Peer Effects in Mutual Funds

JESSE BLOCHER*

September 19, 2013

ABSTRACT

This paper documents temporary abnormal returns in mutual fund performance due to peer effects among mutual funds associated by similar asset holdings. With a network specification of instrumental variables to control for correlated shocks to associated funds, I find that flows to and from peer mutual funds funds account for 1.6% of mutual fund quarterly excess return which reverses 1.1% in the following year. Temporary abnormal returns may explain mutual fund performance persistence in the absence of frictions inhibiting reallocation of investor funds across mutual funds.

*This paper is based on my dissertation and was previously circulated as "Contagious Capital: A Network Analysis of Interconnected Intermediaries". Special thanks to my committee (Jennifer Conrad, Adam Reed, Joey Engelberg, Greg Brown, Pab Jotikasthira and Peter Mucha) for timely and helpful advice on this project. I am also grateful to Nick Bollen, Rick Sias, Ethan Cohen-Cole, Tim Conley, Matt Bothner, Eric Ghysels, Chris Lundblad, Ed Van Wesep and Matt Ringgenberg for useful comments. Thanks also to seminar participants at the SEC, Michigan, Boston College, Georgia, Case Western Reserve University, Tulane, FRB-NY, and FRB-Boston for useful feedback, as well as conference attendees at the 2011 Washington University at St. Louis Corporate Finance Conference PhD poster session and the 2011 Financial Research Association meetings. Blocher is from the Owen Graduate School of Management, Vanderbilt University, 401 21st Avenue South, Nashville, TN 37203, Contact info: jesse.blocher@owen.vanderbilt.edu or www.owen.vanderbilt.edu/blocher.

Mutual funds remain the dominant form of investing by retail investors with assets under management recently eclipsing $13 trillion for the first time (ICI 2013). Mutual fund investments represent 23% of individual household assets and approximately 50% of retirement assets (ICI 2013). Thus, understanding the effects of capital flows on subsequent mutual fund performance and pricing remains a question of first order importance.

While the primary consensus of research on mutual fund managers indicates an inability to outperform after expenses (e.g. Carhart, 1997, Fama and French, 2010), there are studies which show persistence in performance, particularly in the short term (e.g. Bollen and Busse, 2004, Cremers and Petajisto, 2009, Amihud and Goyenko, 2013). Work that attempts to explain this persistence assumes, implicitly or explicitly, that the source of persistence is frictions which inhibit the allocation of funds by rational agents among competing mutual funds (Berk and Green, 2004, Franzoni and Schmalz, 2013). Regulators follow this theory, pushing for greater transparency among mutual funds in order to overcome these frictions.[1]

This paper provides empirical evidence that persistence in performance can be generated by temporary price effects from capital flows directed by rational, return-chasing mutual fund investors (Berk and Green, 2004). To illustrate, suppose investors move cash into a mutual fund which has recently performed well and that portfolio manager subsequently invests primarily in his existing portfolio (Lou, 2012), creating buying pressure that generates short-term abnormal returns. These price effects may spillover to mutual funds with similar holdings (or 'peers'). While it is unlikely that this effect could be caused by a single mutual fund, multiple funds with overlapping holdings could provide enough flow-based buying pressure

[1]For instance, in 2009, the SEC released rule 33-8998 requiring summary information to be prominently displayed and disclosures to be immediately available on websites to speed transmission of relevant information.

to affect prices. Aggregate flows have already been shown to affect stock prices (Lou, 2012, Jotikasthira, Lundblad, and Ramadorai, 2012, Coval and Stafford, 2007), and return chasing is a well-known phenomena (e.g. Edelen and Warner, 2001).

This paper makes three contributions. First, I show that localized mutual fund flows predict short-term mutual fund performance which subsequently reverses. Second, I show that this effect is due to peer effects among mutual funds with overlapping holdings rather than groupings of otherwise unrelated mutual funds based on some characteristic. To disentangle the 'peer' effect from the 'group' effect, I rely on the recent but growing econometrics of peer effects literature for identification which I further confirm with a network permutation test of randomized neighbors. My third contribution is methodological: this network methodology allows a level of analysis between the whole market and the individual mutual fund by focusing on peer groups – the market mesostructure. I analyze a distribution of economic significance across mutual funds rather than just an average effect on all funds, and show that this exponential distribution can have very large, localized effects even if the average effect is modest.

Peer flows, which are the combined effect of fund flows into and out of mutual funds with similar holdings are associated with a 1.6% quarterly abnormal return as defined by Daniel, Grinblatt, Titman, and Wermers (1997), hereafter called DGTW return. Approximately two-thirds of the effect reverses in the subsequent year. While these results use peers defined by raw portfolio weights, I obtain similar results with peer measures defined by benchmark adjusted weights and weights based on changes in portfolio holdings.

I show that this temporary abnormal return is due to spillover effects from peer funds by analyzing the coefficient estimates from a spatial autoregression (SAR) of peer flows on contemporaneous mutual fund flows.[2] In this specification, the amplification effect of peer funds is revealed with the coefficient for spillover effects always exceeding that of the direct effect. This pattern holds for all explanatory variables such as past flows and changes in market share (Spiegel and Zhang, 2013) and across the three different peer measures. The economic significance of spillover effects is 85% of the direct effect on average. This average effect, however, masks an underlying exponential distribution of outcomes in which direct effects at the 75th percentile are larger than the initial shock, and at the 90th percentile several multiples of the initial shock. This effect is only slightly reduced for spillover effects – managers who do not receive a shock feel effects that are almost as large as those who do.

The primary challenge for this study is identification of a causal peer effect distinct from a correlated shock to both the fund and its peers, which I address in three ways.

First, I apply a two-step GMM estimation (Kelejian and Prucha, 1998, Lee, 2003, Bramoullé, Djebbari, and Fortin, 2009). In the first step, I employ network-based instruments to identify the endogenous peer mutual fund flows in a specification to predict mutual fund flows. In the second step, I compute predicted peer flow as the predicted flow from the first step with the same peer similarity measure, and use it as the sole instrument for peer flows to infer effects on returns and spillover effects among fund flows.

Second, I run a network permutation test with randomized peers. Since randomized peers show no effect, I conclude that network structure factors in the result over latent market-wide forces.

Third, I control for style investing (e.g. Teo and Woo, 2004, Barberis and

[2]Spatial auto-regressions are typically used to measure spillover effects of traffic and pollution in geographic contexts.

Shleifer, 2003) by computing a specification using only peer funds between groups (i.e. omitting peers within the same style or category) and obtain the same outcome as the full sample. Isolating my analysis to peers only within the same Morningstar category further validates that my peer effect is distinct from a category effect: the sign of the effect reverses in line with the competitive result of Wahal and Wang (2011).

The notion of peers among mutual funds is not new. Cohen, Coval, and Pastor (2005) propose a performance measurement based on a peers, which we extend to investigate an influence process rather than benchmarking performance. Intuitively, this study can also be thought of as measuring the effects of Kiyotaki and Moore (1997) but with financial assets rather than durable assets and across security space rather than physical geography. The peer flow effect also explains how mutual fund performance can be explained by momentum even when mutual fund managers do not seem to be implementing a momentum strategy (Carhart, 1997). Performance persistence induced by slow price adjustment seems at odds with Wermers (1999) who finds that mutual fund herding speeds price adjustment. I reconcile this by suggesting that temporary abnormal returns generate performance persistence. I also show how money can be both smart and dumb (Zheng, 1999, Frazzini and Lamont, 2008): peer flows are positively correlated with fund performance in the short run and negatively correlated in the longer term.

This paper proceeds as follows. First I provide a brief review of related literature in Section I. Section II describes my measurement of peer effects in detail and gives some preliminary, motivating results. Section III describes my identification strategy including the two-step GMM/IV procedure and permutation test. Section IV discusses results, including the interpretation of network coefficients and their economic significance as well as distributional implications and robustness to style-based effects. Section V concludes.

I. Literature Review

Cohen, Coval, and Pastor (2005) propose measuring mutual fund performance based on peer performance. We extend this idea by noting that while peer associations can be categorizations, peers also influence each other. For example, a professor at a prestigious university may be well regarded due to her affiliation, but her performance is likely to depend on relationships with peers at that university – their skills, experience, and knowledge spill over onto her and vice versa in a mutually reinforcing cycle. This study can be thought of as measuring these types of spillover effects among mutual funds.

In Kiyotaki and Moore (1997), agents financing durable assets (e.g. land) with secured debt drives geographic spillover. While mutual funds do not use leverage, the fact that they provide daily, on-demand withdrawals at the NAV of the fund makes them comparable to overnight repo financed investment vehicles and thus susceptible to withdrawals forcing the sale of assets (Coval and Stafford, 2007). Similar to Kiyotaki and Moore, mutual fund holdings are both what drive mutual fund performance and are the collateral for investor funds. These analogies motivate my work as an empirical test of Kiyotaki and Moore (1997) among mutual funds connected in security space.

The dynamic described here also helps explain the evidence of mutual fund herd behavior (e.g. Wermers, 1999, Sias, 2004) without an information cascade (Bikhchandani, Hirshleifer, and Welch, 1992) since experimental evidence indicates that information cascades do not exist in the presence of a flexible market price (Drehmann, Oechssler, and Roider, 2005).[3] Instead, I show that mutual funds pur-

[3]Çelen and Kariv (2004) disentangle herding behavior from information cascades.

chasing their own portfolio with inflows leads to common trades by similar mutual funds which is measured herd-like behavior without requiring an information-based mechanism.

Dasgupta, Prat, and Verardo (2011) model a similar temporary abnormal return process where managers receive enough reputational payoff to purchase negative expected return securities, thus inducing them to overshoot fundamentals. My innovation is to show that returns can be positive in expectation for two reasons. First, informed agents cannot contemporaneously observe the source of price pressure (mutual fund flows) due to delays in disclosure and thus cannot predict the reversal. Second, fundamental value for a long-only equity investment is not clearly observable, so reinforcing revisions in relative valuation during a run up in prices leads investors crowding into a common position to push prices away from fundamentals while still maintaining a positive expected return (Stein, 2009).[4] That these returns are abnormal can only be seen ex post when prices reverse, as I show.[5]

The existence of temporary abnormal returns does not preclude manager skill (as evidenced in Coval and Moskowitz (2001), Kacperczyk, Sialm, and Zheng (2005), and Cohen, Frazzini, and Malloy (2008), among others). Indeed, its existence may explain why identifying skill is challenging, the results are mixed, and there are strong opinions on both sides of the argument - different groupings may find positive performance, negative performance, or no net result depending both on time horizon and categories chosen for grouping.

[4]Stein (2009) calls these 'un-anchored' strategies in which increased arbitrage capital flows do not by definition reduce mispricing as they would in a spread trade, for instance. Drehmann, Oechssler, and Roider (2005) provide experimental evidence that even in a simple, controlled environment, convergence to a fundamental value is slow and non-monotonic.

[5]The only countervailing force to this price appreciation is short sellers subject to limits to arbitrage (Shleifer and Vishny, 1997) and synchronization risk (Abreu and Brunnermeier, 2002, 2003)

II. Measuring Peer Effects

In this section, I introduce my peer measure and provide simple portfolio sorts showing that *Peer Flow* (defined as flows to and from peer mutual funds funds) affects mutual fund flows and returns independent of past mutual fund performance and past fund flows. First, I describe my data.

A. Data

My primary dataset is from Morningstar and contains the flows, returns, and full portfolio holdings of U.S. Open Ended funds from 1998 to 2009.[6] Flows of funds are a simple dollar value per fund, quarterly, and are net flows summarizing all subscriptions and redemptions across the relevant time period. This data includes reported values for both fund flows and portfolio returns in contrast to other studies which must compute fund flows from returns and changes in total net assets. Also included are equity style computed by Morningstar which places a fund on the three by three grid of Large, Mid, and Small cap and Value, Growth, and Blend as well as the Morningstar category which includes styles in addition to other categorization like industry segment, dividend funds, international funds, lifecycle funds, etc. Also included is information on fund inception date used to compute fund age and expense ratio for the fund as reported in the annual report.

To focus on equity funds as is common in the literature, I only keep funds with at least 75% of their holdings in equities. To alleviate incubation bias I only keep funds greater than $5 million in assets, hold at least 10 securities, and are at least one year old. I combine this data with CRSP by CUSIP when necessary to obtain stock characteristics. DGTW returns are computed as in Daniel, Grinblatt,

[6]Elton, Gruber, Blake, Krasny, and Ozelge (2010) perform a thorough comparison of the Morningstar holdings data with the more commonly used data from Thomson Reuters and find it to be very similar and without survivorship bias.

Titman, and Wermers (1997) and Wermers (2003), downloaded from Russ Wermers' website. *Flow* is defined as fund flow divided by total net assets as is typical (e.g. Coval and Stafford, 2007) and differenced to remove common group effects, discussed in detail in Appendix A. *Cash %* is defined as currency, treasuries, and other cash-like holdings, also divided by total net assets. Change in market share is total net assets divided by all mutual fund assets less the lagged value of the same ratio (Spiegel and Zhang, 2013), and Turnover is the minimum of positive and negative share changes divided by the average of this and the prior period's total net assets (Carhart, 1997). Stock variables (such as portfolio weights, assets, and cash) are measured at time t-1 and flow variables (such as fund flows and returns) are measured at time t, thus accumulating from t-1 to t. This is to avoid any mechanical correlations between things like cash holdings and fund flows since total assets or cash at time t may be highly related to fund flows accumulated from t-1 to t. Summary statistics are in Table I.

To compute fund benchmarks, I obtain Vanguard Exchange Traded Fund (ETF) holdings, also from Morningstar. These contain stock holding identifying information, shares, market value, and portfolio weights at a monthly frequency. ETFs are an excellent proxy for index weights because they are index-tracking products with data available at relatively high frequency. Vanguard is a particularly good company to choose as the family of ETFs since they have a long standing reputation for delivering low cost exposure to popular benchmark indexes.[7] In addition, ETFs provide a very broad set of possible index benchmarks including things like sector indices, dividend-based indices, and various small, mid, and large cap value, blend and growth indices which allows a higher probability

[7]Berk and Van Binsbergen (2012) in recent work use Vanguard Index Mutual Funds as benchmarks; the choice here of ETFs over index mutual funds is trivial in comparison because Vanguard's ETFs are technically a share class of their popular index mutual funds and so holdings information is identical.

of assigning the appropriate benchmark. The list of ETFs is in Table II.

