Available online at www.sciencedirect.com
SCIENCE DIRECT
Finance Research Letters 1 (2004) 154–170
Finance Research Letters
www.elsevier.com/locate/frl

Decomposing the persistence of international equity flows

Kenneth A. Froot [a,*], Jessica Tjornhom Donohue [b]

[a] *Harvard University, Cambridge, MA, USA*
[b] *State Street Associates, Cambridge, MA, USA*

Received 4 June 2004; accepted 14 June 2004

Abstract

The portfolio flows of institutional investors are widely known to be persistent. What is less well-known, however, is the source of this persistence. One possibility is the 'informed trading hypothesis:' that persistence arises from autocorrelated trades of individual investors who believe they have information about value and who face an imperfectly liquid market. Another possibility is that there are asynchroneities with respect to investment decisions across funds, across investments, or both. These asynchroneities could be due to wealth effects (across investments for a single fund), investor herding (across funds for a single investment), or generalized contagion (across funds and across investments). We use daily data on institutional flows into 21 developed countries by 471 funds to measure and decompose aggregate flow persistence. We find that the informed trading hypothesis explains about 75% of total persistence, and that the remaining amount is attributable entirely to cross-fund own-country persistence. While asynchroneities across funds investing in the same country are important, asynchroneities across countries, either within a given fund, or across funds, are not important. The cross-fund flow lags we identify might result from different fund investment processes, or from some funds mimicking others' decisions. We reject the hypothesis that wealth effects explain persistence.

Keywords: International investment; Investor behavior; Portfolio investment; Portfolio flows

* Corresponding author.
E-mail addresses: kfroot@hbs.edu (K.A. Froot), jtdonohue@statestreet.com (J.T. Donohue).

1544-6123/$ – see front matter
doi:10.1016/j.frl.2004.06.002

1. Introduction

If there were a single characteristic that describes the portfolio flows of institutional investors, it would seem to be persistence. A number of authors in widely varying contexts have found this result. Studies looking at individual Asian equity markets have found persistence in foreigner's flows (e.g., Choe et al. 1999, 2001; Kim and Wei, 2000, and Seasholes, 2000). Studies of institutional investor flows across a number of markets have also found strong persistence on country or regional levels (see Richards, 2002; Froot et al., 2001, and Kaminsky et al., 2000). In addition, studies of mutual fund flows in the US show persistence at relatively high frequencies. These empirical findings are robust, not only across different databases, but also after conditioning on other variables. For example, a number of these studies demonstrate that flow persistence does not appear strongly diminished by controlling for past returns, even though past returns and past flows are correlated.

Persistence in net purchases by institutional investors would not seem very important if institutional flows had no stable relationship with prices. However, this is not the case. Considerable research has documented that current returns are strongly positively correlated with institutional flows, that current institutional flows tend to react positively to past returns, and that current flows are positively predictive of future returns. These associations suggest that institutional flows may be bound up with return momentum in equities, which is, to date, one of the broadest and most prevalent empirical anomalies in asset pricing.[1]

While persistence in institutional flow is clear in the data, the underlying mechanisms driving it are not. In models of informed trading (such as Kyle, 1985), informed order flow is conditionally autocorrelated. Traders with positive information attempt to disguise it, rationing their purchases at any given time and deferring some into the future, in order to reduce total price impact. The same holds if the information is merely perceived, rather than actual, i.e., if there is trader overconfidence.

This type of mechanism suggests that the persistence in institutional order flow is related to the scope of (real or perceived) information. Under what we call the informed trading hypothesis, each fund's purchases of a given country's equities are likely to be own-autocorrelated, provided that the funds attain private information about the prospects of firms within that country. Autocorrelation emerges because individual fund managers have either company- or country-specific information and dispense it slowly and optimally into prices.

However, such *own-fund own-country* persistence is likely to be only one piece of the total persistence in aggregated flow data. The most general alternative explanation is lack of simultaneity across investors. Investors may process information at different rates, wait for different signals over time, or have different lags associated with infrastructure,

[1] On the profitability and breadth of momentum effects see Jegadeesh and Titman (1993) and Rouwenhorst (1998). A number of papers study the relationship between institutional flows and equity returns both in the US and internationally. See Cohen et al. (2001), Froot et al. (2001), Froot and Ramadorai (2004), Grinblatt and Keloharju (2000), Grinblatt et al. (1995), Lakonishok et al. (1992), Nofsinger and Sias (1999), Richards (2002), Wermers (1999, 2002).

bureaucracy, or decision-making. These lags may operate across investments within a single fund—for example, shocks to fund wealth may result in rebalancing transactions that take time to complete (see Kyle and Xoing, 2001). Lags may also occur across funds, as one manager may respond to others' decisions, often cited as the basis for 'herding.'

These alternative mechanisms induce persistence in aggregated flows even in the absence of own-fund own-country persistence due to the informed trading hypothesis. Clearly, if they are present, these alternative mechanisms will need to be better understood.

The purpose of this paper is to decompose persistence in institutional investor equity flow across funds and countries. We measure this decomposition and use it to test for deviations from the informed trading hypothesis. These deviations can provide considerable information on the mechanisms behind flow persistence. A good analogy to what we do comes from the literature decomposing the persistence of equity index returns. That work shows virtually all of the (positive) autocorrelation of historical US equity indexes to be attributable to non-contemporaneous cross-stock return correlations—own-stock autocorrelations are zero. While the literature has not arrived at a definitive model of the leads and lags of individual stocks versus the index, the decomposition is crucial for testing a variety of theories (e.g., non-trading, short sales constraints, informational inefficiencies, etc.). Our problem in flows is considerably richer, since in addition to the cross-country dimension we have an added cross-fund dimension.[2]

To preview our results, we find that after conditioning on own-fund own-country persistence, the magnitude of cross effects is still statistically and economically an important component of overall persistence. We find that *cross-fund own-country* persistence, the largest additional source, contributes approximately 25% of total persistence. There is only very weak evidence of persistence of flows across countries (from *own-fund cross-country* and *cross-fund, cross-country* persistence). There is 'excess' persistence in the data, but it comes almost exclusively from cross-fund, own-country effects. We therefore reject the informed trading null in favor of the alternative that there is a slow rippling of flow across funds in a given country, but not within a fund or across funds across countries. Thus, we find no evidence to support the Kyle and Xoing (2001) hypothesis that wealth effects within a fund result in cross-country non-contemporaneously correlated flows (sometimes invoked to explain 'contagion').

