 monthly returns (*Risk*), the previous growth rate of net flow for all funds in the same objective as the fund (*Lagged Objective Flow*), and the growth rate of net flow for the fund in the previous period (*Lagged Cash Flow*), measured in the same terms as the dependent variable. The number of funds and the average adjusted R² are also reported.

The evidence in Panel A is very similar to that in Table 4. On the one hand, previous returns and previous fund flows have positive and statistically significant effects on the fund flows during the following period, while the effect of the portfolio's risk is significantly negative. On the other hand, there are significant differences in the mean coefficients of some explanatory variables. For instance, the effects of fund size and of previous fund flows are significantly lower in the model in which the dependent variable is the implied flows (coefficients of -1.735 and -0.03, respectively).

Turning to Panel B, we have some very interesting results. Firstly, previous returns have a positive and statistically significant effect (coefficient of 0.014) on the actual net cash flows (as in (2)). Nevertheless, they seem to have a negative (coefficient of -0.150) effect on the implied flows, estimated as in (1). Moreover, the effect of the net expense ratio on the implied flows seems to be significantly positive (coefficient of 22.797), but it is non-significant (p-value of 0.260) when a more accurate measure of net cash flows is considered. Also, most of the differences in the mean coefficients from both models are statistically significant. The same evidence is found for the annual-based analysis. In other words, the distortion generated by implied flows is higher when longer windows are used in their estimation.

Overall, the evidence related to analyses for different windows indicates that the use of implied flows, despite being a good approximation of the actual net cash flows experienced by the fund, could lead to wrong conclusions on the determinants of the investors' flows, especially if they are estimated during longer periods (e.g., two-quarterly or annually).

3.3. Does the effect of the determinants of fund cash flow change during bullish and bearish periods?

To observe whether there are any differences in the estimates among different sub-periods, we distinguish between bearish and bullish periods, and study the effect of the determinants of investors' flows. Only monthly flows are studied due to the lack of sufficient information for a consistent analysis on a quarterly or annual basis.¹ Table 6 shows the results of these analyses. Specifically, Panel A and Panel C report the results for two bearish periods (2000-2003 and 2008-2009), while Panel B and Panel D present the results for two bullish periods (from 2004 to 2007, and from 2010 to the end of the sample period).

¹ For instance, the 2008-2009 period only covers eight quarterly observations.

Table 6
The determinants of fund flows during bullish and bearish periods

Panel A: Bearish period. January 2000–December 2003

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	132.497	(0.000)	77.203	(0.009)	55.294	(0.150)
Log lag TNA	-17.515	(0.000)	-13.699	(0.000)	-3.816	(0.093)
Expense Ratio	10.386	(0.475)	22.043	(0.219)	-11.657	(0.613)
Age	0.037	(0.061)	0.053	(0.008)	-0.016	(0.577)
Return	0.048	(0.000)	0.059	(0.000)	-0.010	(0.404)
Risk	-0.120	(0.175)	-0.004	(0.957)	-0.115	(0.345)
Lagged Objective Flow	0.396	(0.000)	0.270	(0.000)	0.126	(0.072)
Lagged Cash Flow	0.004	(0.731)	0.030	(0.001)	-0.026	(0.060)
Number of funds	1,278		1,278			
Adjusted R ²	0.258		0.270		-0.011	(0.261)

Panel B: Bullish period. January 2004–December 2007

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	180.238	(0.000)	138.747	(0.000)	41.491	(0.010)
Log lag TNA	-21.553	(0.000)	-17.032	(0.000)	-4.521	(0.001)
Expense Ratio	-5.479	(0.411)	-3.422	(0.593)	-2.057	(0.824)
Age	0.081	(0.000)	0.074	(0.000)	0.008	(0.738)
Return	0.108	(0.000)	0.087	(0.000)	0.021	(0.097)
Risk	-0.386	(0.000)	-0.183	(0.090)	-0.203	(0.165)
Lagged Objective Flow	0.207	(0.000)	0.208	(0.000)	-0.001	(0.980)
Lagged Cash Flow	0.097	(0.000)	0.127	(0.000)	-0.030	(0.003)
Number of funds	1,886		1,886			
Adjusted R ²	0.283		0.301		-0.018	(0.033)

Panel C: Bearish period. January 2008– December 2009

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	52.936	(0.044)	42.962	(0.112)	9.974	(0.791)
Log lag TNA	-10.351	(0.000)	-10.828	(0.000)	0.477	(0.714)
Expense Ratio	17.491	(0.414)	30.633	(0.166)	-13.142	(0.669)
Age	0.038	(0.048)	0.022	(0.250)	0.015	(0.573)
Return	0.089	(0.000)	0.101	(0.000)	-0.013	(0.130)
Risk	0.023	(0.649)	0.032	(0.521)	-0.009	(0.894)
Lagged Objective Flow	0.043	(0.118)	0.029	(0.299)	0.014	(0.710)
Lagged Cash Flow	-0.080	(0.000)	-0.071	(0.000)	-0.009	(0.444)
Number of funds	1,170		1,170			
Adjusted R ²	0.225		0.231		-0.006	(0.565)