B. Peer Measures

While there may be agreement that peers influence each other, the correct measure for peers is not obvious. In order to provide as general and robust a result as possible, I compute results for three different measures of peers showing the same outcome. While all three measures are based on holdings, they capture different aspects of a peer relationship and identify somewhat different groups of peers.

The first measure of similarity is based on raw portfolio weights, designated as *Raw* or *Raw Peer* throughout. This measure uses the simple market value weighted portfolio of security holdings, which sums to 1 by definition for each mutual fund portfolio.

The second similarity measure, called *Bench* or *Bench Peer*, takes those raw weights and subtracts off the fund's assigned benchmark. I assign the Vanguard ETF with maximal similarity as the benchmark for each fund, where benchmark similarity is computed the same way as peer similarity, discussed below. I then subtract the ETF holdings weights from the associated mutual fund holdings weights to obtain a benchmark deviation, which for each holding can theoretically vary from -1 to 1, but will not sum to 1 for each mutual fund portfolio.[8]

The third peer measure uses change in holdings, which I call *Delta* or *Delta Peer*. I start with the change in shares from period $t-1$ to t (one quarter in my sample) and multiply it by the average stock price for the quarter, which is the

[8]I computed this two ways, first using only those securities held by the mutual fund, and second using the entire set of stocks held by either the mutual fund or matched ETF. The two different calculations were correlated greater than 0.9 so I keep the former since it is less likely to produce erroneous similarity in the case of a poorly matched benchmark ETF. The downside is that it omits stocks held by the ETF but not the MF which may be an intentional benchmark deviation. These are likely very small errors.

same across all funds. This allows stocks to be price weighted without inducing similarity based on quarterly stock returns. I then compute a set of weights by dividing each price weighted share change by the sum of all price weighted share changes to get a vector that sums to 1.

For each of the three peer measures, I construct the similarity between two portfolios i and j, denoted s_{ij}, as the dot product between the security weight vectors of each portfolio manager i and j, divided by the product of the Euclidean norm of each vector.[9] Then, for each similarity measure, I construct *Peer Flow* for each manager i by computing a weight vector such that the similarity measure s_{ij} for each other manager j is divided by the sum over all of that fund's similarities, setting self-similarity s_{ii} to 0. For example, consider a portfolio manager with three neighbors having similarities of 0.1, 0.2, and 0.1, such that the weights are .25, .50, and .25, respectively. If those neighbor's flows (divided by total net assets) are 0.01, 0.05, and 0.10, respectively, then *Peer Flow* is $(.25 * .01) + (.50 * .05) + (.25 * .10) = 0.0525$.[10]

I compute other peer variables such as peer size (total net assets) in the same way.[11] The primary advantage of using this peer weighting procedure is that peer

[9] Note that this similarity measure is the same as the cosine of the angle between the two vectors in security space, and for centered data is identical to the computed correlation between the two vectors. See more discussion of cosine similarity vs correlation computations in Reca, Sias, and Turtle (2011). Detailed definitions of the norm and a full derivation is available in Appendix B. This notion of portfolio distance is intuitively and mathematically similar to that of social distance as in Conley and Topa (2002).

[10] While not identical, this formulation of similarity is correlated 0.98 with the similarity measure by Cohen, Coval, and Pastor (2005) which uses peer mutual funds to identify skilled mutual fund managers. They compute what I would identify in my framework as "Peer Alpha" which uses a similar weighting system with Jensen's alpha to identify closely related high performing funds as a measure of fund manager skill.

[11] All of the *Peer* variables are the same as structural equivalence variables in sociology originally defined by Burt (1987) and used more recently by Bothner (2003). The primary difference is that in sociology, these variables are typically lagged to provide identification, whereas herein I use an IV/GMM specification for better identification with my data. Lagged variables are poorly identified here because two out of three similarity measures (Raw and Bench) have temporal auto-regression coefficients of .80 and .76, respectively. For lagged peer variables to provide identification in this context, we would require close to 100% portfolio turnover each period which

statistics are very similarly distributed to the corresponding 'own' statistics, as seen in Table I.

These three peer measures (Raw, Bench, and Delta) combine to measure the dynamic I propose, which has a passive part and an active part. The passive part of the dynamic is when a fund's peers buying and/or selling activity affects common holdings. In this case, no action is required on the part of the fund manager to impact his fund return. The *Raw* and *Bench* peer measures both capture this passive effect, *Delta Peer* does not.

The active part occurs when a manager receives inflows or redemptions and must act. Funds acting in concert buying and selling securities is a key driver of price pressure, which is captured by *Delta* but less so by *Bench* or *Raw*. *Raw* peers capture some of this effect because previous research has indicated that fund managers sell in a 1:1 ratio with their holdings and buy in a 0.6:1 ratio (Lou, 2012).

Bench Peers capture the essence of active management: deviations from the benchmark are the choices made by a manager. *Raw* and *Delta* peer measures do not and thus may pick up similarities in underlying benchmarks.

For reference, *Raw* and *Bench* similarities are correlated 0.39, while *Raw* and *Delta* are correlated 0.11. *Bench* and *Delta* are correlated 0.16. I provide additional correlations of the different *Peer Flow* measures in Table III

C. Preliminary Intuitive Results

The primary identification challenge is disentangling peer influence from a correlated group effect – i.e. does fund A affect fund B or are both fund A and B impacted by an unobserved shock? I address this econometrically in Section III but here I perform portfolio sorts to show how peer flows, measured in different

does not happen in mutual funds (Ethan Cohen-Cole, personal communication). Instead, I used peer's lagged flows in my instrumentation strategy, discussed in detail later.

ways, are independent effects from other common predictors of mutual funds.

An important phenomena in this context is flow-based price pressure (Lou, 2012): prices move in response to institutional flows. Thus, it could be that groups of funds which simultaneously receive fund inflows subsequently outperform regardless of any common holdings. A second phenomena is return chasing (Edelen and Warner, 2001): fund flows follow performance. It could be that my measured peer effect is simply the result of groups of funds with good past performance all obtaining correlated flows, again regardless of any common holdings. I address these two phenomena by sorting lagged flows and lagged performance each with *Peer Flow* to show its independent effect on both fund flows and performance.

First, I address flow-based price pressure by analyzing mutual fund returns. Table IV shows results for *Raw*, *Bench*, and *Delta* peer variables in panels A, B, and C respectively, sorted by *Peer Flow* vertically and then either by lagged fund flow or lagged DGTW return horizontally. Reported is equal-weighted DGTW return for each quintile, in basis points.[12] At the bottom of each column (either lagged flow or lagged performance), we see the isolated, differential peer effect by subtracting quintile 1 from quintile 5. This difference is large, positive and statistically significant at the 1% level in each lagged flow or lagged performance quintiles across all peer specifications, showing a strong effect of *Peer Flow* holding lagged performance or flows constant. In contrast, the same is not always true within peer flow quintiles, when investigating the 1-5 difference column. In many cases, that difference is small and/or statistically insignificant and sometimes even negative and significant. This indicates that there is less predictive power in past flows or past performance once *Peer Flows* are taken into account and clearly indicates that *Peer Flow* is not simply a proxy for past flows or past performance

[12]Unreported analysis give the same result for Carhart alpha both as lagged performance measure and fund return measure instead of DGTW return.

when investigating mutual fund returns.

Second, I address the return-chasing phenomena by analyzing fund flows in the same framework. The results are in Table V, with *Raw*, *Bench*, and *Delta* peers in panels A, B, and C, respectively. As before, lagged fund flows and lagged performance (DGTW return) are sorted horizontally and *Peer Flow* vertically. This time, the variable reported is abnormal fund flow in basis points.[13] We see the same significant positive difference across *Peer Flow* within lagged fund flow and lagged performance quintile, indicating an important role for *Peer Flow* in predicting mutual fund flows. We also see strong evidence of return chasing independent of Peer Flow with the 5-1 column positive and significant within *Peer Flow* quintiles in all cases as well. This indicates that lagged flows and lagged returns play a role of equal importance as *Peer Flow* impacting fund flows.

The results so far indicate that the linkages among mutual funds provide additional information about mutual fund returns and flows, but are the common holdings the right link? To provide evidence of this, I briefly change the analysis to look at stocks rather than mutual funds. Table VI shows returns for groups of stocks held in common by mutual funds in different terciles of Peer Flows.[14] The top tercile is a group of top 50 stocks held by funds with large peer inflows and the bottom tercile is top 50 stocks held by funds with large peer outflows, with the middle being top 50 stocks of funds having close to net zero peer flow. We see that the top 50 stocks held in common by funds with large peer flows see a significantly positive return. Similarly, those top 50 stocks held by bottom tercile peer flow

[13]Fund flow is differenced globally as discussed in Appendix A, though the results are robust to this choice of differencing.

[14]Results presented for top 50 holdings, but I find the same outcome using the top 10 and top 25 holdings in unreported results, so this choice of threshold is not material. Each bin is defined as a group of stocks all of which are in the top 50 holdings of a fund by weight (Raw, Bench, Peer), and then held exclusively by funds in one of the three Peer Flow terciles. Thus, funds in the top 50 holdings of mutual funds in different terciles are omitted.

funds show large negative raw and abnormal returns, with statistically significant differences between the two. This indicates that they are a likely conduit for peer effects.[15]

Returning to our analysis of mutual funds, I present graphical evidence of reversals consistent with the temporary nature of the above measured abnormal returns. Figure 1 plots DGTW excess returns through time across two groups of mutual funds sorted by Peer Inflows (defined as top 40%) and Peer Outflows (defined as bottom 40%) with the middle 20% omitted. In the figure, we see the strong initial effect at time 0 already shown in Table IV. The figure then tracks these groups of funds through time, showing significant cumulative reversal over the following 3-4 quarters. This is true regardless of which peer measure chosen, though the size of the effect is smallest with the *Delta* peer specification, which is perhaps unsurprising given the more transient nature of that measure.

While these results provide some measure of confidence that *Peer Flow* is an effect independent of past flows or performance which subsequently reverse, they are not conclusive because they do not control for other factors (size of the fund, fund fees, etc.) nor do they provide econometric rigor to show causality. To do that, we need a more thorough identification approach.

III. Identification and Methodology

The primary challenge of this study is identification of peer effects. This is because *Peer Flow* is an endogenous effect – if the average flow of my peers affects me, then my fund flow affects them. This endogenous effect is hard to distinguish from exogenous effects where a fund and its peers have common exogenous charac-

[15] This result is consistent with Antón and Polk (2010) who find that overlapping mutual fund holdings increase comovement of the stocks held in common.

teristics, for instance fund age. It is also hard to distinguish from correlated effects, for instance fund size, style or category, since funds in these groups may experience inflows or outflows in correlated patterns. Starting with Manski (1993) who identified the problem, there has been a long literature focusing on econometric specifications and network conditions necessary to identify endogenous effects.

The first step toward identification is spatial differencing to remove group means, which is necessary but not sufficient (see Appendix A). To achieve full identification of causal peer effects, I draw on a two-step GMM/IV specification (Kelejian and Prucha, 1998, Lee, 2003). Among multiple instrument specifications, I also employ two-step peer (my neighbor's neighbor) variables as instruments as in Bramoullé, Djebbari, and Fortin (2009) within the above two-step GMM/IV specification. Finally, I discuss a network permutation test designed to establish a baseline for the analysis and remove the possibility of market-wide effects which may be driving peer effects as an omitted variable.

A. Two-Step GMM Methodology

Kelejian and Prucha (1998) document a novel two-step GMM/IV procedure for identifying peer effects. In the first step, you specify instruments predicting the endogenous variable as the dependent variable, which is fund flow in this study. Then, with that specification you generate a set of predicted values for the endogenous variable, which you then use as the instrument in the second step GMM/IV specification.[16] The first step GMM specification is as follows:

[16] To minimize confusion, I will refer to each of these specification as Step 1 and Step 2, and within each we have the typical GMM Stage 1 and Stage 2 regressions. So listed are Step 1 Stage 1 and 2, then Step 2 Stage 1 and 2.

$$PeerFlow = Instruments \tag{1}$$

$$Flow = \widehat{PeerFlow} + OtherVariables \tag{2}$$

I then compute $Peer\left(\widehat{Flow}\right)$ with the predicted value $\widehat{Flow}$ as the instrument in the second step, exactly identified GMM specification:

$$PeerFlow = Peer\left(\widehat{Flow}\right) \tag{3}$$

$$Flow = \widehat{PeerFlow} + OtherVariables \tag{4}$$

A proper GMM/IV specification uses instruments which are correlated with endogenous regressors but orthogonal to the error term. For each GMM specification, I include four key statistics which validate the instrument specification. The *KP LM Stat* is the Kleinbergen-Paap Wald Statistic which tests for weak instruments. Weak instruments are not correlated 'enough' with the endogenous regressors and can lead to biased results. This test should reject the null hypothesis so strong instruments should show a large test statistic and low p-value (less than 0.05). The *Hansen J statistic*, also known as the test of overidentifying restrictions, tests the validity of the overidentification of the model, which is correctly specified when we cannot reject the null hypothesis. This is a test of whether the instruments are correlated with the error term of the second stage regression, which should not be so. Thus, a very small J statistics and correspondingly large p-value which fail to reject the null indicate a properly overidentified GMM specification.[17] Note that

[17] For more information on the derivations of these test statistics, see Hayashi (2000) or some other comparable statistics text. Also consult the help files with Stata's xtivreg2 procedure, which I used for all of my GMM estimations (Schaffer and Stillman, 2007).

the Hansen J test is not available in the Step 2 GMM specifications because they are exactly identified with a single instrument, which is as intended.