Empirically, our findings appear quite robust. With over 12 million fund/country/day flow data points, statistical power is not really an issue; all of the above rejections of the null are very highly statistically significant, while the failures to reject reflect extremely (economically) small point estimates. Second, the results are essentially unchanged whether persistence is measured in daily or weekly data. Third, there is no impact on the results of conditioning on other variables that have been identified as important short run determinants of flows, such as lagged returns.

The rest of the paper is structured as follows. Section 2 discusses the decomposition. Section 3 describes the data. Section 4 provides some basic descriptive statistics. Section 5 discusses the main results and Section 6 concludes.

---

[2] See, for example, Lo and MacKinlay (1988) and Froot and Perold (1995).

2. Decomposition

We begin with an established fact, and one that we further confirm in our data: that institutional net order flow aggregated across funds and countries is highly autocorrelated:

$$f_t = \delta + \alpha f_{t-\tau} + \varepsilon_t, \quad \alpha > 0, \tag{1}$$

where f is net dollars of flow (dollar bought–dollars sold) aggregated across funds ($i = 1, \ldots, I$) and countries ($k = 1, \ldots, K$), $f_t = \sum_i \sum_k f_{i,k,t}$ and normalized in some way.

With respect to normalization of the flows there is no clearly dominant solution. One approach, used commonly, is to normalize the underlying value of flow by country market capitalization as a way of controlling for differences in market capacity, i.e., $f_{i,k,t} = F_{i,k,t}/M_{k,t-1}$, where $F_{i,k,t}$ is the dollar amount of net flow into country k by fund i, summed across all transactions on date t, and $M_{k,t}$ is the dollar market capitalization of the kth country.

While it is typical to weight flows across managers by dollar amount, we can use our disaggregated data to weight flow across managers differently. One simple approach that puts different markets and funds on a more similar footing is a digital normalization. It treats all fund/countries with net buys (sells) on a given day as having the same flow magnitude, i.e., $f^{\mathrm{d}}_{i,k,t} = \mathrm{sign}(F_{i,k,t}/M_{k,t-1})$, where sign(·) returns either 1, 0, or −1. A second approach to normalization uses the net buy count for each fund/country/date. Letting $B_{i,k,t}$ and $S_{i,k,t}$ represent the number of buys and sells, respectively for each fund/country/date, we define a flow count normalization to be $f^{\mathrm{c}}_{i,k,t} = (B_{i,k,t} - S_{i,k,t})/(B_{i,k,t} + S_{i,k,t})$. In much of what follows, we rely on $f^{\mathrm{d}}_{i,k,t}$ and $f^{\mathrm{c}}_{i,k,t}$, since these are comparably scaled across both funds and countries.

Once we have chosen a particular normalization, we next need to characterize the sources of persistence. This becomes a four-dimensional problem if we want to characterize generally the non-contemporaneous cross correlation between $f_{i,k,t}$ and $f_{j,l,t-\tau}$. To be specific, the τth-order normalized cross covariance is given by $\rho^{\tau}_{i,k,j,l} = \mathrm{cov}(f_{i,k,t}, f_{j,l,t-\tau})/\mathrm{var}(f_t)$, and the corresponding covariance matrix by $\Gamma(1)$. The τth-order autoregressive coefficient of total flow above, $\alpha(\tau)$, is given by

$$\alpha(\tau) = \sum_i \sum_j \sum_k \sum_l \frac{\mathrm{cov}(f_{i,k,t}, f_{j,l,t-\tau})}{\mathrm{var}(f_t)}. \tag{2}$$

In order to make some headway here in reducing the dimensionality of the problem, we divide up these terms. We can afford to do this: since we employ data on 471 funds and 21 countries, $\Gamma(\tau)$ has over 97 million elements—$(471 \times 21)^2$—for a single lag, τ.

We therefore divide things using the simple distinction between own versus cross correlations in each dimension. This brings us down to just 4 components, two in each dimension. Specifically, we group the $\Gamma(\tau)$ matrix as

$$\alpha(\tau) = \sum_k \sum_i \left(\frac{\mathrm{cov}(f_{i,k,t}, f_{i,k,t-\tau})}{\mathrm{var}(f_t)} + \sum_{j \neq i} \frac{\mathrm{cov}(f_{i,k,t}, f_{j,k,t-\tau})}{\mathrm{var}(f_t)} \right.$$

$$+\sum_{l\neq k}\frac{\mathrm{cov}(f_{i,k,t},\, f_{i,l,t-\tau})}{\mathrm{var}(f_t)}+\sum_{j\neq i}\sum_{l\neq k}\frac{\mathrm{cov}(f_{i,k,t},\, f_{j,l,t-\tau})}{\mathrm{var}(f_t)}\Bigg).$$

To save space, assume that each of the covariances above is constant, so that we can estimate a single parameter for each. That is, we impose the following four restrictions, each corresponding to a specific type of covariation:

(1) Equal *own-fund own-country* covariations:

$$\frac{\mathrm{cov}(f_{i,k,t},\, f_{i,k,t-\tau})}{\mathrm{var}(f_t)}=\alpha_{\mathrm{oo}}(\tau)\quad \text{for all } i,k \text{ pairs.}$$

Own-fund, own-country persistence is probably easiest to interpret. It comes from persistence in each fund's purchases of a country's equities. For example, Japanese equity inflows may be persistent because a given fund's purchases of Japanese equities today will on average continue for several days. Own-fund own-country persistence is what we would expect from an informed investor in the Kyle (1985) model. It is also what we would expect from an overconfident investor, who behaves as though he has information.[3]

(2) Equal *cross-fund own-country* covariations:

$$\sum_{j\neq i}\frac{\mathrm{cov}(f_{i,k,t},\, f_{j,k,t-\tau})}{\mathrm{var}(f_t)}=\alpha_{\mathrm{co}}(\tau)\quad \text{for all } i\neq j \text{ and for all } k.$$

This component is driven by non-synchronized purchases across funds investing in a given country's equities. For example, suppose a given fund buys Japanese equities today. While that fund may not on average purchase more Japanese equities tomorrow, other funds may tend to purchase Japanese equities at that time. We call this *cross-fund own-country* persistence.