Table 6 (Continued)**Panel D: Bullish period. January 2010– July 2015**

Independent Variables	Implied Flows		Net Cash Flows		Mean Differences	
	Mean	p-value	Mean	p-value	Mean	p-value
Intercept	99.440	(0.000)	82.797	(0.000)	16.643	(0.069)
Log lag TNA	-12.324	(0.000)	-10.321	(0.000)	-2.004	(0.009)
Expense Ratio	-1.240	(0.741)	-1.054	(0.788)	-0.186	(0.973)
Age	0.019	(0.024)	0.021	(0.011)	-0.002	(0.848)
Return	0.061	(0.000)	0.064	(0.000)	-0.003	(0.707)
Risk	-0.121	(0.004)	-0.105	(0.017)	-0.016	(0.790)
Lagged Objective Flow	0.106	(0.010)	0.098	(0.038)	0.008	(0.898)
Lagged Cash Flow	0.105	(0.000)	0.121	(0.000)	-0.017	(0.047)
Number of funds	1,898		1,898			
Adjusted R ²	0.184		0.196		-0.012	(0.067)

This table shows the average results and the significance (in parentheses) of the coefficient estimates across the OLS time-series regressions for each fund in the sample. Results are presented for each sub-period. The first and second columns report the results of the regression of the *Implied Flows*, defined as the estimated monthly net cash flow in relative terms. The third and fourth columns show the results of the regression of the *Net Cash Flows*, measured as the net percentage fund flow using flows into and out of the fund in the same period. The fifth and sixth columns report the results of the mean differences of each coefficient and their statistical significance. The explanatory variables are the log of the funds under management in the previous period (*Log Lagged TNA*), the net expense ratio borne by the fund during the previous year (*Expense Ratio*), the age of the fund since inception, in months (*Age*), the monthly return of the fund in the previous period (*Return*), the risk of the fund measured as the standard deviation of the last 12 monthly returns (*Risk*), the growth rate of net flow for all funds in the same objective as the fund in the previous period (*Lagged Objective Flow*), and the growth rate of net flow for the fund in the previous period (*Lagged Cash Flow*), measured in the same terms as the dependent variable. The number of funds and the average adjusted R² are also reported.

Results for the bullish periods are in line with those in Table 4. For instance, the higher the fund's previous returns, the higher the level of net flows attracted (the mean coefficients range from 0.061 to 0.108). Previous flows also have positive and statistically significant effects, and the risk borne by the fund's portfolio impacts negatively on both net cash flow measures. Moreover, the inherent error assumed by implied flows generates significant differences in the effect of some explanatory variables, such as previous fund size and previous fund flows.

Regarding bearish periods (Panel A and Panel C), we also find similar evidence for the effect of previous returns on the fund flows (their coefficient is significantly positive). In contrast, it seems

that the risk borne by the portfolio has no significant effect on the fund flows during these sub-periods. Also, the differences in the mean coefficients of both models are not significant at the 5% level.

In sum, implied flows, that is, flows indirectly estimated using data on fund size and return, seem to be a good measure of the actual net cash flows experienced by the fund. However, this measure implicitly assumes an error in its calculation. And this error can lead to differences in the estimate of the fund investors' response to some related variables, especially during bullish periods.

4. Conclusion

Explaining the variability in the cash flows of a mutual fund has attracted much attention from both professionals and academics. Accordingly, estimating the effect of some determinants on the fund investors' demand plays an important role in the mutual fund management. For instance, mutual fund returns in the portfolio attract investors' flows. In addition, previous cash flows into the fund also have positive and significant effects on investment decisions. In contrast, the effect of the risk borne in the portfolio is significantly negative, at least during bullish periods. Nevertheless, the effect of these determinants can change depending on which measure of net cash flows is used.

Many authors, because of the unavailability of the data for inflows and outflows in some countries, estimate net cash flows that occurred during a period using fund size and returns information. According to them, these implied flows correspond to the cash flows that are not due to dividends and capital gains. Notwithstanding, this measure implicitly assumes that all the flows occur at the end of the period, and that all dividends are reinvested in the fund. This is an approximation of the actual cash flows into and out of the mutual funds and, therefore, causes an inherent noise in their calculation.