B. Two-Step Peers and other Instrument Specifications

I have estimated a total of nine instrument specifications which all give similar results. Each are properly specified in that they give valid results from tests of weak instruments and overidentification.[18] They are as follows:

1. Peer Lagged Flow. (contemporaneous peers, lagged flows)
2. Lagged Peer Flow. (lagged peer relationships and lagged flows)
3. Two Step Peer Flow. (contemporaneous fund flow of two-step peers)
4. Two Step Peer Flow and Peer Lagged Flow. (1 and 3 combined)
5. Two Step Peer Flow, Two Step Size. (average size of two step peers)
6. Two Step Peer Flow, Two Step Size, Peer Lagged Flow. (1 and 5 combined)
7. Peer Lagged Flow, Two Step Peer Lagged Flow, Two Step Size.
8. Peer Lagged Flow, Two Step Peer Lagged Flow.
9. Peer Lagged Flow, Two Step Size.

In specifications 1 through 6, the square of each term is also included for better overidentification. In the interest of brevity, I only report results for Specification 1 and 7, but results for the others are available upon request. Specification 1 has the advantage of being the most parsimonious, intuitive specification, as well as using a lagged variable which is a very common instrument in the literature. Specification 7 uses two-step peers as instruments without any squared terms, and is closest to the specification in Bramoullé, Djebbari, and Fortin (2009) which is the most rigorous way to identify causal peer effects econometrically.[19]

[18]This is true in most cases. I ran all nine specifications across three peer measures and three differencing methods with both flows and returns as dependent variables. I also checked the between-within specifications discussed later. This is over 50 regressions for each instrument specification, and most gave valid results, with the weakest being the 'within' set of specifications.

[19]In unreported results, I attempt to directly follow Bramoullé, Djebbari, and Fortin (2009) who include all two step peer variables as instruments but could never pass the Hansen J test of overidentifying restrictions (i.e. instruments were correlated with second stage residuals indicat-

Bramoullé, Djebbari, and Fortin (2009) prove that using two-step peers as instruments solves the peer effects identification problem, based on the fact that there exist "intransitive triads" in most networks which allows separation of A's influence on B from B's influence on A. An intransitive triad is present if A connects to B and B to C, but A is not connected to C. Thus, A can instrument for B's influence on C since any influence A has on C must be through the common relation with B. In network terminology, A and C are *Two-Step* neighbors. For instance, a U.S. technology fund may be connected to a mid-cap fund through common mid-cap technology holdings, and that mid-cap fund may also be connected to a Latin American fund through mid-cap Latin American holdings. Thus, the lagged flows of the Latin American fund can instrument for the mid-cap fund's influence on the U.S. technology fund since they are only connected through their common mid-cap neighbor.

Not all two-step neighbors form intransitive triads, however. Two-step neighbors can only serve as an instrument if they satisfy the exclusion restriction – that the instrument is only correlated with the dependent variable through the endogenous regressor. To address these concerns, Bramoullé, Djebbari, and Fortin (2009) specify a rank test which establishes that peer influence effects are identified through two-step neighbors, which my network satisfies in all time periods.[20]

ing that they did not satisfy the exclusion restriction). Bramoullé, Djebbari, and Fortin (2009) do not show results for this test on their data. As such, I pare down the list of instruments to maximize explanatory power while attempting to use variables which are as exogenous as possible. This had the benefit of balancing overidentification with weak instruments.The provided list was my first and only try so this is not the result of a specification search (Leamer 1978).

[20]Proposition 5. This is simply a rank test of the matrix of network connections. I take the identity matrix, my Peer Weight matrix, and my Two-Step Peer Weight matrix (see the appendix for details), reorder them to be in column vectors, and show that the resulting matrix has at least rank 3. I repeat this for all three peer measures. More details of this computation available on request.

C. Permutation Test

I provide a network permutation test as a baseline. Because we are analyzing equity mutual funds who are all subject to market-wide movements (i.e. they have a 'beta' in some sense) it may be that there are unobserved effects which drive peer relations and cannot be otherwise controlled. Another way of thinking about this is ensuring that my null hypothesis is zero.

To that end, I run a network permutation test with each set of peer measures. In each case, I randomly reassign peers in the network, but hold all else constant. I then create *Peer Flow* from the new randomized peers and run a simple OLS regression since peer-based instruments no longer work with random peer assignment. This test is sufficient to show an insignificant baseline result because with the real data network, the relationship in an OLS regression is strongly positive and statistically significant.

Table VII summarizes the results of 1,000 permutations, and shows no significant effect. There are fewer than 1% statistically significant results at the 99th percentile level and less than 5% statistically significant results at the 95th percentile level. The average coefficient found is less than 0.01 and the average T statistic is less than 0.1 (both in absolute value). This non-result sets a clean slate with a clear null hypothesis of zero effect of peer flows on mutual funds.

IV. Results

I now discuss my identified, rigorous results consistent with peer fund flow effects inducing temporary abnormal mutual fund returns. First, we look at abnormal fund returns and subsequent reversals, then investigate the spillover effect of peer flows in a regression with fund flows as the dependent variable. The nature of this regression is such that, similar to an autoregression with lagged dependent variables, the coefficients require interpretation to fully capture the economic significance. Finally, I discuss robustness checks which focus on peers only between styles or within styles to show that my peer effect is independent of documented style-based comovement. First, a brief discussion of the econometrics of my base specification.

A. Econometrics and Controls

Recent papers by Petersen (2009) and Gormley and Matsa (2013) have demonstrated the importance of paying careful attention to econometric specifications and not simply "following the literature" when computing standard errors and including fixed effects, respectively. In light of these, I find it necessary to include both time and fund fixed effects and further cluster my standard errors in both time and fund dimensions, though most of the literature has used Fama-MacBeth or an OLS specification with fewer standard error corrections without fixed effects.[21]

To capture group effects and following Sirri and Tufano (1998), I include a *Style Flow Diff* control variable which is the average flow into that fund's style, differenced with the average global fund flow (ex that style's flow) similar to the

[21]A Breusch-Pagan test and an F test on RSS of regressions with and without time and fund fixed effects show that it is necessary to include some type of fixed or random effects. A Hausman test verifies that fixed effects are necessary over random effects (Kennedy, 2003). Clustering standard errors in both time and manager dimensions produces large changes in standard errors indicating that this is a necessary step (Petersen, 2009). With the same pooled OLS and Fama-MacBeth framework as Coval and Stafford (2007), I get results qualitatively similar to them and others who have investigated this relation such as Lou (2012) and Ferreira, Keswani, Miguel, and Ramos (2011). Results from these tests as well as a table comparing the varying differences in specification are available upon request. Recall that my dataset is different from the other studies cited and as such these test results may or may not extend to their specifications so I am making no claims about their results. The inclusion of group fixed effects is strongly encouraged by Gormley and Matsa (2013) over simple group means as a method of controlling for group effects. I do this in unreported results with no material change in result.

fund-level differencing (Appendix A). In unreported results, I include a control using Morningstar Category-based groups with similar results. Since I am using fund flows as a percentage of fund total net assets, I follow Spiegel and Zhang (2013) and include change in market share (*Mkt Shr Chg*), which accounts for total fund flows. Also included are typical controls for both flow and return regression such as *Fund Size*, computed as the logarithm of total net assets, *Fund Age* which is the log of the age of the fund in years, expense ratio in basis points, and turnover as defined in Carhart (1997). I also include *Volatility* which is the trailing annual rolling volatility of raw mutual fund returns as well as controls for cash holdings (Simutin, 2009) and a portfolio-based liquidity measure aggregated across stocks held as computed in Amihud (2002).[22]

B. Temporary Abnormal Returns

Table VIII shows the results for abnormal returns which subsequently reverse. Models 1 and 2 are for *Raw* peers, 3 and 4 for *Bench* peers, and 5 and 6 for *Delta* peers. All show the same basic result, which is a positive and significant relationship between excess DGTW return and *Peer Flow* contemporaneously in the first model, and a negative and significant relationship between the subsequent annual DGTW return and the same *Peer Flow* variable in the second model. Excess DGTW return is defined as the mutual fund gross return less mutual fund DGTW return, which is the portfolio weighted individual equity DGTW returns. In the contemporaneous return regressions *Peer Flow* is instrumented by $Peer(\widehat{Flow})$ with the predicted values from from the step 1 GMM regression. In the reversal regressions (2, 4, and 6), simple OLS suffices since there is little endogeneity

[22]In unreported results, I substitute Amihud liquidity for other liquidity measures such as spreads (bid minus ask over midpoint) and stock turnover (average daily volume divided by shares outstanding) with similar results, available on request.

between a quarterly lagged peer flow variable and subsequent annual return.

Economically, the effect in Models 1, 3, and 5 is 1.6%, 1.4%, and 0.5% in quarterly excess return. The economic significance of the reversal in models 2, 4, and 6 is -1.1%, -0.75% and -0.37% respectively, which corresponds with reversals of 69%, 54% and 75%. In unreported results with four factor alpha (Carhart, 1997) as the performance measure, I get similar results.

In Table IX we see the same result, this time with the two-step instrument specification. Note that Models 2, 4, and 6 are identical to those in the previous table and only reproduced for easy comparison.

C. Flows and Spillover

Flow results are in Table X for both the Step 1 and Step 2 GMM/IV regressions. Models 1 and 2 are the Step 1 and Step 2 for *Raw* peers, 3 and 4 the same for *Bench* peers and 5 and 6 for *Delta* peers.[23] The result of interest is *Peer Flow* in the first row – in all cases it is positive and significant. We obtain the same result with the two-step peer instrumentation specification in Table XI, which give further evidence that the causal effect is fully identified.

Interpreting this result is not trivial, however, because this is a spatial auto-regression (SAR) specification such that *Flow* appears on both sides of the equals sign, requiring transformation similar to an temporal auto-regression with a lagged dependent variable.[24] The primary effect of the SAR specification is that the coefficient on every explanatory variable is not a scalar but an $N \times N$ matrix (given N mutual funds). Thus, there are separate estimated coefficients for the

[23]Coval and Stafford (2007) employ both lagged flows and lagged returns as predictors, but I instead follow (Lou, 2012) who uses lagged Carhart four-factor alpha. Alpha is computed using a 12 month rolling average regression, accounting for alpha over the past 12 months, so only a single lag is necessary. Lagged excess DGTW provide similar results but require four lags since it is quarterly.

[24]I discuss some details around estimating an SAR model with GMM/IV in Appendix C

effect of each mutual fund on each other mutual fund. It is these matrix coefficients which allow the direct measurement of spillover effects as well as the analysis of the distribution of economic effects across the population of mutual funds.

To compute these matrix coefficients, I begin by rewriting the general specification in Equation 4 in matrix form, without the instrumentation:

$$F = \rho W F + X\beta + \epsilon \tag{5}$$

in which F is the $N \times 1$ vector of fund flows at time t. W is a row-stochastic transformation of the $N \times N$ portfolio similarity matrix S at time t, such that $PeerFlow = W \cdot F$ and ρ is the estimated coefficient on $PeerFlow$. X represents all other explanatory variables for simplicity. The result is

$$(I_N - \rho W)\, F = X\beta + \epsilon \tag{6}$$

$$F = X_1\tilde{\beta}_1 + X_2\tilde{\beta}_2 + \ldots + X_L\tilde{\beta}_L + \epsilon \tag{7}$$

for each $l = 1 \ldots L$ explanatory variables. Each actual estimated coefficient is

$$\tilde{\beta}_{l,N\times N} = (I_N - \rho W)^{-1}\, \beta_l \tag{8}$$

which is an $N \times N$ matrix.[25]

To interpret the network coefficients, I divide each matrix coefficient into direct effects, represented by the diagonal, and spillover effects which reside on the off-diagonal. The results are in Table XII, with Panels (a), (b), and (c) showing results for *Raw Peer*, *Bench Peer*, and *Delta Peer*, respectively. The first column is the scalar coefficient estimate, β, without the network transformation. Next are

[25]Since there is a W for each time t, I compute $\tilde{\beta}_{l,N\times N}$ at each time t then report the average results for each manager across time.

the direct effects, computed as the average of the diagonal of the matrix coefficient, $\tilde{\beta}$. These are the effects that each mutual fund has on itself (the traditional analysis) including feedback effects from spillover to others which is propagated back to the originator. Next, spillover effects are computed as the average of all off-diagonal entries in the matrix coefficient, $\tilde{\beta}$, capturing the average effect that a shock to a mutual fund has on other mutual funds. The final column reproduces the standard error of the coefficient estimate for reference as an indicator of statistical significance.[26] While the direct effects only slightly exceed the simple coefficient estimate, the spillover effects are very large. For the Lagged Flow variables, spillover effects are double the direct effects, and for the rest of the variables like lagged alpha and fund size, they are an order of magnitude or more greater than direct effects. This pattern holds across the three peer measures.

Next we turn to economic significance in Table XIII. In this table, I have simulated a one standard deviation shock in each explanatory variable on one third of the population of managers. In the first column, I show the average effect on those who do *not* experience this shock, which is a measure of the economic significance of spillover. The second column is the average direct effect on those managers who directly receive the shock. The third column is the ratio of the spillover to direct effect, and the final two columns divide the effect by the standard deviation of Flow, the dependent variable in the regression, to standardize the size of the effect and more easily allow for comparison. As above, Panels (a), (b), and (c) show results for *Raw*, *Bench* and *Delta* peers, respectively.

The results are striking – the economic significance of spillover effects are, on average, 85% of the direct effect across the board. We see this uniformity within a single panel because each coefficient is divided by the same matrix as seen in the

[26] This entire exposition follows LeSage and Pace (2009) and is standard for an SAR model.

above exposition. There is no mechanical reason why we should see such similar results across panels, however, so the fact that spillover effects are remarkably similar regardless of peer specification provides strong evidence that peer effects are quite important regardless of precisely how they are measured.

Most explanatory variables give economic significance between 20%-50% of a standard deviation, which is a moderate to large effect. The largest effect is that of fund size, which we discount since a one standard deviation shock to fund size is unlikely in a single quarter. Liquidity, expense ratio, and fund age have small effects and turnover has a large (negative) effect, almost a full standard deviation.

