# Contagious Capital: A Network Analysis of Interconnected Intermediaries

JESSE BLOCHER*

January 12, 2012

**ABSTRACT**

I measure the effects of capital flow contagion in financial markets by analyzing portfolio managers linked through interconnected asset holdings. My novel, network-based specification provides estimates of shocks to common predictor variables 50-75% higher than existing estimates of manager's capital flows which ignore network relationships. This additional impact arises because my network specification includes the effect of spillover onto immediate neighbors and beyond, leading to feedback loops. My findings seem to result from crowded trades (popular, short-term investment strategies) since network connections do not show strong persistence and relatively small changes in asset allocation toward more concentrated positions may increase interconnection considerably.

*Special thanks to Jennifer Conrad, Adam Reed and Joey Engelberg, Greg Brown and Pab Jotikasthira for timely and helpful advice on this project. I am also grateful to Rick Sias (Discussant at FRA conference), Chris Lundblad, Ed Van Wesep and Matt Ringgenberg for useful comments. Thanks to Matthew Bothner (ESMT Management/Sociology) and Peter Mucha (UNC Applied Mathematics) for valuable guidance on network measures and computational algorithms, respectively. Ethan Cohen-Cole and Tim Conley provided help and references on identification and econometrics. James Moody and seminar participants at the Duke Network Analysis Center (dnac.ssri.duke.edu) provided useful feedback, as did seminar participants at the SAMSI (www.samsi.info) program on Complex Networks. Blocher is from the University of North Carolina - Chapel Hill Kenan-Flagler Business School, Finance Area, Campus Box 3490 McColl Building, Chapel Hill, NC 27599-3490. Contact info: jesse_blocher@unc.edu or www.unc.edu/~jabloche.

Since the beginning of the recent financial crisis, the concept of too-interconnected-to-fail has grown in importance, leading regulators to identify portfolio overlaps of financial intermediaries as a potential source of systemic risk.[1] Practitioners have also shown concern, suggesting that "...there may be more crowded trades than most investors realize. If investors exit at the same time, market movements could be chaotic."[2] Among academics, Stein (2009) identifies "crowding", or similar portfolios among sophisticated investors, as a risk in financial markets and Brunnermeier and Sannikov (2011) identify portfolio overlaps as a destabilizing mechanism in financial markets.

In this paper, I show that crowded trades may induce capital flow contagion among these interconnected portfolio managers. Capital flow contagion occurs when the withdrawals and forced sales experienced by one investment manager provoke capital outflows and asset sales from other funds with similar portfolio holdings through the depressed prices of commonly held assets (Brunnermeier and Pedersen (2009)).[3] However, my approach accounts for broader network propagation effects and feedback loops, not just pairwise connections. To do this, I employ a novel instrumental variables specification to estimate contemporaneous capital flow contagion effects in steady-state across the network.

Compared to the analysis of disconnected, independent portfolio managers common in the literature, I find that coefficient estimates of common predictors of fund flows increase by 50-75% when network relationships are taken into account.[4]

[1] See speech by The Bank of England's Executive Director for Financial Stability Andrew Haldane at http://www.bankofengland.co.uk/publications/speeches/2010/speech433.pdf

[2] Bank of America-Merrill Lynch report, http://ftalphaville.ft.com/blog/2011/06/01/581676/the-calm-before-the-volatility-storm/ as quoted in the Financial Times, 1 June 2011.

[3] Brunnermeier and Pedersen (2009) model withdrawals and forced sales as a "market liquidity/funding liquidity" spiral, my extension is to consider what happens to other investors holding the assets being sold.

[4] Existing literature predicting fund flows assumes each fund to be independent (e.g. Sirri and Tufano (1998)). The common predictors of fund flows I consider are past returns, fund category average flows, and cash holdings.

This increase is due to two contagion processes I am able to incorporate with the full network of interconnections. First, "own" effects increase by up to 20% due to feedback loops in which a shock to a manager propagates out and back via a sequence of connected peers. Second, spillover effects (assumed to be zero in non-network specifications) are substantial and increase estimates by an additional 30-55%. Spillover effects are similar to a network externality in which a shock to a manager spills over onto his neighbors, such that an unsuspecting manager may find his portfolio under stress due to funding problems by others holding a similar set of positions. Coval and Stafford (2007) and Lou (2010) have established that fund flows impact asset prices. My innovation is to consider the effect that these fund flows may have on the fund flows of neighbors holding those same assets, since the same flow-performance relationship holds (Chevalier and Ellison (1997)) even if asset prices change due to a peer's forced sale.