This study shows that, although this method seems to be a good measure, there is indeed a deviation in this rough estimate of cash flows in relation to the actual fund flows. The higher the

fund return (in absolute terms), the greater the differences generated, presumably due to the no appreciation of the flows experienced by the fund. Moreover, smaller funds and funds experiencing higher levels of inflows and outflows are also proportional to this error in the flows estimate.

Accordingly, this rough measure causes an error when estimating the effect of the explanatory determinants of the fund flows, such as the return or the flows experienced by the fund. This inaccuracy is more important during bullish periods, especially when longer time horizons are considered in estimating the fund flows.

In conclusion, implied flows are a good approximation to the actual cash flows experienced by the fund during a period, especially when there is no information related to the fund inflows and outflows. Nevertheless, we have to consider that their calculation is not always accurate. And this lack of precision can lead to distorted results of the analysis where implied flows are considered.

References

- Andreu, L., and Sarto, J. L., 2016. Financial consequences of mutual fund mergers. *The European Journal of Finance*, 22, 529–550. doi: 10.1080/1351847X.2013.858055
- Barber, B. M., Odean, T., and Zheng, L., 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. *Journal of Business*, 78, 2095–2119. doi: 10.1086/497042
- Bhattacharya, U., Lee, J. H., and Pool, V. K., 2013. Conflicting family values in mutual fund families. *Journal of Finance*, 68, 173–200. doi: 10.1111/j.1540-6261.2012.01797.x
- Cashman, G. D., Deli, D. N., Nardari, F., and Villupuram, S., 2012. Investors do respond to poor mutual fund performance: Evidence from inflows and outflows. *Financial Review*, 47, 719–739. doi: 10.1111/j.1540-6288.2012.00346.x
- Christoffersen, S. E., Evans, R., and Musto, D. K., 2013. What do consumers' fund flows maximize? Evidence from their brokers' incentives. *Journal of Finance*, 68, 201–235. doi:

10.1111/j.1540-6261.2012.01798.x

Cooper, M. J., Gulen, H., and Rau, P. R., 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance*, 60, 2825–2858. doi:

10.1111/j.1540-6261.2005.00818.x

Edelen, R. M., 1999. Investor flows and the assessed performance of open-end mutual funds.

Journal of Financial Economics, 53, 439–466. doi: 10.1016/S0304-405X(99)00028-8

Feng, X., Zhou, M., and Chan, K. C., 2014. Smart money or dumb money? A study on the selection ability of mutual fund investors in China. *The North American Journal of*

Economics and Finance, 30, 154-170. doi: 10.1016/j.najef.2014.09.004

Friesen, G. C., and Sapp, T., 2007. Mutual fund flows and investor returns: An empirical examination of fund investor timing ability. *Journal of Banking and Finance*, 31, 2796–

2816. doi: 10.1016/j.jbankfin.2007.01.024

Gruber, M. J., 1996. Another puzzle: The growth in actively managed mutual funds. *Journal of*

Finance, 51, 783–810. doi: 10.1111/j.1540-6261.1996.tb02707.x

Guercio, D. D., and Reuter, J., 2014. Mutual fund performance and the incentive to generate alpha. *Journal of Finance*, 69, 1673–1704. doi: 10.1111/jofi.12048

Huang, J., Wei, K. D., and Yan, H., 2007. Participation costs and the sensitivity of fund flows to past performance. *Journal of Finance*, 62, 1273–1311. doi: 10.1111/j.1540-

6261.2007.01236.x

Ivković, Z., and Weisbenner, S., 2009. Individual investor mutual fund flows. *Journal of*

Financial Economics, 92, 223–237. doi: 10.1016/j.jfineco.2008.05.003

Jayaraman, N., Khorana, A., and Nelling, E., 2002. An analysis of the determinants and shareholder wealth effects of mutual fund mergers. *Journal of Finance*, 57, 1521–1551. doi:

10.1111/1540-6261.00468

- Keswani, A., and Stolin, D., 2008. Which money is smart? Mutual fund buys and sells of individual and institutional investors. *Journal of Finance*, 63, 85–118. doi: 10.1111/j.1540-6261.2008.01311.x
- Peng, C. L., Chen, M. L., Shyu, S. D., and Wei, A. P., 2011. When is money likely to be smart? Evidence from mutual fund investors in Taiwan. *Investment Analysts Journal*, 40, 13–25. doi: 10.1080/10293523.2011.11082534
- Sirri, E. R., and Tufano, P., 1998. Costly search and mutual fund flows. *Journal of Finance*, 53, 1589–1622. doi: 10.1111/0022-1082.00066
- Zhao, X., 2005. Exit decisions in the US mutual fund industry. *Journal of Business*, 78, 1365–1401. doi: 10.1086/430863
- Zheng, L., 1999. Is money smart? A study of mutual fund investors' fund selection ability. *Journal of Finance*, 54, 901–933. doi: 10.1111/0022-1082.00131