The average effects in Table XIII only tell part of the story, however, because the distribution of effects is exponential, not normal. Table XIV presents a range of percentiles across the distribution of outcomes, both on the directly shocked managers and those who experience the resulting spillover. Displayed are percentiles of fund flow divided by its standard deviation, but only for lagged flow, lagged performance, changes in market share, and trailing annual volatility since these are important variables which may be subject to unexpected shocks.[27] Panels (a), (b), and (c) show results for *Raw Peer*, *Bench Peer*, and *Delta Peer*, respectively.

The distribution of direct effects and spillover effects are essentially the same regardless of peer measures. Greater differences exist instead across the four variables summarized. Lagged flow effects at or above the 75$^{\text{th}}$ percentile give an effect greater than a standard deviation which increases to 2 or 3 times a standard deviation at the 95$^{\text{th}}$ percentile. The effect of changes in market share are almost as big as flow, followed by lagged performance as captured by alpha. Each of these shows a greater than one standard deviation effect at the 90$^{\text{th}}$ percentile with multiple standard deviations at the 99$^{\text{th}}$ percentile. The effect of trailing annual volatility

[27] Results for the remaining variables available on request.

is less, but the upper tail of the distribution still exceeds a full standard deviation.

D. Robustness to Style Investing effects

It could be that my peer measures are identifying funds in the same style which may receive correlated flows to generate my result (e.g. Boyer, 2011, Teo and Woo, 2004, Barberis and Shleifer, 2003). Because the existing *Style Flow Diff* control may not be fully convincing, I conduct the following test. I partition my sample in to *within* and *between* groups, where *within* only includes peers who share the same style and *between* only includes peers which span across styles. You can visualize this partition as a block diagonal $N \times N$ mutual fund similarity matrix (with N fund managers) sorted by style where the within specification includes only the 9 style blocks on the diagonal and the between specification only includes the remaining blocks off the main diagonal. All other computations remain the same.

The results in Tables XV and XVI show positive and significant results for *Peer Flow* in all specifications for the two different instrument specifications. Each table has nine models in groups of three by peer measure. The first is the full sample, reproduced from earlier results for reference, the second is the *between* partition, the third is *within.* Models 1-3 are for *Raw* peers, 4-6 for *Bench* peers, and 7-9 for *Delta* peers in both tables. In each case, we see one of two patterns: either the *between* regression coefficient on *Peer Flow* is very close to the reference coefficient including all peers or the *within* coefficient is significantly less than the reference coefficient including all peers. Either of these patterns is evidence that peers with common styles are not driving the result either because I can remove them with little effect or because by themselves they have a much smaller effect than in the larger sample.

Further, Table XVII shows evidence consistent with Wahal and Wang (2011) who investigate competition among mutual funds. Competitive flows should have a negative coefficient on *Peer Flow* because flows into a competitive peer are associated with outflows for me. To investigate this apparent paradox, I isolate my analysis to a situation which is most likely to reflect a competitive environment, similar to Wahal and Wang (2011). I focus on Morningstar category, since it A. is an easy categorization device for investors and B. includes the same categories of funds as Wahal and Wang (2011). Instead of the global differencing used thus far, I use category-based differencing such that each fund's flow is measured as the difference between its flow and the average flow into or out of the category to which it belongs. This way I capture shifts in flows among funds within the same category even with minimal net flows.

The *within* result in Model 3 of Table XVII captures the effect – a negative and significant result for the flow coefficient. The entire sample and *between* subset remain positive and significant, even with the category-based differencing. I obtain insignificant results predicting fund returns, which follows because rebalancing flows should cancel out any flow-based price effects on the common holdings since one fund is buying and the other selling.

This results shows the consistency of my overall result: in general, I find feedback effects associated with positive-reinforcing spillover effects, but when I isolate my analysis to a scenario where competitive effects are likely to be strongest, I find that the sign of the result reverses, showing that competitive effect, consistent with the existing literature.

V. Conclusion

In the wake of the recent financial crisis, the peer effects among market participants has become an important new area of research. Employing a network-based specification, I show that interconnected mutual funds induce fund flow spillover effects inducing an abnormal return of 1.6% per quarter, substantially reversed in the subsequent year. I identify the influence of portfolio manager's fund flows on each other by exploiting the network structure as an instrument. I show that spillover effects are almost as large as the direct effects on average but can be much larger since the distribution of effects is exponential.

An interesting implication of this work is the idea of 'optimal frictions' in financial markets. While the literature has focused on removing frictions to speed adjustments, it may be that too little friction leads to abnormal returns which subsequently reverse to the detriment of price efficiency.

Network methods are becoming more popular in corporate finance (e.g. Cohen, Frazzini, and Malloy, 2008, Ahern and Harford, 2010, Lewellen, 2012) and market microstructure (e.g. Cohen-Cole, Kirilenko, and Patacchini, 2010), although little has been done to apply network methods to securities markets. My network approach allows an analysis of peer influence processes, bringing structure to cross-sectional analysis previously only available in the time-series. While I have applied it to portfolio interconnections, it may also have broad applicability to other areas such as interbank lending (Cohen-Cole, Patacchini, and Zenou, 2011) and stock market volatility (Greenwood and Thesmar, 2011). In a time when bailouts are motivated not because of too-big-to-fail but because of too-interconnected-to-fail, understanding and quantifying the interconnections among market participants is a vital pursuit.

References

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

———, 2003, Bubbles and Crashes, *Econometrica* 71, 173–204.

Ahern, Kenneth R, and Jarrad Harford, 2010, The Importance of Industry Links in Merger Waves, *Unpublished working paper*.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Y, and R Goyenko, 2013, Mutual Fund's R2 as Predictor of Performance, *Review of Financial Studies* 26, 667–694.

Antón, Miguel, and Christopher Polk, 2010, Connected Stocks, *Unpublished working paper* pp. 1–56.

Barberis, N, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Berk, Jonathan, and Richard Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan, and Jules Van Binsbergen, 2012, Measuring Managerial Skill in the Mutual Fund Industry, *SSRN eJournal*.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, *Journal of Political Economy* 100, 992–1026.

Bollen, Nicolas P B, and Jeffrey A Busse, 2004, Short-Term Persistence in Mutual Fund Performance, *Review of Financial Studies* 18, 569–597.

Bothner, Matthew S, 2003, Competition and Social Influence: The Diffusion of the Sixth-Generation Processor in the Global Computer Industry, *The American Journal of Sociology* 108, 1175–1210.

Boyer, Brian, 2011, Style Related Comovement: Fundamentals or Labels?, *The Journal of Finance* 66, 307–332.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, 2009, Identification of peer effects through social networks, *Journal of Econometrics* 150, 41–55.

Burt, Ronald S, 1987, Social Contagion and Innovation: Cohesion Versus Structural Equivalence, *The American Journal of Sociology* 92, 1287–1335.

Carhart, Mark M, 1997, On Persistence in Mutual Fund Performance, *The Journal of Finance* 52, 57–82.

Çelen, Boğaçhan, and Shachar Kariv, 2004, Distinguishing Informational Cascades from Herd Behavior in the Laboratory, *The American Economic Review* 94, 484–498.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The small world of investing: Board connections and mutual fund returns, *Journal of Political Economy* 116, 951–979.

Cohen, Randolph B, Joshua D Coval, and Lubos Pastor, 2005, Judging Fund Managers by the Company They Keep, *The Journal of Finance* 60, 1057–1096.

Cohen-Cole, Ethan, Andrei Kirilenko, and Eleonora Patacchini, 2010, Are Networks Priced? Network Topology and Systemic Risk in High Liquidity Markets, *Unpublished working paper*.

Cohen-Cole, Ethan, Eleonora Patacchini, and Yves Zenou, 2011, Systemic Risk and Network Formation in the Interbank Market, *Carefin Working Paper*.

Conley, Timothy G, 1999, GMM estimation with cross sectional dependence, *Journal of Econometrics* 92, 1–45.

———, and Giorgio Topa, 2002, Socio-economic distance and spatial patterns in unemployment, *Journal of Applied Econometrics* 17, 303–327.

Coval, Joshua D, and Tobias J Moskowitz, 2001, The Geography of Investment: Informed Trading and Asset Prices, *Journal of Political Economy* 109, 811–841.

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

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How Active Is Your Fund Manager? A New Measure That Predicts Performance, *Review of Financial Studies* 22, 3329–3365.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *The Journal of Finance* 52, 1035–1058.

Dasgupta, Amil, Andrea Prat, and Michela Verardo, 2011, The price impact of institutional herding, *Review of Financial Studies* 24, 892–925.

Drehmann, Mathias, Jörg Oechssler, and Andreas Roider, 2005, Herding and Contrarian Behavior in Financial Markets: An Internet Experiment, *The American Economic Review* 95, 1403–1426.

Edelen, Roger M, and Jerold B Warner, 2001, Aggregate price effects of institutional trading: a study of mutual fund flow and market returns, *Journal of Financial Economics* 59, 195–220.

Elhorst, J Paul, 2010, Spatial Panel Data Models, *Handbook of Applied Spatial Analysis* pp. 377–407.

Elton, Edwin J, Martin J Gruber, Christopher R Blake, Yoel Krasny, and Sadi O Ozelge, 2010, The effect of holdings data frequency on conclusions about mutual fund behavior, *Journal of Banking and Finance* 34, 912–922.

Fama, Eugene F, and Kenneth R French, 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, *The Journal of Finance* 65, 1915–1947.

Ferreira, Miguel A, Aneel Keswani, Antonio F Miguel, and Sofia Brito Ramos, 2011, The Flow-Performance Relationship Around the World, *Unpublished working paper*.

Franzoni, Francesco A, and Martin C Schmalz, 2013, Fund Flows in Rational Markets, *SSRN Electronic Journal.*

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

Gormley, Todd A., and David A. Matsa, 2013, Common Errors: How to (and Not to) Control for Unobserved Heterogeneity, *Review of Financial Studies.*

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Hamilton, James, 1994, *Time Series Analysis* (Princeton University Press: Princeton, NJ).

Hayashi, Fumio, 2000, *Econometrics* (Princeton University Press).

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, 2007, Whom You Know Matters: Venture Capital Networks and Investment Performance, *The Journal of Finance* 62, 251–301.

Jotikasthira, Chotibhak, Christian T Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of financial shocks, *The Journal of Finance.*

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the Industry Concentration of Actively Managed Equity Mutual Funds, *The Journal of Finance* 60, 1983–2011.

Kelejian, Harry H., and Ingmar R. Prucha, 1998, A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, *The Journal of Real Estate Finance and Economics* 17, 99–121.

———, 2002, 2SLS and OLS in a spatial autoregressive model with equal spatial weights, *Regional Science and Urban Economics* 32, 691–707.

Kennedy, Peter, 2003, *Guide to Econometrics* (The MIT Press) 5 edn.

Kiyotaki, Nobuhiro, and John Moore, 1997, Credit Cycles, *Journal of Political Economy* 105, 211–248.

Lee, Lung fei, 2003, Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances, *Econometric Reviews* 22, 307–335.

LeSage, James, and R Kelley Pace, 2009, *Introduction to Spatial Econometrics* (Chapman and Hall/CRC: Boca Raton, FL).

Lewellen, Stefan, 2012, Executive Compensation and Peer Effects, *Unpublished working paper* pp. 1–62.

Lou, Dong, 2012, A Flow-Based Explanation for Return Predictability, *Review of Financial Studies* 25, 3457–3489.

Manski, Charles F, 1993, Identification of Endogenous Social Effects: The Reflection Problem, *The Review of Economic Studies* 60, 531–542.

Petersen, Mitchell A, 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

Reca, Blerina, Richard W Sias, and Harry J Turtle, 2011, Hedge Fund Herding: The Apologists' Evidence, *SSRN Electronic Journal*.

Schaffer, Mark E, and Steven Stillman, 2007, Enhanced routines for instrumental variables/GMM estimation and testing, *Stata Journal*.

Shleifer, Andrei, and Robert W Vishny, 1997, The Limits of Arbitrage, *The Journal of Finance* 52, 35–55.

Sias, Richard W, 2004, Institutional Herding, *Review of Financial Studies* 17, 165–206.

Simutin, Mikhail, 2009, Excess Cash and Mutual Fund Performance, *SSRN Electronic Journal*.

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

Spiegel, Matthew I, and Hong Zhang, 2013, Mutual fund risk and market share-adjusted fund flows, *Journal of Financial Economics*.

Stein, Jeremy C, 2009, Presidential address: Sophisticated investors and market efficiency, *The Journal of Finance* 64, 1517–1548.

Teo, Melvyn, and Sung-Jun Woo, 2004, Style effects in the cross-section of stock returns, *Journal of Financial Economics* 74, 367–398.

Wahal, Sunil, and Albert (Yan) Wang, 2011, Competition among mutual funds, *Journal of Financial Economics* 99, 40–59.

Wermers, Russ, 1999, Mutual Fund Herding and the Impact on Stock Prices, *The Journal of Finance* 54, 581–622.

———, 2003, Is Money Really "Smart"? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence, *Unpublished working paper*.

Zheng, Lu, 1999, Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability, *The Journal of Finance* 54, 901–933.

VI. Appendix

A. Spatial Differencing

An introductory time series analysis class will introduce early on the necessity of using first differences in any time series analysis with lagged dependent variables as regressors (Hamilton, 1994). This matters because most financial and economic time series are not stationary – i.e. they have a time trend – and this trend will induce a false correlation across time if it is not differenced out. We have the same issue with 'spatial lags', another term for the peer measures employed herein. A spatial lag is intuitively similar to a temporal lag: a one step spatial lag is a measured value of a peer or neighbor who is directly connected to the initial agent of interest. Similarly, there could be a two step neighbor who is my neighbor's neighbor, all in the same time, t. So, to measure the effect that each fund manager's flows have on each other, I need to difference out any common flow and only consider the abnormal flow to or from each manager.This removes the possibility of an induced spatial lag due to common groupings or other market-wide events or trends.