(3) Equal *own-fund cross-country* covariations:

$$\sum_{l\neq k}\frac{\mathrm{cov}(f_{i,k,t},\, f_{i,l,t-\tau})}{\mathrm{var}(f_t)}=\alpha_{\mathrm{oc}}(\tau)\quad \text{for all } k\neq l \text{ and for all } i.$$

Flows into a given country from a given fund may be correlated with past flows into other countries from the same fund. A fund buying Japanese equities today might buy Australian equities tomorrow. Reasons for this source of persistence include a substitution effect towards Australian equities as Japanese equity prices rise, an implementation lag in getting to Australian equities, an emerging appreciation that the news for Japan also may apply to Australia, etc.

(4) Equal *cross-fund cross-country* covariations:

$$\sum_{j\neq i}\sum_{l\neq k}\frac{\mathrm{cov}(f_{i,k,t},\, f_{j,l,t-\tau})}{\mathrm{var}(f_t)}=\alpha_{\mathrm{cc}}(\tau)\quad \text{for all } i\neq j \text{ and } k\neq l.$$

[3] We do not test in this paper the information content of institutional investor trades, so we are agnostic here about whether persistent trades are the result of information or overconfidence.

This is the most dispersed form of persistence. Purchases of a given country by one fund may over time diffuse toward purchases by other funds of other countries. Peer and herding issues may be important here as well, to the extent they span investment opportunities other than countries. For example, suppose that funds focus on diversifying across corporate sectors, rather than countries. Then cross-fund delays in investing (due to either implementation issues or peer concerns) would necessarily show up as cross-country delays as well.[4]

Together, these four restrictions can be used to decompose the aggregate autocorrelation of total flows, shown in Eqs. (1) and (2):

$$\alpha(\tau) = IK\big(\alpha_{oo}(\tau) + \alpha_{co}(\tau) + \alpha_{oc}(\tau) + \alpha_{cc}(\tau)\big). \tag{3}$$

These four components of α can be estimated using OLS in the individual regressions

$$f_{i,k,t} = c + a(L)x + \varepsilon_t, \tag{4}$$

where $a(L)$ takes on the values, $\alpha_{oo}(\tau)$, $\alpha_{oc}(\tau)$, $\alpha_{co}(\tau)$, and $\alpha_{cc}(\tau)$, when x takes on the values

$$c_{oo} f_{i,k,t-\tau}, \quad c_{oc} \sum_{j \neq i} f_{j,k,t-\tau}, \quad c_{co} \sum_{l \neq k} f_{i,l,t-\tau}, \quad \text{and} \quad c_{cc} \sum_{j \neq i} \sum_{l \neq k} f_{j,l,t-\tau},$$

respectively, and where the c's are constants of proportionality such that in all four cases the standard deviation of x equals that of aggregate flow, f_t.

2.1. *Is flow persistence driven entirely by informed or overconfident trading?*

This decomposition provides perspective on the magnitude of the individual own- and cross-effects driving aggregate flow persistence. We need to go an additional step, however. The informed/overconfident trader hypothesis that we discuss above suggests that own-purchases are serially correlated. Traders get slowly into positions and the magnitude of their trades is a function of the perceived difference between value and price. In the continuous auction environment of Kyle (1985), market depth is constant. In expectation, as the informed trader pushes price towards perceived value, trade size declines. In this sense flow is stationary and persistent with respect to shocks to perceived value.

If we take the informed trader story as our null hypothesis, we would predict that cross-country and cross-fund persistence emerge as a result. These additional sources of persistence emerge because perceived opportunities may be contemporaneously correlated across countries. In the presence of own-fund own-country persistence, contemporaneous correlation across funds and countries will translate into non-contemporaneous correlation.

To see this, take the simple case in which flows for a given fund/country are autoregressive and stationary, and have iid news (or overconfidence) shocks:

$$f_{i,k,t} = \theta_{i,k}(L) f_{i,k,t-1} + \xi_{i,k,t}, \tag{5}$$

[4] There is increasing evidence that sector allocations are as important, or even more important than country allocations for diversifying risk. See, for example, VanRoyen and Page (2002).

where L is the lag operator. Given stationarity, it follows that $f_{i,k,t}$ can be written as a moving average process, $f_{i,k,t} = \phi_{i,k}(L)\xi_{i,k,t}$, where $\phi_{i,k}(L) = (1 - \theta_{i,k}(L))^{-1}$. Even though the ξ shocks are serially uncorrelated across all funds and countries (i and k), they may be contemporaneously correlated across both funds and countries. As a result, even small amounts of contemporaneous correlation between $\xi_{i,k,t}$ and $\xi_{j,l,t}$ can generate important non-contemporaneous cross-country and/or cross-fund correlations between $f_{i,k,t}$ and $f_{j,l,t-\tau}$ under the informed/overconfident trader hypothesis.