However, spatial lags are more complicated than temporal lags because space is two dimensional and so the level to be differenced out is not as obvious. Thus, to remove these effects, we need to identify spatial groups which denote the 'common' part of the flow to be differenced out. I do this in three ways: globally (G), by style (S) and by Morningstar Category (C). I follow the practice common in the literature and compute the differenced flow as the observed own flow minus the group flow less the own flow:

$$K\_Flow\_Diff_i = Flow_i - \sum_{\substack{\forall j \neq i \\ \forall j \in k}} Flow_j, i \in k \tag{9}$$

Where k is the group definition, either all funds for global differencing (G), or all funds in a given style (S) or Morningstar Category (C). *Flow* is already transformed prior to this differencing by dividing through by total net assets in order to normalize it across funds and minimize the disruption caused by large funds. [28] Note that I only report results for global differencing because it is the most commonly used in the peer effects literature and all give similar results. The others are available on request. The exception is in the final robustness check of competitive mutual fund effects where I report Category differences.

B. Details on how to compute Similarity Measure

As noted in the text, similarity between two portfolios i and j, denoted s_{ij}, as the dot product between the security holding weight vectors of each portfolio manager i and j, divided by the product of the Euclidean norm of each vector. s_{ij} is computed at each

[28] I perform the differencing after normalizing flow because the differencing creates strange results with dollar levels due to the tremendous difference in fund size which conflates the differenced measure with a size effect since all small funds have large negative spatial differences in fund flows anytime they are grouped with other large funds, who have corresponding very small differences.

time t, but I suppress the time subscript for expositional ease. Specifically, where h_i is a vector of portfolio weights for manager i, the similarity between managers i and j is defined as

$$s_{ij} = \frac{h_i \cdot h_j}{|h_i|\,|h_j|} \tag{10}$$

For each manager i, the Euclidean norm is defined across M securities as

$$|h_i| = \sqrt{\sum_{m=1}^{M} h_{im}^2} \tag{11}$$

Deriving this same measure in matrix form, let H be the $M \times N$ holdings matrix, with portfolio managers as each column, and each row consisting of the weight between 0 and 1 each manager places on that security. My portfolio similarity matrix is then

$$S = \frac{H^T H}{|H| \cdot |H|} \tag{12}$$

in which each s_{ij} already defined above is an element of symmetric similarity matrix $S_{N\times N}$ and h_i are the columns of H. The norm of the matrix H is a Euclidean column norm, such that for each column j, the norm of H_j is defined as

$$|H_j| = \sqrt{\sum_{m=1}^{M} h_{jm}^2} \tag{13}$$

I then compute *Peer Flow* as the dot product of the weight vector and the corresponding vector of fund flows for each manager. Formally, peer weights are computed as

$$PeerWeight_{ij} = \frac{s_{ij}}{\sum_k s_{ik}}, k \neq i \tag{14}$$

and *Peer Flow* is thus

$$PeerFlow_i = \sum_k PeerWeight_{ik} Flow_k \tag{15}$$

In matrix form, if W is a row-stochastic transformation of S, such that each row sums to 1, then $PeerFlow = W \cdot Flow$ in which both $PeerFlow$ and $Flow$ are $N \times 1$ vectors and W is an $N \times N$ matrix at time t.[29]

Two-step neighbors are computed as $B = S^2$, with matrix multiplication (as opposed to element-by-element) where the diagonal of S has already been set to 0 to avoid duplicating one-step and two-step neighbors.[30] In summation notation, the equivalent

[29] As noted in the paper, my measure is very similar to Cohen, Coval, and Pastor (2005). The difference is that they use a doubly-stochastic matrix which is both row and column normalized such that each row and each column sums to one. Mine is only row stochastic.

[30] A nonzero diagonal indicates a 'self-loop.' If S has a nonzero diagonal, a 'two-step' neighbor could be i connecting to i (a self loop) and then i connecting to j, which therefore duplicates a

product is

$$b_{ij} = \sum_{q=1}^{N} s_{iq} s_{qj},\; q \neq i, j \tag{16}$$

with the diagonal of B also set to zero such that a manager cannot be his own two-step neighbor.[31] If $\widetilde{W}$ is the row-stochastic, $N \times N$, two-step weighting matrix derived from B, then $TwoStepPeerFlow = \widetilde{W} \cdot Flow$ or as a summation:

$$\widetilde{w}_{ji} = \frac{b_{ji}}{\sum_k b_{jk}} \tag{17}$$

$$TwoStepPeerFlow_j = \sum_k \widetilde{w}_{jk} Flow_k \tag{18}$$

C. Estimating an SAR Model

I produce estimates using Generalized Method of Moments, whereas most specifications of this type in the spatial econometrics literature estimate models via Maximum Likelihood. Conley (1999) notes that maximum likelihood specifications in which spatial dependence is measured with error are misspecified. While measurement error is unlikely to be a problem with geographical measures of distance typical of the spatial econometrics literature, my measure of distance in security space may be much less precise. Fortunately, Kelejian and Prucha (2002) show that with panel data, both OLS and GMM estimators are consistent, and thus represent the appropriate estimation approach. Elhorst (2010) includes a short discussion on MLE vs IV/GMM estimators, noting that while the use of IV/GMM is promising, it is still new to the spatial econometrics literature and needs further research.[32]

The primary issue in using a GMM specification to estimate a SAR model is that the auto-regression coefficient should be bounded between -1 and 1. In an MLE framework, this can be explicitly incorporated as a constraint on the optimization process. In GMM it is technically possible but computationally and mathematically difficult with no good solution. Elhorst (2010) suggests simply trying different specifications within the limits of what theoretically should be included in order to obtain an outcome which provides a coefficient within those bounds. I have attempted to do that here and presented the best results I have. In cases where the coefficient of the Step 2 GMM specification exceed 1 (Model 4 in Tables X and XI, I simply set it to 0.99 for interpretation in the subsequent section. This is only a problem if there is a different local optima for the estimation with a coefficient on the interval $(-1, 1)$ which is radically different that 0.99, for instance less than 0. I tried multiple specifications with varying controls (and transforms of those controls) and rarely got coefficient estimates much different from those reported, so this seems unlikely.

one-step neighbor. This is a common adjustment in network analysis.

[31]The diagonal of B must now be set to 0 because for every one-step neighbor, a manager is his own two-step neighbor. For instance, i connects to j, but then j also connects back to i, such that for every connection like this i is his own two-step neighbor.

[32]Spatial Econometrics primarily uses MLE because they have only one network – geography – and thus no panel data. In these cases, only MLE is appropriate (Elhorst, 2010).

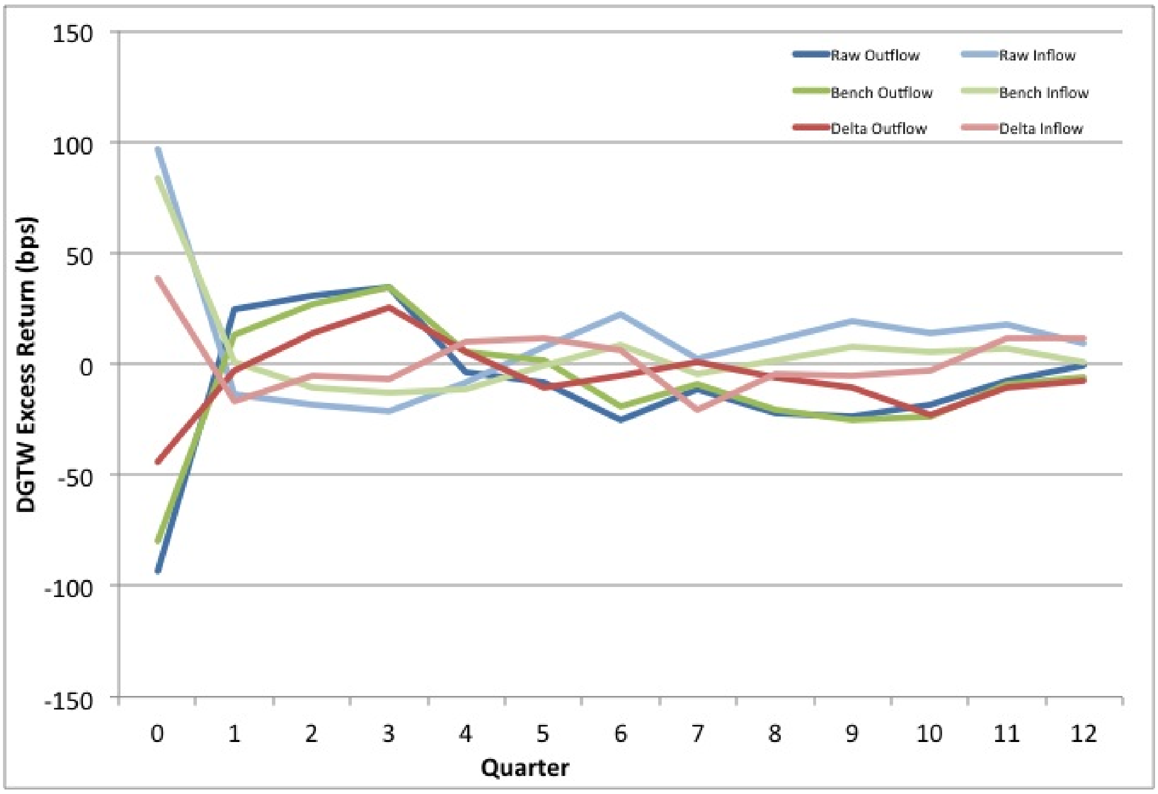

Figure 1: Performance of Peer Flow portfolios through time. Portfolios based on Peer Flow are formed at time 0. Outflow is defined as the lower two quintiles, Inflow the upper two quintiles. Portfolios held for the subsequent 12 quarters. Return is Excess DGTW return in basis points. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Results for *Raw Peers* are in blue, *Bench Peers* in green, and *Delta Peers* in red.

Table I: Mutual Fund Summary Statistics

Total Net Assets are fund assets in $millions, as are fund flows. Cash % is cash like instruments in dollars over total net assets. Carhart Alpha is a four-factor alpha computed annually, rolling. Turnover is the minimum of positive and negative turnover, divided by the average of total net assets between this and the prior period as defined in Carhart (1997). Change in Mkt Share equals the difference between this period and last period's total net assets over market-wide total net assets, as defined in Spiegel and Zhang (2013). Volatility is rolling annual trailing fund return volatility. Number of holdings is a count of holdings, expense ratio is as reported in the fund's annual report. Flow % is dollar fund flow divided by total net assets. G Diff is globally differenced flow. Style Flow Diff is the average mutual fund style flow less the global average flow ex that fund's flow, where style is defined on a three by three matrix based on market cap and growth-blend-value. Category Flow Diff is the same except defined by Morningstar category rather than fund style, which is a more granular definition using styles as well as sectors, countries, regions, etc. Raw Pr is a peer measure based on raw portfolio weights, Bench Pr is a peer measure based on portfolio weights differenced from a benchmark, and Delta Pr is a peer measure based on dollar-weighted changes in holdings.

	Count	Mean	Std Dev	Min	Median	Max
Total Net Assets ($M)	51,376	1,435	5,694	5.003	293.8	19,3453
Fund Flow ($M)	51,376	-0.074	206.4	-4417	-1.89	8,170
Cash %	51,376	0.033	0.034	-0.003	0.023	0.276
Fund Return	51,376	0.014	0.113	-0.756	0.021	0.832
Carhart Alpha	51,376	-0.001	0.009	-0.082	-0.001	0.106
Fund Age (yr)	51,376	8.232	6.391	1.000	6.800	77.000
Turnover	51,376	0.027	0.024	0.000	0.021	0.443
Change in Mkt Share	51,376	-0.028	1.044	-47.06	-0.007	26.93
Volatility (Annual)	51,376	0.050	0.028	0.003	0.043	0.275
Number of Holdings	51,376	126.5	140.4	10	86	3,622
Expense Ratio (bps)	51,376	138	49	-51	138	749
Flow %	51,376	-0.014	0.128	-1.000	-0.014	0.735
G Diff Flow %	51,376	-0.017	0.128	-1.029	-0.017	0.776
Style Flow %	51,376	0.002	0.023	-0.175	0.002	0.213
Style Flow Diff	51,376	-0.000	0.018	-0.176	-0.002	0.236
Category Flow %	51,376	0.001	0.033	-0.501	-0.000	0.356
Category Flow Diff	51,376	-0.001	0.030	-0.516	-0.002	0.362
Raw Pr G Diff	51,376	-0.002	0.011	-0.099	-0.003	0.101
Bench Pr G Diff	51,240	-0.005	0.012	-0.107	-0.005	0.298
Delta Pr G Diff	51,374	0.000	0.023	-0.306	-0.001	0.281

Table II: Table of Vanguard ETFs

This is a list of Vanguard ETFs used as benchmarks to determine each Mutual Fund's benchmark peer weights by subtracting off the matched ETF weights from the raw mutual fund portfolio weights.

ETF Name	Ticker
Vanguard Consumer Discretionary ETF	VCR
Vanguard Consumer Staples ETF	VDC
Vanguard Energy ETF	VDE
Vanguard Financials ETF	VFH
Vanguard Health Care ETF	VHT
Vanguard Industrials ETF	VIS
Vanguard Information Technology ETF	VGT
Vanguard Materials ETF	VAW
Vanguard Telecom Services ETF	VOX
Vanguard Utilities ETF	VPU
Vanguard S&P 500 ETF	VOO
Vanguard Growth ETF	VUG
Vanguard Value ETF	VTV
Vanguard Mid-Cap ETF	VO
Vanguard Mid-Cap Growth ETF	VOT
Vanguard Mid-Cap Value ETF	VOE
Vanguard Small Cap ETF	VB
Vanguard Small Cap Growth ETF	VBK
Vanguard Small Cap Value ETF	VBR
Vanguard Dividend Appreciation ETF	VIG
Vanguard High Dividend Yield Indx ETF	VYM
Vanguard Large Cap ETF	VV
Vanguard Extended Market Index ETF	VXF
Vanguard Total Stock Market ETF	VTI
Vanguard REIT Index ETF	VNQ
Vanguard Total Intl Stock Idx ETF	VXUS
Vanguard Total World Stock Index ETF	VT
Vanguard MSCI EAFE ETF	VEA
Vanguard MSCI Europe ETF	VGK
Vanguard MSCI Pacific ETF	VPL
Vanguard FTSE All-Wld ex-US SmCp Idx ETF	VSS
Vanguard FTSE All-World ex-US ETF	VEU
Vanguard FTSE Emerging Markets ETF	VWO

Table III: Flow Correlations

Flow % is dollar fund flow divided by total net assets. G Diff is globally differenced flow as discussed in Appendix A. Style Flow Diff is the average mutual fund style flow less the global average flow ex that fund's flow. Category Flow Diff is the same except defined by Morningstar category rather than fund style. Raw Pr is a peer measure based on raw portfolio weights, Bench Pr is a peer measure based on portfolio weights differenced from a benchmark, and Delta Pr is a peer measure based on dollar-weighted changes in holdings.