Notice, however, that if we can control properly for the own-autoregressive correlations in $f_{i,k,t}$, then the remaining own-flow components will be uncorrelated across funds, countries, and time. That is, after controlling for the own-autoregressive part of $f_{i,k,t}$, we are left with $\xi_{i,k,t}$. These own-flow innovations are uncorrelated with past own- and cross-flow innovations. That is, $\xi_{i,k,t}$ is uncorrelated with $\xi_{j,l,t-\tau}$, for all values of i, j, k, l, and $\tau > 0$.

Consider, then the regression

$$f_{i,k,t} = c + a_{oo}(L) f_{i,k,t-1} + \frac{a_{co}(L)}{I-1} \sum_{j \neq i} f_{j,k,t-1} + \frac{a_{oc}(L)}{K-1} \sum_{l \neq k} f_{i,l,t-1} + \frac{a_{cc}(L)}{(I-1)(K-1)} \sum_{j \neq i} \sum_{l \neq k} f_{j,l,t-1} + \varepsilon_{i,k,t}, \tag{6}$$

where we have made the coefficients easier to compare with one another by dividing by the number of funds and countries over which we sum (i.e., by $(I - 1)$ and $(K - 1)$, respectively). The informed trader hypothesis suggests that $a_{oo}(L) > 0$, and that $a_{co}(L) = a_{oc}(L) = a_{cc}(L) = 0$. Our alternative hypotheses are that one or more of these latter coefficients are different from zero. These coefficients represent the extent to which there is excess covariation in flows across funds and countries. Essentially, if these latter coefficients are different from zero, then there must be some other source of flow correlation beyond what is driven by the informed trader hypothesis.

For example, suppose that $a_{co}(L) > 0$, so there is excess cross-fund, own-country persistence. This suggests that some funds react to the same news as other funds with a lag when investing in country k. As mentioned above, this lag may be driven by implementation and decision-making lags and delays. Note, however, that unlike the total covariance, $\alpha_{co}(L)$, $a_{co}(L)$ is a partial covariance that controls are other sources of persistence.

For estimation, we use a version of Eq. (6) that allows for different own-fund own-country persistence profiles across countries, i.e., we allow a_{oo} to vary with k. We also ran versions of Eq. (6) including additional terms of lagged own-country returns and US returns, based on evidence that past returns help forecast flows.[5] Inclusion of returns had an extremely tiny economic and negligible statistical effect on both flow coefficients and R^2. Thus, returns may be important, but their absence does not cloud any conclusions about flow persistence. To save space, we make these additional results available on request.

[5] See papers by Froot et al. (2001), Richards (2002), and Choe et al. (1999) and the references therein for evidence on the predictability of flows by returns.

3. Data

The flow data used in our analysis are derived from proprietary data provided by State Street Corporation. State Street is the world's largest global custodian, with over $9 trillion of assets under custody. We extract data for a set of 930 distinct funds (without names or identifying characteristics to protect anonymity) from a total of almost 10,000 funds, using the criteria that a fund must trade equities incorporated in 21 or more distinct countries. Because our focus is on active, not passive funds, we reduce the set of funds by choosing only those that, when they are active in the dataset, trade at least 75% of the days during their active period. We focus on a set of 21 developed markets leaving us a sample of 471 funds.

Our country designations are somewhat unusual, in that they are driven by an equity issuer's country of incorporation. This definition allows us to include the trading ADRs and GDRs, categorized according to the company's country of incorporation, rather than according to where the security is traded. We focus on the 21 developed countries based on the number of transactions that exist in the reduced dataset.[6] The set of countries includes Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, New Zealand, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the United States. With 2166 days in the sample since 1994, our ultimate data set has over 12 million observations of net flows on a given day, into a given country, by a given fund.

In addition to the flow data, we also employ equity market returns. Returns are calculated using MSCI equity indices for each country used in our sample.

4. Descriptive statistics

Figure 1 provides a 'heatmap' of fund trading by country. It shows the number of trades for each fund/country over the sample period, relative to the maximum number of trades for that fund in any country over the sample across all countries. It is clear that the major developed countries account for a large percentage of the trading. Several developed countries, however, have very sparse transactions, including Belgium and Denmark.[7]

Tables 1.1–1.3 report some descriptive statistics for our net flow measures: dollar flows; digital indicator flows (1 for inflow, 0 for no flow, −1 for outflow); and buy–sell ratio flows (count of buys less sells normalized by buys plus sells). There are several points worth making.

First, the mean net flow by a single fund into a single country on a single day is just over $6400. Naturally, because this is a net flow, it is near zero. The daily standard deviation of own-fund own-country flows is much larger, approximately $1,198,840. The mean net inflow while small is nevertheless highly statistically significant. Thus, during our

[6] If there were less than a total of 2000 transactions for a given country of incorporation using the original set of 930 funds, it was dropped from our data.

[7] Recall that our country definition of each stock refers to the country of incorporation.

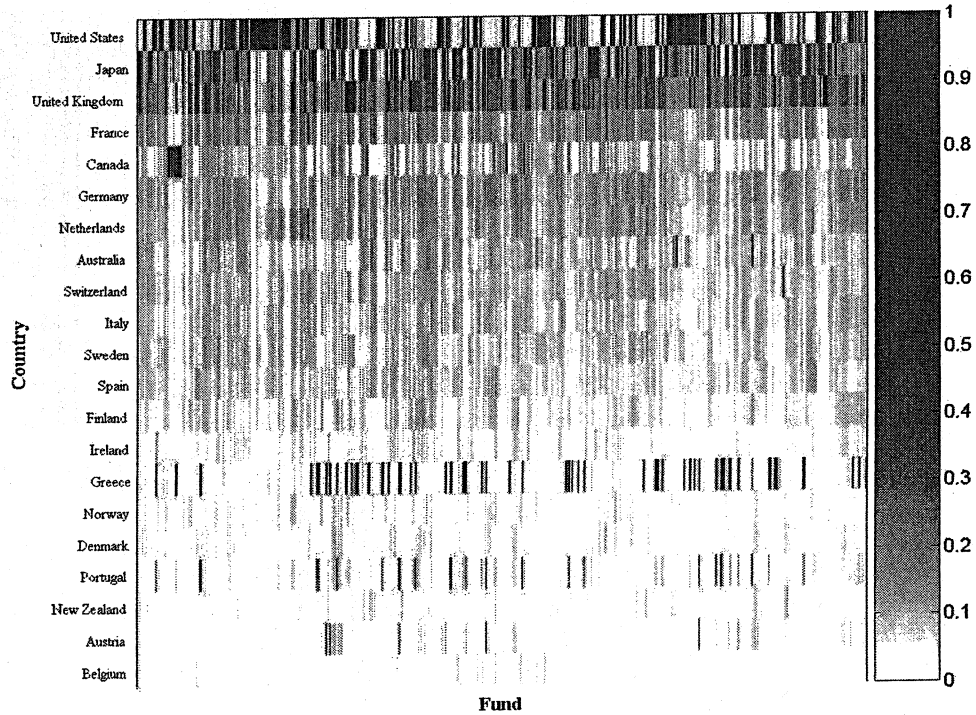

Fig. 1. Number of trades by individual funds for each country (relative to individual fund total number of trades over all countries).

Table 1.1
Descriptive statistics for net US dollar flows

	$F_{i,k,t}$	$\sum_{j\neq i} F_{j,k,t-\tau}$	$\sum_{l\neq k} F_{i,l,t-\tau}$	$\sum_{j\neq i}\sum_{l\neq k} F_{j,l,t-\tau}$
μ^*	6.40	1,988.56	128.08	39,771.22
σ^*	1,198.84	30,696.54	6042.06	160,352.92
N	12,728,330	12,728,330	12,728,330	12,728,330
ρ_1	0.1208	0.2514	0.1813	0.3089
ρ_2	0.0291	0.0737	0.0313	0.0799

Notes. The table provides the summary data on the net flows of institutional investors, across funds and countries. The flows cover 21 developed markets. In the first column we report for fund i and country k at time t $F_{i,k,t}$, the net US dollar net flow. The second column sums the net dollar flows over all funds j not equal to i for a given country k at each time period. The third variable gives, for fund i, the sum of all of its flows into countries other than k. Finally, the last term is the sum of flows over all country and fund pairs that do not include fund i and country k. The summary data we report include the mean, μ, the standard deviation, σ, and the first- and second-order autocorrelations, ρ_1 and ρ_2, stacking the data across i and k. The first and second autocorrelations are calculated by regressing the variable in question on two lags of itself for each (i, k) pair and then averaging across all active funds in country k. Mean and standard deviations are in thousands of US dollars.

period, there are net mean inflows recorded in our data; domestics on average are selling to international investors around the world.

Table 1.2
Descriptive statistics for digital signal

	$f^{d}_{i,k,t}$	$\sum_{j\neq i} f^{d}_{j,k,t}$	$\sum_{l\neq k} f^{d}_{i,l,t}$	$\sum_{j\neq i}\sum_{l\neq k} f^{d}_{j,l,t}$
μ	0.0049	1.3128	0.0974	26.2569
σ	0.3380	16.1972	1.9009	83.0693
ρ_1	0.2669	0.3804	0.2693	0.3559
ρ_2	0.0920	0.1596	0.0897	0.1075

Notes. The table provides the summary data on the sign or digital signal of net flows of institutional investors, across funds and countries where a net inflow is represented as a +1, a net outflow as −1, and no flow is 0. The flows cover 21 developed markets. In the first column we report for fund i and country k at time t, f^{d}_{ikt}, the net US dollar flow signal for country k. The second column sums the net dollar flow signals over all funds j not equal to i for a given country k at each time period. The third variable gives, for fund i, the sum of all of its flow signals into countries other than k. Finally, the last term is the sum of flow signals over all country and fund pairs that do not include fund i and country k. The summary data we report include the mean, μ, the standard deviation, σ, and the first- and second-order autocorrelations, ρ_1 and ρ_2, stacking the data across i and k. The first and second autocorrelations are calculated by regressing the variable in question on two lags of itself for each (i, k) pair and then averaging across all active funds in country.

Table 1.3
Descriptive statistics for the buy–sell ratio

	$f^{c}_{i,k,t}$	$\sum_{j\neq i} f^{c}_{j,k,t}$	$\sum_{l\neq k} f^{c}_{i,l,t}$	$\sum_{j\neq i}\sum_{l\neq k} f^{c}_{j,l,t}$
μ	0.0044	1.1741	0.0880	23.4816
σ	0.3033	15.1090	1.7771	78.4616
ρ_1	0.2887	0.3957	0.3069	0.4094
ρ_2	0.0992	0.1727	0.1000	0.1250

Notes. The table provides the summary data on the buy–sell ratio, the ratio of the number of buy transactions minus the number of sell transactions as a percentage of total buy and sell transactions, of net flows of institutional investors, across funds and countries. The flows cover 21 developed markets. In the first column we report for fund i and country k at time t, f^{c}_{ikt}, the buy–sell ratio for country k. The second column sums the buy–sell ratio over all funds j not equal to i for a given country k at each time period. The third variable gives, for fund i, the sum of all of its buy–sell ratios into countries other than k. Finally, the last term is the sum of buy–sell ratios over all country and fund pairs that do not include fund i and country k. The summary data we report include the mean, μ, the standard deviation, σ, and the first- and second-order autocorrelations, ρ_1 and ρ_2, stacking the data across i and k. The first and second autocorrelations are calculated by regressing the variable in question on two lags of itself for each (i, k) pair and then averaging across all active funds in country k.

Second, as expected, flows are persistent, as suggested by the partial autocorrelation coefficients in Tables 1.1–1.3. Both first- and second-order autocorrelations are consistently and statistically positive. For own-fund own-country flows, these autocorrelations are about 12%, which is not economically very large, but given the number of data points is extremely significant (the standard error is much less than 1%). It is interesting to note, however, that the own-fund own-country autocorrelations are considerably higher (27–29%) in the digital and buy/sell count flow indicators than in the raw flows themselves. This is because the scale, but not the direction, of dollar transactions, even for a given fund and given country jumps around considerably in the data. As a result of these 'outlier' datapoints, directional indicators appear more persistent.