	Total Avg Flow %	Style Flow Diff	Cat Flow	G Diff Flow % Flow %	Raw Pr G Diff Flow %	Bench Pr G Diff Flow %	Delta Pr G Diff
Total Avg Flow %	1.000						
Style Flow Diff	-0.003	1.000					
Cat Flow Diff	-0.028	0.408	1.000				
G Diff Flow %	-0.023	0.117	0.183	1.000			
Raw Pr G Diff Flow %	-0.238	0.452	0.526	0.158	1.000		
Bench Pr G Diff Flow %	-0.156	0.440	0.522	0.175	0.873	1.000	
Delta Pr G Diff Flow %	0.025	0.246	0.295	0.259	0.457	0.438	1.000

Table IV: DGTW Return Portfolios

Excess DGTW Return in basis points sorted in quintiles by Peer Flow and Lag Flow (differenced) or Lag DGTW Return, by Peer Measure (Top is raw peers, middle is benchmark peers, bottom is delta peers. Significance is denoted at the 1, 5, and 10% level.

		Lag Flow						Lag DGTW Return					
		Outflow				Inflow		Negative				Positive	
	Pr Flow	1	2	3	4	5	Diff 5-1	1	2	3	4	5	Diff 5-1
	(Out) 1	-136	-119	-109	-169	-160	-24	-106	-135	-125	-131	-176	-69***
	2	-68	-81	-72	-82	-105	-37***	-106	-74	-53	-67	-74	32*
A. Raw	3	4	-11	-7	-17	-24	-29**	-25	-41	-19	5	44	69***
	4	85	46	39	31	32	-54***	43	22	6	40	104	61***
	(In) 5	215	199	176	183	205	-10	110	120	73	94	299	189***
	Average	-15	-15	-3	-3	7	21***	-47	-43	-28	-2	92	139***
	Diff 5-1	351***	318***	285***	352***	365***		216***	256***	198***	225***	475***	
	(Out) 1	-140	-117	-110	-157	-134	6	-113	-131	-104	-124	-156	-43*
	2	-41	-65	-54	-72	-94	-53***	-57	-62	-49	-50	-71	-15
B. Bench	3	2	-10	-21	-21	-23	-25*	-14	-41	-34	-5	36	50***
	4	63	29	28	19	24	-40**	18	5	-4	32	94	76***
	(In) 5	221	189	177	166	177	-43*	81	96	68	106	311	231***
	Average	-15	-14	-3	-4	6	22***	-43	-42	-28	-1	91	134***
	Diff 5-1	361***	305***	287***	323***	312***		193***	227***	172***	230***	467***	
	(Out) 1	-66	-40	-59	-129	-126	-60***	-84	-83	-70	-59	-2	83***
	2	-47	-46	-47	-79	-104	-57***	-73	-82	-64	-31	-2	71***
C. Delta	3	0	-35	-35	-29	-36	-36**	-62	-41	-35	-17	33	95***
	4	-14	18	10	19	9	23	-8	-19	-15	17	73	81***
	(In) 5	211	164	137	117	119	-92***	66	32	62	89	274	208***
	Average	-15	-12	-6	-5	6	21**	-46	-45	-27	0	90	136***
	Diff 5-1	277***	204***	195***	246***	245***		150***	115***	132***	149***	276***	

Table V: Fund Flow Portfolios

Abnormal Fund Flow (normalized by net assets, differenced) in basis points sorted in quintiles by Peer Flow and Lag Flow (differenced) or Lag DGTW Return, by Peer Measure (Top is raw peers, middle is benchmark peers, bottom is delta peers. Significance is denoted at the 1, 5, and 10% level.

		Lag Flow						Lag DGTW Return					
		Negative				Positive		Negative				Positive	
	Pr Flow	1	2	3	4	5	Diff 5-1	1	2	3	4	5	Diff 5-1
	(Out) 1	-9.30	-4.99	-2.92	-0.41	3.02	12.32***	-5.12	-3.32	-3.18	-2.49	-2.55	2.57***
	2	-7.78	-4.29	-2.13	0.24	4.70	12.48***	-4.57	-2.78	-1.73	-1.45	-0.29	4.28***
A. Raw	3	-6.76	-3.70	-1.92	1.13	6.50	13.26***	-3.12	-1.63	-1.30	-0.35	1.29	4.40***
	4	-6.40	-3.12	-1.14	1.44	7.38	13.79***	-2.36	-0.78	0.11	0.43	1.79	4.15***
	(In) 5	-5.09	-2.36	-0.50	2.19	9.60	14.70***	-0.30	0.36	0.99	2.08	4.24	4.55***
	Average	-7.44	-3.81	-1.72	1.05	6.74	14.18***	-3.73	-1.97	-1.12	-0.23	1.61	5.34***
	Diff 5-1	4.20***	2.62***	2.42***	2.60***	6.58***		4.82***	3.67***	4.17***	4.57***	6.79***	
	(Out) 1	-9.48	-5.25	-3.06	-0.88	1.64	11.12***	-5.71	-3.60	-3.55	-3.68	-2.83	2.88***
	2	-7.76	-4.25	-2.08	0.39	5.09	12.84***	-4.07	-2.78	-2.25	-1.19	0.04	4.11***
B. Bench	3	-6.99	-3.54	-1.95	1.07	6.11	13.10***	-3.28	-1.96	-0.82	-0.61	1.14	4.41***
	4	-5.94	-3.25	-1.16	1.48	8.00	13.94***	-1.90	-0.60	-0.01	0.79	2.05	3.95***
	(In) 5	-4.45	-2.20	-0.43	2.29	9.94	14.38***	0.37	0.78	1.22	2.53	4.46	4.09***
	Average	-7.41	-3.82	-1.73	1.04	6.78	14.19***	-3.72	-1.98	-1.13	-0.24	1.69	5.41***
	Diff 5-1	5.04***	3.05***	2.63***	3.17***	8.30***		6.08***	4.38***	4.77***	6.21***	7.29***	
	(Out) 1	-9.15	-4.90	-3.18	-2.31	-3.78	5.37***	-6.63	-5.00	-4.60	-4.55	-3.99	2.63***
	2	-9.07	-4.64	-3.11	-1.22	0.48	9.55***	-5.95	-4.58	-4.11	-3.39	-2.68	3.27***
C. Delta	3	-7.26	-4.07	-2.02	0.70	5.39	12.65***	-3.63	-1.92	-1.53	-1.11	0.64	4.28***
	4	-4.88	-2.26	-0.35	2.13	8.94	13.82***	0.14	0.74	1.68	2.59	3.42	3.28***
	(In) 5	-1.11	-0.43	0.62	3.40	10.70	11.80***	1.70	3.18	4.14	5.01	6.22	4.52***
	Average	-7.31	-3.76	-1.70	1.01	6.55	13.86***	-3.69	-1.96	-1.02	-0.20	1.39	5.09***
	Diff 5-1	8.04***	4.47***	3.80***	5.72***	14.47***		8.33***	8.18***	8.74***	9.57***	10.21***	

Table VI: Top 50 stock holdings by peer flow tercile

Each group is a set of stocks who are in the top 50 holdings of a fund by weight, where weight could be *Raw*, *Bench* or *Peer* as designated. Once the top 50 are identified, they are sorted in to high, medium and low Peer Flow bins only when they are exclusively held by funds in that tercile. This means funds in the top 50 but held by funds in differing terciles are omitted. Excess is excess DGTW Return, Raw is raw return, both in percent. Tercile 1 is outflows, Tercile 3 is inflows, Tercile 2 is close to zero net flow. Differences in bold are significant at 1% level.

Raw Peer Weights			Bench Peer Weights			Delta Peer Weights		
PrFlw	Excess	Raw	PrFlw	Excess	Raw	PrFlw	Excess	Raw
(Out) 1	-1.0%	-1.3%	(Out) 1	-1.3%	-1.2%	(Out) 1	0.2%	1.8%
2	0.6%	2.9%	2	0.4%	2.3%	2	0.2%	2.3%
(Inf) 3	1.9%	8.1%	(In) 3	2.2%	8.6%	(In) 3	0.5%	6.2%
Diff 3-1	**2.9%**	**9.5%**	Diff 3-1	**3.5%**	**9.8%**	Diff 3-1	0.3%	**4.4%**

Table VII: Simulation results

Data is the same as in Table VIII and X, except peer flow variable generated by randomizing peer assignments, run 1,000 times. Avg Peer Coeff is the average of the coefficient on Peer Flow, defined by the different peer measures noted at the top of each column. Avg Peer T Stat is the average T statistic of that coefficient. Results at the 99th and 95th percentile are simple counts of how many of the 1,000 runs showed a coefficient significantly different from zero with a one-tailed t test at the given level of significance. For instance, at the 99th percentile, if less than 1% of results are significant than the result is no different than random noise.

	Raw Weight Peer		**Benchmark Peer**		**Delta Peer**	
	Flow	**DGTW**	**Flow**	**DGTW**	**Flow**	**DGTW**
Avg Peer Coeff	−0.006	0.004	−0.001	−0.006	−0.004	−0.004
Avg Peer T Stat	−0.026	0.001	−0.003	−0.022	−0.054	−0.043
99th Pctl: Right Tail						
% Sig	0.4%	0.6%	1.1%	0.5%	0.7%	0.7%
99th Pctl: Left Tail						
% Sig	0.8%	1.3%	0.4%	0.8%	0.8%	0.7%
95th Pctl: Right Tail						
% Sig	3.8%	4.0%	2.8%	3.3%	2.9%	3.3%
95th Pctl: Left Tail						
% Sig	3.8%	3.9%	4.0%	3.6%	3.5%	3.6%

Table VIII: Return Overshoot Result with Reversal

Models 1 and 2 are Raw Peers, Models 3 and 4 Benchmark, Models 5 and 6 Delta. Odd numbered models have quarterly Excess DGTW return estimated by GMM, even numbered models have the holding period return for the subsequent four quarters estimated by OLS. Peer flow is defined with the given peer measure. Instrument is predicted peer flow from the Step 1 GMM/IV regression with the main peer lagged flow instrument. Flow % is dollar fund flow divided by total net assets, globally differenced. Flow and Return lags 2-4 are included but omitted, as is Cash %. A portfolio-weighted Amihud measure is also included but omitted. L Fund Size is the log of Total Net Assets in $millions. Turnover is as defined in Carhart (1997). Mkt Share Change is as defined in Spiegel and Zhang (2013). Volatility is rolling annual trailing fund return volatility. Number of holdings is a count of holdings, expense ratio is as reported in the fund's annual report. Time and Fund Fixed Effects included, standard errors clustered by both time and fund. KP LM stat tests the null of weak instruments. T statistics are in parentheses and significance is denoted at the 1, 5, and 10% level.

	Raw Peer		**Bench Peer**		**Delta Peer**	
	(1) DGTW	(2) Lead	(3) DGTW	(4) Lead	(5) DGTW	(6) Lead
Peer Flow	1.6707***	-1.0030***	1.2307***	-0.6108***	0.1728***	-0.1514***
	(4.16)	(-2.65)	(5.47)	(-2.60)	(4.61)	(-2.82)
Lag1 Flow	-0.0036	0.0044	-0.0046	0.0049	-0.0039	0.0046
	(-0.82)	(0.73)	(-1.02)	(0.83)	(-0.90)	(0.77)
Lag1 Return	0.0167	-0.1546	0.0258	-0.1895	0.0569	-0.1767
	(0.23)	(-1.26)	(0.34)	(-1.56)	(0.74)	(-1.39)
L Fund Size	-0.0112***	-0.0428***	-0.0105***	-0.0436***	-0.0105***	-0.0431***
	(-7.16)	(-15.94)	(-6.81)	(-14.87)	(-6.67)	(-15.74)
Log Fund Age (yrs)	0.0020	0.0111*	0.0020	0.0124*	0.0016	0.0112*
	(0.97)	(1.74)	(1.01)	(1.81)	(0.74)	(1.77)
Mkt Shr Chg (bps)	0.0073***	-0.0017*	0.0072***	-0.0019*	0.0076***	-0.0018*
	(5.79)	(-1.72)	(5.49)	(-1.81)	(5.68)	(-1.81)
Group Flow	-0.0440	0.0562	0.0306	-0.0172	0.2996***	-0.1398
	(-0.40)	(0.32)	(0.40)	(-0.12)	(4.90)	(-1.06)
Volatility	-0.0793	0.0380	-0.0759	0.0647	-0.0697	0.0221
	(-0.27)	(0.10)	(-0.25)	(0.16)	(-0.23)	(0.05)
Exp Ratio	-0.0088*	-0.0268	-0.0098*	-0.0160	-0.0099**	-0.0264
	(-1.84)	(-1.31)	(-1.83)	(-0.80)	(-2.23)	(-1.28)
Turnover	0.0003	0.0012**	0.0003	0.0015**	0.0004	0.0009
	(0.93)	(2.04)	(0.86)	(2.54)	(1.32)	(1.63)
Observations	46723	36028	43343	33094	46722	36027
R Squared	0.15	0.08	0.15	0.09	0.11	0.08
Fund clusters	3,432	2,897	3,289	2,699	3,432	2,897
Time clusters	44	40	44	40	44	40
KP LM Stat	29.85		29.45		27.69	
KP LM p value	0.0000		0.0000		0.0000	

Table IX: Return Overshoot Result with Reversal - robust instrument specification. All is the same as Table VIII except the instrument specification used includes two peer lagged flow and two peer size, following Bramoullé, Djebbari, and Fortin (2009).