The third point to make is that the persistence is greater for higher levels of aggregation across funds, countries or both. We might expect this because larger aggregations create greater scope for cross-persistence to emerge. For example, for the digital indicator across all countries, own-fund own-country first-order autocorrelation is 26.7%, own-fund cross-country autocorrelation is 26.9%, cross-fund own-country autocorrelation is 38.0%, and cross-fund cross-country is 35.6%.[8]

5. Results

Before interpreting the regressions results, we note that the standard errors are simple OLS. Driving this choice is that we have very many data points, over 21 million, so that many variables appear extremely statistically significant, with t-statistics that range from 10 to 1000. Adjustments that are often made to OLS standard errors to account for cross-sectional or autocorrelations of the residuals are unlikely to reverse t-statistics of this magnitude, even in the presence of strong correlations. Indeed, here we have the presence of only weak contemporaneous correlation, and, with the use of lagged variables to eliminate autocorrelation, very weak serial dependence. As a consequence, we report OLS standard errors, but interpret them conservatively. We take the informal view that a variable that cannot achieve a t-statistic of, say, 4, with the power of 21 million observations is probably too small to matter, so that, as in many large sample studies, marginal statistical significance is immaterial.

The results of our decomposition of aggregate multi-fund, multi-country flows, $f_{i,k,t}$ are shown in Tables 2.1–2.2. The correlation of aggregate flows is approximately, 40% with a standard deviation of approximately 0.4%. This is about the same level of flow persistence found by Froot et al. (2001) and Richards (2002) for international investors.

Table 2.1 then shows how the 0.40 total breaks down across the four different sources of flow persistence. Much as in the return literature, the vast size of the cross section (both

Table 2.1
Covariance decomposition

		$\alpha_{oo}(\tau)$	$\alpha_{co}(\tau)$	$\alpha_{oc}(\tau)$	$\alpha_{cc}(\tau)$
1st order	40.4722	2.7266	21.9617	1.9247	13.8591
Percentage		6.74%	54.26%	4.76%	34.24%

Notes. This table reports the decomposition of total flow autocorrelation into four components: own lag, lagged cross-fund own-country signals, lagged own-fund cross-country signals, and cross-fund cross-country signals. The decomposition is based on the equation

$$\alpha(\tau) = \big(\alpha_{oo}(\tau) + \alpha_{oc}(\tau) + \alpha_{co}(\tau) + \alpha_{cc}(\tau)\big).$$

We use digital signals of underlying net flows for this decomposition. The results are obtained by regressing a single lag of each of the four component variables on, $f^{d}_{i,k,t}$, the net US dollar flow signal for country k. We also report the composition in terms of percentages.

[8] Since we have many more funds than countries in our data, cross-fund own-country aggregations tend to be more highly aggregated than cross-country own-fund aggregations.

Table 2.2
Autoregressive behavior of digital equity flow signals

	f_t	$a_{oo}(L)$	$a_{co}(L)$	$a_{oc}(L)$	$a_{cc}(L)$
1st order	0.3466	0.2649	0.2985	0.1224	0.0961
	(0.0047)	(0.0003)	(0.0023)	(0.0011)	(0.0088)
2nd order	0.0733	0.1044	0.1310	0.0237	0.0014
	(0.0049)	(0.0003)	(0.0024)	(0.0011)	(0.0093)
3rd order	0.1236	0.0753	0.1062	0.0237	0.0585
	(0.0047)	(0.0003)	(0.0023)	(0.0011)	(0.0088)
R^2	0.2927	0.1176	0.0037	0.0019	0.0002
SE	14.8410	0.3170	0.3369	0.2934	0.3375
N	45,465	12,728,331	12,728,331	12,728,331	12,728,331

Notes. The table reports the first, second, and third order autoregressive coefficient for the digital signals where a net inflow is represented as a +1, a net outflow as −1, and no flow is 0. Each column represents the results for the autoregression. The first column reports the results for the aggregated digital flow autocorrelation across all countries and funds. Columns 2 to 5 give the results for each of the four digital variables discussed in Table 1.2. The equation being estimated in all cases is

$$f_t^{\mathrm{d}} = c + a(L) f_{t-1}^{\mathrm{d}} + \varepsilon_t.$$

The subscripts on the autoregressive coefficients indicate the flow variable used in the autoregression; α_{oo} is the coefficient for the regression of own-fund, own-country digital signals, α_{co} is the coefficient for cross-fund own-country signals, α_{oc} is for own-fund cross-country signals, and α_{cc} is for cross-fund cross-country signals. Standard errors are reported in parentheses.

across funds and countries) implies that much of the autocorrelation of total flows is driven by the cross components, rather than own flows. Indeed, Table 2.1 shows that own-fund own-country flows account for only about 0.027, or about 7% of the overall total 0.40. Own-fund cross-country flows account for another 0.019, another 5% or so. So own-fund flows appear relatively unimportant in explaining total flow persistence. While there are more funds than countries, this is nevertheless a relatively small contribution from the cross-country effects.

The large contributions to total flows necessarily come from the two cross-fund components, and do so about equally from cross-fund own-country, and cross-fund cross-country components. These account for 0.219 and 0.140, respectively, of the total 0.40. The simple interpretation of this finding would be that the informed trader effects are not very important in explaining flow persistence; we should instead look to lags—particularly across funds and countries—in implementation and decision-making.

However, this conclusion would be naïve, since some portion of the cross effects might emerge under the informed trading hypothesis. The only way to find out is to examine the multivariate regression results from Eq. (6) where we estimate the size of the cross effects conditioning on own persistence.

When we estimate Eq. (6), we do so in weekly as well as daily data. These results are in Table 3, where there are several things to notice. First, it is clear that own-fund own-country persistence remains very powerful indeed. The first-order correlations generally are in the range of the high 20s to low 30s (and t-statistics in the hundreds or thousands). The first-order autocorrelation at the daily frequency is 26.3%. Second- and third-order partial autocorrelations show a similar pattern at 10.4 and 7.5%, respectively. Under our

Table 3
Persistence of institutional investor's flows

	$a_{oo}(L)$	$a_{co}(L)$	$a_{oc}(L)$	$a_{cc}(L)$
Daily data				
1st order	0.2630	0.0795	0.0224	−0.0063
	(0.0003)	(0.0022)	(0.0010)	(0.0083)
2nd order	0.1042	0.0394	−0.0204	−0.0257
	(0.0003)	(0.0023)	(0.0010)	(0.0088)
3rd order	0.0747	0.0369	−0.0044	0.0370
	(0.0003)	(0.0022)	(0.0010)	(0.0083)
R^2	0.1179			
SE	0.3170			
N	12,728,331			
Weekly data				
1st order	0.3040	0.0968	−0.0299	−0.0032
	(0.0006)	(0.0045)	(0.0023)	(0.0171)
2nd order	0.0575	0.0506	0.0217	−0.0242
	(0.0007)	(0.0049)	(0.0023)	(0.0181)
3rd order	0.0530	0.0536	0.0335	0.0248
	(0.0006)	(0.0045)	(0.0022)	(0.0170)
R^2	0.1213			
SE of regression	0.9785			
N	2,541,504			

Notes. This table shows the results of a regression of own-fund own-country digital signals from underlying flow data on its own lags, lagged cross-fund own-country signals, lagged own-fund cross-country signals and cross-fund cross-country signals. The equation estimated is

$$f^{d}_{i,k,t} = c + a_{oo}(L) f^{d}_{i,k,t-1} + a_{co}(L) \sum_{j \neq i} f^{d}_{j,k,t-1} + a_{oc}(L) \sum_{l \neq k} f^{d}_{i,l,t-1} + a_{cc}(L) \sum_{j \neq i} \sum_{l \neq k} f^{d}_{j,l,t-1} + \varepsilon_{i,k,t}.$$

These results are for the set of 21 different developed country equity markets in our sample. Results are reported for both the daily and weekly frequency. Standard errors are reported in parentheses.

null hypothesis, own-fund own-country persistence should be directly (negatively) linked to liquidity, and that appears to be the case in the data.

The same estimates for weekly data show a slightly higher first-order autocorrelation, at 30.4%, but lower second- and third-order correlations. The weekly estimates are probably somewhat more reliable. Given that the flows happen around the world, their daily timing is harder to pin down. All these numbers continue to be of very high statistical significance.

In terms of the cross effects in Table 3, there are a number of interesting points to notice. First, by far the most important cross effect is the cross-fund, own-country coefficient. In daily data the first-order coefficient is approximately 8%, with another 4% added by second- and third-order coefficients. In the weekly data, the cross-fund own-country coefficients come in slightly stronger, at 9 and 5%, respectively.

The other cross terms are far more mixed. The own-fund cross-country coefficients are not consistently positive. While the first-order daily coefficient is 2.2%, the second-order coefficient is −2.0%. In the weekly data, the first-order coefficient is negative, at

Table 4
Persistence with country specific coefficients

	$a_{oo}(L)$	$a_{co}(L)$	$a_{oc}(L)$	$a_{cc}(L)$	$b_o(L)$	$b_c(L)$
1st order	0.1350	0.0928	−0.0346	0.0061	0.1343	−0.1076
	(0.0005)	(0.0010)	(0.0005)	(0.0038)	(0.0046)	(0.0062)
2nd order	0.1550	0.0533	0.0244	−0.0134	0.0508	−0.1565
	(0.0010)	(0.0011)	(0.0005)	(0.0040)	(0.0046)	(0.0063)
3rd order	0.1501	0.0561	0.0345	0.0139	0.0376	0.1221
	(0.0010)	(0.0010)	(0.0005)	(0.0037)	(0.0046)	(0.0062)
R^2	0.1241					
SE	0.2137					
N	2,541,504					

Notes. This table shows the results of a regression of weekly own-fund own-country digital signals from underlying net flows of institutional investors on its own lags, cross-fund own-country signals, own-fund cross-country signals, cross-fund cross-country signals, own-country returns, and returns in the US market as the proxy for cross-country returns. This specification includes separate own-fund own-country autoregressive coefficients for each market, k. The equation estimated is

$$f^{d}_{i,k,t} = c_{i,k} + a_{oo;k}(L) f^{d}_{i,k,t-1} + a_{co}(L) \sum_{j \neq i} f^{d}_{j,k,t-1} + a_{oc}(L) \sum_{l \neq k} f^{d}_{i,l,t-1} + a_{cc}(L) \sum_{j \neq i} \sum_{l \neq k} f^{d}_{j,l,t-1} + b_o(L) r_{i,t-1} + b_c(L) r_{j,t-1} + \varepsilon_{i,k,t}.$$

These results in the upper panel are for the 21 different developed country equity markets and the lower panel for the 15 different emerging country equity markets in our sample. Standard errors are reported in parentheses.

−3.0%. These results suggest that there is little own-fund cross-country persistence after accounting for own-fund own-country persistence.

The absence of any effect also applies to cross-fund cross-country persistence. Here the coefficients in both daily and weekly data are small (at 1–4%) and are measured relatively imprecisely, many not meeting even standard levels of statistical significance. Some of the lag coefficients are also negative.

Table 4 provides estimates of Eq. (6). This specification is the same as that in Table 3, except that we allow own-fund own-country persistence to vary by country. This may help account for differences in liquidity across countries. Because there are now 21 own-fund own-country persistence coefficients (one for each country), we report in the table the average of these coefficients. Table 4 also includes lagged returns as additional explanatory variables. However, as previously stated, this results in only very minor impacts on lagged flow coefficients.