	Raw Peer		Bench Peer		Delta Peer	
	(1) DGTW	(2) Lead	(3) DGTW	(4) Lead	(5) DGTW	(6) Lead
Peer Flow	1.6755***	-1.0030***	1.2318***	-0.6108***	0.1705***	-0.1514***
	(4.16)	(-2.65)	(5.47)	(-2.60)	(4.61)	(-2.82)
Lag1 Flow	-0.0036	0.0044	-0.0046	0.0049	-0.0038	0.0046
	(-0.82)	(0.73)	(-1.02)	(0.83)	(-0.89)	(0.77)
Lag1 Return	0.0166	-0.1546	0.0258	-0.1895	0.0570	-0.1767
	(0.23)	(-1.26)	(0.33)	(-1.56)	(0.74)	(-1.39)
L Fund Size	-0.0112***	-0.0428***	-0.0105***	-0.0436***	-0.0105***	-0.0431***
	(-7.16)	(-15.94)	(-6.81)	(-14.87)	(-6.67)	(-15.74)
Log Fund Age (yrs)	0.0020	0.0111*	0.0020	0.0124*	0.0016	0.0112*
	(0.97)	(1.74)	(1.01)	(1.81)	(0.74)	(1.77)
Mkt Shr Chg (bps)	0.0073***	-0.0017*	0.0072***	-0.0019*	0.0076***	-0.0018*
	(5.79)	(-1.72)	(5.49)	(-1.81)	(5.68)	(-1.81)
Group Flow	-0.0451	0.0562	0.0304	-0.0172	0.3002***	-0.1398
	(-0.41)	(0.32)	(0.39)	(-0.12)	(4.91)	(-1.06)
Volatility	-0.0793	0.0380	-0.0759	0.0647	-0.0697	0.0221
	(-0.27)	(0.10)	(-0.25)	(0.16)	(-0.23)	(0.05)
Exp Ratio	-0.0088*	-0.0268	-0.0098*	-0.0160	-0.0099**	-0.0264
	(-1.84)	(-1.31)	(-1.83)	(-0.80)	(-2.23)	(-1.28)
Turnover	0.0003	0.0012**	0.0003	0.0015**	0.0004	0.0009
	(0.93)	(2.04)	(0.86)	(2.54)	(1.31)	(1.63)
Observations	46723	36028	43343	33094	46722	36027
R Squared	0.15	0.08	0.15	0.09	0.11	0.08
Fund clusters	3,432	2,897	3,289	2,699	3,432	2,897
Time clusters	44	40	44	40	44	40
KP LM Stat	29.85		29.45		27.72	
KP LM p value	0.0000		0.0000		0.0000	

Table X: Flow GMM/IV Regression: Steps 1 and 2

Models 1 and 2 are Raw Peers, Models 3 and 4 Benchmark, Models 5 and 6 Delta. Odd numbered models are Step 1 regressions with multiple instruments, even numbered models are Step 2 regressions with a single Peer(Predicted Flow) instrument. Peer flow is defined with the given peer measure. Instrument in Step 1 regression the main peer lagged flow instrument, Step 2 uses the predicted output from Step 1 to form the Peer(Predicted Flow) instrument. Flow % is dollar fund flow divided by total net assets, globally differenced. Flow lags 2-4 are included but omitted, as is Cash % and expense ratio. A portfolio-weighted Amihud measure is also included but omitted. L Fund Size is the log of Total Net Assets in $millions. Turnover is as defined in Carhart (1997). Mkt Share Change is as defined in Spiegel and Zhang (2013). Volatility is rolling annual trailing fund return volatility. Style Flow Diff is differenced style average flow. Time and Fund Fixed Effects included, standard errors clustered by both time and fund. Hansen J stat is a test of overidentification for which the null hypothesis is that instruments are uncorrelated with stage 2 regression, KP LM stat tests the null of weak instruments. T statistics are in parentheses and significance is denoted at the 1, 5, and 10% level.

	Raw Peer		**Bench Peer**		**Delta Peer**	
	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow	(6) Flow
Peer Flow	1.3629***	0.9057***	1.1678***	1.1220***	0.8910***	0.8637***
	(6.51)	(10.22)	(5.58)	(12.58)	(16.54)	(23.32)
Lag1 Flow	0.0746***	0.0759***	0.0706***	0.0690***	0.0688***	0.0682***
	(4.96)	(5.04)	(4.51)	(4.38)	(4.59)	(4.56)
Lag1 Alpha	0.6892***	0.7343***	0.6968***	0.6863***	0.7010***	0.6928***
	(6.50)	(6.25)	(6.40)	(6.26)	(6.20)	(6.07)
Log Fund Size	-0.0303***	-0.0300***	-0.0300***	-0.0305***	-0.0301***	-0.0298***
	(-10.46)	(-10.21)	(-10.29)	(-10.43)	(-10.84)	(-10.69)
Log Fund Age (yrs)	-0.0106**	-0.0106**	-0.0116**	-0.0112**	-0.0085*	-0.0094*
	(-2.11)	(-2.08)	(-2.23)	(-2.15)	(-1.74)	(-1.90)
Mkt Shr Chg (bps)	0.0111***	0.0113***	0.0113***	0.0112***	0.0103***	0.0106***
	(4.99)	(4.95)	(4.70)	(4.65)	(4.80)	(4.91)
Style Flow Diff	0.2569***	0.3716***	0.2654***	0.2806***	0.3589***	0.3622***
	(4.18)	(8.41)	(3.96)	(6.27)	(9.67)	(10.42)
Volatility	0.0985	0.0904	0.1074	0.0997	0.0965	0.0937
	(1.41)	(1.34)	(1.45)	(1.36)	(1.46)	(1.42)
Turnover	-0.0147***	-0.0148***	-0.0154***	-0.0153***	-0.0130***	-0.0131***
	(-15.30)	(-15.15)	(-15.31)	(-15.05)	(-13.86)	(-13.78)
Observations	52281	52281	48454	48454	52278	52278
R Squared	0.09	0.09	0.09	0.09	0.10	0.10
Fund clusters	3,832	3,832	3,684	3,684	3,831	3,831
Time clusters	44	44	44	44	44	44
Hansen J stat	0.04	0.00	2.50	0.00	1.54	0.00
J p value	0.8476		0.1138		0.2146	
KP LM Stat	19.58	22.99	22.34	26.81	30.24	27.83
KP LM p value	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
IV Step	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2

Table XI: Flow GMM/IV Regression: Steps 1 and 2 with robust instruments. All is the same as Table X except the instrument specification used includes two peer lagged flow and two peer size, following Bramoullé, Djebbari, and Fortin (2009).

	Raw Peer		**Bench Peer**		**Delta Peer**	
	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow	(6) Flow
Peer Flow	1.4426***	0.9115***	1.0864***	1.1222***	0.7898***	0.8588***
	(6.92)	(10.28)	(5.44)	(12.57)	(17.45)	(23.20)
Lag1 Flow	0.0714***	0.0759***	0.0669***	0.0690***	0.0791***	0.0682***
	(1.99)	(2.22)	(2.05)	(2.17)	(2.00)	(2.05)
Lag1 Alpha	0.6788***	0.7337***	0.6464***	0.6863***	0.7351***	0.6935***
	(6.43)	(6.25)	(6.09)	(6.26)	(6.44)	(6.07)
Log Fund Size	-0.0299***	-0.0300***	-0.0294***	-0.0305***	-0.0277***	-0.0298***
	(-10.39)	(-10.21)	(-10.54)	(-10.43)	(-10.06)	(-10.69)
Log Fund Age (yrs)	-0.0113**	-0.0106**	-0.0103**	-0.0112**	-0.0106**	-0.0094*
	(-2.24)	(-2.08)	(-2.00)	(-2.15)	(-2.15)	(-1.90)
Mkt Shr Chg (bps)	0.0107***	0.0113***	0.0112***	0.0112***	0.0088***	0.0106***
	(4.84)	(4.95)	(4.69)	(4.65)	(4.16)	(4.91)
Style Flow Diff	0.2339***	0.3701***	0.2889***	0.2806***	0.3882***	0.3635***
	(3.83)	(8.39)	(4.38)	(6.27)	(11.10)	(10.45)
Volatility	0.1141*	0.0905	0.1070	0.0997	0.1399**	0.0935
	(1.65)	(1.34)	(1.44)	(1.36)	(2.27)	(1.42)
Turnover	-0.0146***	-0.0148***	-0.0156***	-0.0153***	-0.0132***	-0.0131***
	(-15.03)	(-15.15)	(-15.56)	(-15.05)	(-14.09)	(-13.79)
Observations	52281	52281	48454	48454	52278	52278
R Squared	0.09	0.09	0.09	0.09	0.10	0.10
Fund clusters	3,832	3,832	3,684	3,684	3,831	3,831
Time clusters	44	44	44	44	44	44
Hansen J stat	2.71	0.00	3.89	0.00	18.19	0.00
J p value	0.2573		0.1429		0.0001	
KP LM Stat	20.75	22.94	25.91	26.82	32.43	27.87
KP LM p value	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
IV Step	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2

Table XII: Coefficient Analysis of Flow Regression

Transformed coefficients from Models 2, 4, and 6 from Table X for Raw Peer measure, Bench Peer measure, and Delta Peer measure, respectively in Panels a, b, and c. Column 1 is coefficient estimate copied from Models 2, 4, and 6 in Table X, and column 4 is the corresponding standard error of that coefficient estimate. Column 2 is the average of the diagonal of the transformed matrix coefficient, and so summarized the average direct effect each fund has on itself. Column 3 is the average of the off-diagonal of the same matrix coefficient, and so summarized the average spillover effect each fund has on each other.

(a) Raw Peer

	Coeff Estimate	Direct Effect	Spillover Effect	Std Err of Coeff
Lag1 Flow	0.0792	0.0801	1.5343	0.0151
Lag1 Alpha	0.7390	0.7482	14.3237	0.1234
Fund Size	-0.0309	-0.0313	-0.5998	0.0029
Fund Age	0.0005	0.0005	0.0101	0.0032
Mkt Shr Chg	0.0100	0.0101	0.1937	0.0021
Style Flow Diff	0.3930	0.3979	7.6172	0.0418
Volatility	0.1212	0.1227	2.3499	0.0721
Turnover	-0.0158	-0.0160	-0.3055	0.0011

(b) Bench Peer

	Coeff Estimate	Direct Effect	Spillover Effect	Std Err of Coeff
Lag1 Flow	0.0775	0.0800	1.4087	0.0151
Lag 1Alpha	0.6755	0.7464	13.1505	0.1198
Fund Size	-0.0309	-0.0313	-0.5506	0.0028
Fund Age	0.0003	0.0005	0.0092	0.0032
Mkt Shr Chg	0.0097	0.0101	0.1778	0.0021
Style Flow Diff	0.3099	0.3969	6.9933	0.0389
Volatility	0.1197	0.1225	2.1574	0.0708
Turnover	-0.0158	-0.0159	-0.2805	0.0011

(c) Delta Peer

	Coeff Estimate	Direct Effect	Spillover Effect	Std Err of Coeff
Lag1 Flow	0.0712	0.0827	1.7054	0.0148
Lag1 Alpha	0.6991	0.7723	15.9207	0.1195
Fund Size	-0.0307	-0.0323	-0.6666	0.0027
Fund Age	0.0008	0.0005	0.0112	0.0032
Mkt Shr Chg	0.0094	0.0104	0.2153	0.0020
Style Flow Diff	0.3832	0.4107	8.4665	0.0319
Volatility	0.1270	0.1267	2.6119	0.0693
Turnover	-0.0140	-0.0165	-0.3395	0.0010

Table XIII: Economic Significance of Flow Coefficients (Average Effects)

Economic significance of coefficients Table XII for Raw Peer measure, Bench Peer measure, and Delta Peer measure, respectively in Panels a, b, and c. Each row represents a shock to the explanatory variable listed to one third of the funds in the network. The effect of the shock is measured in incremental flow %, differenced globally, which is the dependent variable in the main regression in Table X . The first column is the mean effect on the other two-thirds not experiencing the shock. Column 2 is the mean effect on those who do. The third is the ratio of the two and the last two columns display the effect divided by the standard deviation of the dependent variable to provide a frame of reference.

(a) Raw Peer

	Mean Spillover	Mean Direct Effect	Spillover/ Direct	Direct/ Std Flow %	Spillover/ Std Flow %
Lag1 Flow	0.0553	0.0657	0.8404	0.5200	0.4370
Lag1 Alpha	0.0275	0.0336	0.8199	0.2656	0.2178
Fund Size	-0.1270	-0.1485	0.8555	-1.1742	-1.0045
Fund Age	0.0018	0.0021	0.8676	0.0164	0.0142
Mkt Shr Chg	0.0507	0.0562	0.9016	0.4446	0.4008
Style Flow	0.0427	0.0503	0.8480	0.3982	0.3376
Volatility	0.0157	0.0184	0.8524	0.1452	0.1238
Turnover	-0.1069	-0.1240	0.8618	-0.9809	-0.8453

(b) Bench Peer

	Mean Spillover	Mean Direct Effect	Spillover/ Direct	Direct/ Std Flow	Spillover/ Std Flow
Lag1 Flow	0.0568	0.0671	0.8467	0.5285	0.4475
Lag1 Alpha	0.0313	0.0374	0.8377	0.2948	0.2470
Fund Size	-0.1318	-0.1526	0.8642	-1.2020	-1.0387
Fund Age	0.0018	0.0020	0.8700	0.0161	0.0140
Mkt Shr Chg	0.0477	0.0529	0.9021	0.4164	0.3756
Style Flow	0.0441	0.0515	0.8575	0.4056	0.3478
Volatility	0.0158	0.0184	0.8587	0.1451	0.1246
Turnover	-0.0954	-0.1113	0.8573	-0.8768	-0.7517

(c) Delta Peer

	Mean Spillover	Mean Direct Effect	Spillover/ Direct	Direct/ Std Flow	Spillover/ Std Flow
Lag1 Flow	0.0631	0.0734	0.8607	0.5802	0.4994
Lag1 Alpha	0.0323	0.0383	0.8434	0.3027	0.2553
Fund Size	-0.1426	-0.1634	0.8728	-1.2921	-1.1277
Fund Age	0.0019	0.0022	0.8778	0.0175	0.0153
Mkt Shr Chg	0.0534	0.0591	0.9040	0.4676	0.4227
Style Flow	0.0468	0.0542	0.8639	0.4287	0.3703
Volatility	0.0169	0.0195	0.8661	0.1543	0.1337
Turnover	-0.1127	-0.1291	0.8726	-1.0212	-0.8911

Table XIV: Economic Significance of Flow Coefficients (Distribution)

More detail on the last two columns of Table XIII. For select explanatory variables, percentiles of the underlying distribution of each fund manager's effect are listed, divided through by the standard deviation of globally differenced flow %.

(a) Raw Peer

Direct Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.1203	0.1677	0.2121	0.3479	0.6319	1.1170	1.9159	2.6250	4.4064
Lag1 Alpha	0.0306	0.0598	0.0899	0.1749	0.3388	0.6188	1.1017	1.4485	2.3993
Mkt Shr Chg	0.0409	0.0825	0.1262	0.2581	0.5297	1.0359	1.8812	2.6335	5.4071
Volatility	0.0164	0.0338	0.0492	0.0964	0.1817	0.3372	0.5859	0.8089	1.3404

Spillover Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.0437	0.0857	0.1312	0.2492	0.5175	1.0078	1.8126	2.4756	4.2134
Lag1 Alpha	0.0223	0.0422	0.0636	0.1220	0.2597	0.5093	0.8929	1.2592	2.0957
Mkt Shr Chg	0.0409	0.0784	0.1168	0.2252	0.4740	0.9368	1.6707	2.3300	3.9633
Volatility	0.0125	0.0248	0.0372	0.0703	0.1460	0.2859	0.5146	0.7022	1.1893

(b) Bench Peer

Direct Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.1160	0.1610	0.1988	0.3145	0.5993	1.0991	1.9093	2.5934	4.2579
Lag1 Alpha	0.0300	0.0668	0.0926	0.1743	0.3455	0.6881	1.1949	1.6263	2.6170
Mkt Shr Chg	0.0334	0.0686	0.1021	0.2129	0.4625	0.9833	1.7592	2.5615	5.0508
Volatility	0.0145	0.0317	0.0449	0.0858	0.1698	0.3337	0.5827	0.8097	1.2807

Spillover Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.0362	0.0814	0.1284	0.2540	0.5456	1.0369	1.8442	2.6031	4.5085
Lag1 Alpha	0.0200	0.0434	0.0710	0.1430	0.3061	0.5919	1.0441	1.4747	2.4961
Mkt Shr Chg	0.0294	0.0642	0.1051	0.2123	0.4588	0.8942	1.6044	2.3006	3.9368
Volatility	0.0102	0.0229	0.0365	0.0708	0.1523	0.2899	0.5149	0.7320	1.2526

(c) Delta Peer

Direct Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.1406	0.1846	0.2307	0.3837	0.6853	1.2928	2.2340	2.9819	5.3073
Lag1 Alpha	0.0398	0.0728	0.1010	0.1885	0.3676	0.7337	1.2247	1.6817	2.8311
Mkt Shr Chg	0.0463	0.0779	0.1178	0.2354	0.4969	1.0437	1.8772	2.8026	5.2537
Volatility	0.0207	0.0389	0.0542	0.0992	0.1920	0.3695	0.6520	0.8796	1.5447

Spillover Effect Distribution Percentiles

	1	5	10	25	50	75	90	95	99
Lag1 Flow	0.0523	0.1025	0.1584	0.3095	0.6490	1.2267	2.1992	3.0205	4.9990
Lag1 Alpha	0.0257	0.0506	0.0777	0.1542	0.3280	0.6200	1.1027	1.5148	2.4825
Mkt Shr Chg	0.0380	0.0750	0.1160	0.2311	0.4935	0.9548	1.6920	2.3504	3.9256
Volatility	0.0142	0.0285	0.0438	0.0834	0.1751	0.3308	0.5933	0.8194	1.3500

Table XV: DGTW Bench Peer Flow Between-Within

All variables as defined in Table VIII, as are the econometrics of the specification. Volatility and Style Flow Diff are also included but not displayed due to space considerations. Models 1, 4, and 7 are reproduced from Table X for reference. The subsequent two models are subset to Between style peers and Within style peers to isolate the any effect from style investing. T statistics are in parentheses and significance is denoted at the 1, 5, and 10% level.

	Raw Peer			**Bench Peer**			**Delta Peer**		
	(1) DGTW	(2) DGTW	(3) DGTW	(4) DGTW	(5) DGTW	(6) DGTW	(7) DGTW	(8) DGTW	(9) DGTW
	All	Between	Within	All	Between	Within	All	Between	Within
Peer Flow	1.6707***	0.7649***	4.6075***	1.2307***	1.0771***	0.4794***	0.1728***	1.3465***	0.0800***
	(4.16)	(3.78)	(2.63)	(5.47)	(4.45)	(4.72)	(4.61)	(3.69)	(2.82)
Lag1 Flow	-0.0036	-0.0051	-0.0124**	-0.0046	-0.0034	-0.0044	-0.0039	-0.0118**	-0.0030
	(-0.82)	(-1.31)	(-2.27)	(-1.02)	(-0.86)	(-1.11)	(-0.90)	(-2.31)	(-0.77)
Lag1 Return	0.0167	0.0410	-0.0801	0.0258	0.0351	0.0470	0.0569	0.0212	0.0520
	(0.23)	(0.55)	(-0.97)	(0.34)	(0.46)	(0.59)	(0.74)	(0.28)	(0.67)
Log Fund Size	-0.0112***	-0.0090***	-0.0101***	-0.0105***	-0.0098***	-0.0098***	-0.0105***	-0.0102***	-0.0101***
	(-7.16)	(-6.67)	(-5.41)	(-6.81)	(-6.85)	(-6.64)	(-6.67)	(-7.27)	(-7.00)
Log Fund Age (yrs)	0.0020	0.0033***	0.0048***	0.0020	0.0035***	0.0037***	0.0016	0.0035***	0.0034***
	(0.97)	(2.79)	(2.99)	(1.01)	(2.98)	(3.27)	(0.74)	(2.70)	(2.71)
Mkt Shr Chg (bps)	0.0073***	0.0214***	0.0175***	0.0072***	0.0069***	0.0072***	0.0076***	0.0063***	0.0072***
	(5.79)	(8.06)	(5.94)	(5.49)	(4.78)	(4.82)	(5.68)	(5.19)	(4.98)
Turnover	0.0003	0.0007**	0.0013**	0.0003	0.0002	0.0001	0.0004	0.0026***	0.0003
	(0.93)	(2.43)	(2.45)	(0.86)	(0.58)	(0.21)	(1.32)	(3.26)	(0.90)
Observations	46723	51189	51070	43343	46916	46361	46722	51187	51047
R Squared	0.15	0.15	-0.87	0.15	0.13	0.11	0.11	-0.09	0.10
Fund clusters	3,432	3,688	3,679	3,289	3,489	3,463	3,432	3,688	3,677
Time clusters	44	44	44	44	44	44	44	44	44
KP LM Stat	29.85	28.86	6.21	29.45	31.33	31.00	27.69	23.79	29.46
KP LM p value	0.0000	0.0000	0.0127	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table XVI: Flow Regression divided into All-Between-Within

All variables as defined in Table X, as are the econometrics of the specification. Volatility and Style Flow Diff are also included but not displayed due to space considerations. Models 1, 4, and 7 are reproduced from Table X for reference. The subsequent two models are subset to Between style peers and Within style peers to isolate the any effect from style investing. T statistics are in parentheses and significance is denoted at the 1, 5, and 10% level.

	Raw Peer			**Bench Peer**			**Delta Peer**		
	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow	(6) Flow	(7) Flow	(8) Flow	(9) Flow
	All	Between	Within	All	Between	Within	All	Between	Within
Peer Flow	0.9057***	0.9438***	2.6636***	1.1220***	1.1349***	0.9612***	0.8637***	1.8290***	0.8572***
	(10.22)	(8.46)	(4.00)	(12.58)	(13.74)	(9.32)	(23.32)	(8.83)	(19.80)
Lag1 Flow	0.0759***	0.0754***	0.0676***	0.0690***	0.0762***	0.0713***	0.0682***	0.0650***	0.0670***
	(5.04)	(5.02)	(4.50)	(4.38)	(4.79)	(4.41)	(4.56)	(4.39)	(4.52)
Lag1 Alpha	0.7343***	0.7884***	0.6462***	0.6863***	0.7278***	0.6665***	0.6928***	0.6084***	0.6731***
	(6.25)	(6.04)	(5.16)	(6.26)	(6.07)	(5.49)	(6.07)	(5.39)	(5.52)
Log Fund Size	-0.0300***	-0.0286***	-0.0299***	-0.0305***	-0.0314***	-0.0323***	-0.0298***	-0.0314***	-0.0306***
	(-10.21)	(-10.72)	(-12.87)	(-10.43)	(-11.50)	(-12.15)	(-10.69)	(-12.58)	(-12.27)
Log Fund Age (yrs)	-0.0106**	0.0003	0.0022	-0.0112**	0.0009	0.0019	-0.0094*	0.0013	0.0008
	(-2.08)	(0.09)	(0.72)	(-2.15)	(0.26)	(0.52)	(-1.90)	(0.39)	(0.27)
Mkt Shr Chg (bps)	0.0113***	0.0325***	0.0295***	0.0112***	0.0100***	0.0102***	0.0106***	0.0087***	0.0091***
	(4.95)	(5.85)	(5.23)	(4.65)	(4.39)	(4.41)	(4.91)	(4.60)	(4.64)
Turnover	-0.0148***	-0.0150***	-0.0147***	-0.0153***	-0.0165***	-0.0167***	-0.0131***	-0.0125***	-0.0134***
	(-15.15)	(-14.53)	(-13.79)	(-15.05)	(-15.39)	(-15.48)	(-13.78)	(-10.63)	(-12.94)
Observations	52281	57306	57141	48454	52477	51736	52278	57302	57084
R Squared	0.09	0.11	0.05	0.09	0.09	0.08	0.10	0.07	0.09
Fund clusters	3,832	4,103	4,101	3,684	3,893	3,867	3,831	4,102	4,096
Time clusters	44	44	44	44	44	44	44	44	44
KP LM Stat	22.99	27.76	8.88	26.81	27.14	28.92	27.83	20.43	29.61
KP LM p value	0.0000	0.0000	0.0029	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table XVII: Competitive Effects

Different from all previous specifications, this table uses category-differenced flows, meaning each fund flow % has subtracted from it the Morningstar category average flow less that fund's flow. In keeping with category-based differencing, group flow control is also based on Morningstar categories rather than styles: Category Flow Diff is used in place of Style Flow Diff, computed in a similar manner. Peer measure is based on raw weights. Models 1-3 are flow regressions (Step 2 of the two step GMM/IV sequence) and models 4-6 are DGTW return specifications, also using Peer predicted flow as a single instrument. Model 3 and 6 are the Within subset showing the competitive effects as in Wahal and Wang (2011). All other variables are as defined in Tables X and VIII, and econometrics are also the same.

	(1) Flow	(2) Flow	(3) Flow	(4) DGTW	(5) DGTW	(6) DGTW
Peer Subset	All	Between	Within	All	Between	Within
Peer Flow	1.4427***	0.9691***	-2.4379**	2.2966***	1.0299***	-0.1695
	(4.95)	(4.14)	(-2.33)	(4.36)	(4.47)	(-0.23)
Lag1 Flow	0.0722***	0.0745***	0.0771***	0.0017	-0.0004	0.0005
	(4.90)	(4.98)	(5.12)	(0.53)	(-0.16)	(0.20)
Lag1 Alpha	0.7559***	0.7725***	0.8660***			
	(6.77)	(6.43)	(6.14)			
Lag1 Return				0.0224	0.0340	0.0436
				(0.31)	(0.46)	(0.55)
Log Fund Size	-0.0292***	-0.0283***	-0.0283***	-0.0113***	-0.0094***	-0.0091***
	(-10.26)	(-10.86)	(-10.41)	(-7.09)	(-7.05)	(-7.12)
Log Fund Age	-0.0107**	0.0001	-0.0007	0.0028	0.0036***	0.0033***
	(-2.16)	(0.03)	(-0.22)	(1.39)	(3.22)	(2.96)
Mkt Shr Chg (bps)	0.0109***	0.0314***	0.0327***	0.0072***	0.0206***	0.0210***
	(4.87)	(5.77)	(5.86)	(5.73)	(8.08)	(8.28)
Category Flow Diff	-0.3493***	-0.3300***	-0.4215***	0.3254***	0.3259***	0.3321***
	(-4.85)	(-4.32)	(-7.30)	(6.73)	(7.38)	(4.71)
Volatility	0.1038	0.1742**	0.1707*	-0.0931	0.0271	0.0298
	(1.53)	(2.05)	(1.85)	(-0.31)	(0.09)	(0.10)
Exp Ratio	-0.0160	-0.0114	-0.0166	-0.0095**	-0.0089*	-0.0106**
	(-1.13)	(-0.88)	(-1.23)	(-2.01)	(-1.92)	(-2.29)
Turnover	-0.0148***	-0.0152***	-0.0158***	0.0003	0.0008**	0.0006
	(-14.73)	(-14.10)	(-13.64)	(1.14)	(2.54)	(1.44)
Observations	52244	57254	57245	46698	51149	51137
R Squared	0.07	0.09	0.03	0.14	0.16	0.15
Fund clusters	3,828	4,098	4,097	3,429	3,684	3,682
Time clusters	44	44	44	44	44	44
KP LM Stat	29.00	24.43	8.97	28.65	25.61	12.65
KP LM p value	0.0000	0.0000	0.0027	0.0000	0.0000	0.0004