The estimates in Table 4 show that many countries have less first-order own-fund own-country persistence than reported for all countries combined. However, the effect appears to be more in the timing rather than in the magnitude of the own-fund own-country autocorrelation. While the first-order own coefficients fall, the second- and third-order coefficients rise by approximately offsetting amounts. The sum of the lagged coefficients in Table 4 (approximately 45%) remains essentially unchanged from the comparable weekly estimates in Table 3.

6. Conclusions

This paper has examined the persistence of institutional investor flows into a set of 21 developed countries. We confirm previous findings that the portfolio flows of international investors are highly persistent, with daily autocorrelations of about 40%. We then find that, by a simple additive decomposition, almost 90% of this is attributable to cross-fund components of persistence, with the cross-fund cross-country component being the single most important piece. Own-fund own-country persistence (which comes only from the trace of non-contemporaneous flow covariance matrix) is relatively unimportant, as might be expected from a large cross section of flow data.

How large should these various own and cross components be relative to one another? To answer this, we refer to what we call the 'informed trader hypothesis,' which says that traders with real or perceived information about a country or a stock, will get into their positions slowly, as long as liquidity is less than perfect. This behavior can explain own- as well as cross-effects and gives us a null hypothesis to determine their relative sizes.

When we implement our test of this, conditioning on the magnitude of own-fund own-country persistence, the nature of the persistence decomposition changes importantly. The own-fund own-country components are few in number, but explain much of what is happening in the cross section. Using this metric, approximately 75% of the persistence is attributable to own-fund own-country persistence, i.e., the informed trader model. Of the remaining cross effects, the only one of material importance is that of cross-fund own-country persistence, accounting for most of the remaining 25% of total persistence. This component cannot be explained by individual funds moving slowly into individual investments. Rather, it must be the result of some form of delay that operates across funds for a given country, not within funds across countries or across funds across countries.

Two plausible stories strike us as being consistent with these findings. The first is that there are meaningful implementation and decision-making lags across funds. Information may reach these funds at the same time, but those who act upon it do not do so simultaneously. There may be differences in the way investment decisions are made, or there may be other, fund-specific aspects of the investment process that result in information being expressed in trades at different time.

Another plausible story is that managers follow each other's trades after they have the chance to observe them. This may be very sensible behavior. It could be that fund trades carry positive information about future returns. We do not study that here, though previous studies (e.g., Froot et al., 2001) have found this to be the case. It could also be some funds follow others, not because they are chasing returns, but because other peer funds' allocations act as benchmarks (see Chow, 1995). Staying near the allocation of its peers would reduce a fund's peer risk. Either of these stories could be told as a herding story or as a rational decision-making story.

Finally, our results stand somewhat in contrast with studies that focus on the composition of stock index return persistence rather than flow persistence. In returns, non-contemporaneous index autocorrelations are driven exclusively by cross-stock effects. There are two points to make about this. First, it may be that intermediaries prevent evidence of predictable own-stock persistence from emerging, so that the informed trading

hypothesis cannot be tested well using return data. Second, it may be that there are also important cross-stock effects in flow data. In this paper, we disaggregate only to the country investment level, and we leave such questions about individual stock flow behavior for future research.

Acknowledgments

We are grateful to State Street Corporation for help in obtaining data and to Gang Bai for extraordinarily effective research assistance. The views expressed here are ours, and we alone bear responsibility for any mistakes or inaccuracies.

References

Choe, H., Kho, B.C., Stulz, R., 1999. Do foreign investors destabilize stock markets? The Korean experience in 1997. Journal of Financial Economics 54, 227–264.

Chow, G., 1995. Portfolio selection based on return, risk, and relative performance. Financial Analysts Journal, March–April, 54–60.

Cohen, R., Gompers, P., Vuolteenaho, T., 2001. Who underreacts to cashflow news? Evidence from trading between individuals and institutions. Manuscript. Harvard University.

Froot, K., O'Connell, P., Seasholes, M., 2001. The portfolio flows of international investors. Journal of Financial Economics 59, 151–193.

Froot, K., Perold, A., 1995. New trading practices and short-run market efficiency. Journal of Futures Markets 15, 731–766.

Froot, K., Ramadorai, T., 2004. Currency returns, intrinsic value, and institutional investor flows, with T. Ramadorai. Journal of Finance. In press.

Grinblatt, M., 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.

Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. American Economic Review 85, 1088–1105.

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

Kaminsky, G., Lyons, R., Schmulker, S., 2000. Managers, investors, and crises: mutual fund strategies in emerging markets. World Bank.

Kim, W., Wei, S.J., 2000. Foreign portfolio investors before and during a crisis. Journal of International Economics 56, 77–96.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1336.

Kyle, A., Xoing, W., 2001. Contagion as a wealth effect. Journal of Finance 56, 1401–1443.

Lakonishok, J., Shleifer, A., Vishny, R., 1992. The impact of institutional trading on stock prices. Journal of Financial Economics 32, 23–43.

Lo, A., MacKinlay, C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Nofsinger, J., Sias, R., 1999. Herding and feedback trading by institutional and individual investors. Journal of Finance 54, 2263–2295.

Richards, A., 2002. Big fish in small ponds: the momentum investing and price impact of foreign investors in Asian emerging equity markets. Reserve Bank of Australia.

Rouwenhorst, G., 1998. International momentum strategies. Journal of Finance 53, 267–284.

Seasholes, M., 2000. Smart foreign traders in emerging markets. Working paper. University of California Berkeley.

VanRoyen, A.S., Page, S., 2002. The multiple dimensions of asset allocation: countries, sectors, or factors? Working paper. State Street Associates.

Wermers, R., 1999. Mutual fund herding and the impact on stock prices. Journal of Finance 54, 581–622.

Wermers, R., 2002. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses. Journal of Finance 55, 1655–1695.