I measure these effects with a network-based specification which includes connections between portfolio managers along with their capital flows at time $t$. This contemporaneous specification allows me to estimate cross-sectional steady-state peer influence processes so the effect of each portfolio manager on each other manager is estimated simultaneously. To identify this influence process, I specify the two-step neighbor's capital flow as an instrument. This is a valid instrument if enough two-step neighbors are not themselves connected to the manager of interest.[5]

I also show that the flows of connected neighbors are positively and significantly correlated with a manager's portfolio return; including these networked flow

[5] A "two-step neighbor" is simply my neighbor's neighbor. 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 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.

measures significantly reduces the influence of market returns and fund category average flows as predictors. This is a remarkable result since a portfolio manager's own lagged fund flows show no significance in predicting returns (Frazzini and Lamont (2008)). It is also consistent with a contagion process across managers connected by common holdings, since inflows would induce buying, and outflows selling, of at least a portion of the commonly held portfolio.

To fully identify my network effect, I control for other possible explanations of correlated flows. Specifically, since Sirri and Tufano (1998) show that the size of a mutual fund may influence investor flows due to search costs, I control for both a manager's own total net assets and neighbor's total net assets. In addition, since investor sector rotation strategies or other strategic asset allocation decisions may induce flows to common categories, I include a category average fund flow, similar to a Fama-French industry factor, as a control variable.

Given the result that capital flows seem to be contagious across similar portfolios, I next address the nature of these portfolio connections. It may be that such connections are relatively static, simply the result of natural linkages among varied strategies which are time invariant. But it may also be that portfolio connections are transient and related to crowded trades, such that interconnectedness may grow unobserved. To investigate these two hypotheses, I measure the persistence of network connections through time. Static network connections should show significant autoregressive properties, while transient crowded trades should show no long-term temporal predictability among portfolio connections. I show that these connections among portfolio managers are somewhat persistent short term, with network connections this quarter correlating 0.4 with last quarter's portfolio connections. However, the correlation across years is approximately 0.13, with no correlation after two years. Since fund objectives are likely to persist across several years, this suggests that shorter term connections are at least partially driving my result.

To further investigate the nature of these portfolio connections, I demonstrate that the similarity of two portfolios increases not only in terms of portfolio overlap, but also with concentration in those commonly held assets. That is, two managers who overlap 20% of their portfolio will be twice as connected if that overlap is in one holding than if it is equally held in two holdings. The implication is that a mid-cap fund which holds hundreds of securities may not connect other portfolios together as much as a fund with a few concentrated positions. It also means that small movements toward more concentrated holdings may induce significant connections to the extent that others hold similarly concentrated holdings. Small amounts of overweighting compared to the manager's benchmark may induce more interconnection than a portfolio manager realizes.

First, I establish my hypotheses in the context of existing literature in Section I. Next, in Section II, I describe my empirical approach to measuring capital flow contagion, detailing network formation, measures, and methodology. In Section IV, I discuss my results, including the interpretation of network coefficients and their economic significance. I then further analyze the time-varying properties of my network and its relationship to crowded trades in Section V, after which I conclude with Section VI.

## I. Hypotheses and Background

Quantifying the effects of capital flow contagion through interconnected asset holdings implies hypotheses related to the prediction of mutual fund returns and mutual fund flows. I first develop my hypothesis related to the prediction of portfolio returns which helps establish portfolio overlaps as the mechanism for contagion. Second, I develop two hypotheses related to the spillover effects of manager's fund flows.

I propose that interconnected managers' capital flows influence each other in the following manner: inflows to neighboring portfolio managers induce purchases of their existing portfolio and outflows induce sales, temporarily affecting the prices of those assets bought or sold. But connected portfolio managers holding those same assets should see their portfolio returns affected in a corresponding manner, such that the capital flows of connected portfolio managers positively predict portfolio returns. Subsequently, since negative returns predict outflows and positive returns predict inflows (Chevalier and Ellison (1997)), these affected managers may experience their own inflows or outflows, perhaps beginning a market/funding liquidity spiral (Brunnermeier and Pedersen (2009)). Thus, peer flows predict a portfolio manager's returns suggesting that interconnected portfolios are an important channel for capital flow contagion in financial markets.

The fact that a portfolio manager's fund flows affect the assets he holds is known. Coval and Stafford (2007) show that stocks with significant buying or selling pressure experience subsequent positive and negative returns, respectively. Lou (2010) addresses this question across all fund flows, not just extreme positive and negative flows, and shows that this effect is still significant but asymmetric – he estimates that one dollar of inflows correlates with purchasing 0.6 dollars of the existing portfolio, while one dollar of outflows corresponds to selling one dollar of the existing portfolio.

The "no arbitrage" condition in financial markets indicates that this mispricing should be small or very short-term. But given that source of price pressure may be hidden (e.g., Kyle (1985)), arbitrageurs may not identify a price movement as a deviation from fundamentals, and thus not act to correct it. Arbitrageurs also face synchronization risk (Abreu and Brunnermeier (2002, 2003)), since multiple arbitrageurs may be necessary to absorb the price pressure, as well as other limits to arbitrage (e.g., Shleifer and Vishny (1997)). Indeed, rather than immediately arbitraging an over- or under-pricing, these sophisticated investors may even exacerbate the problem in a predatory manner to increase the mispricing and thus the profitability of a subsequent convergence trade (Brunnermeier and Pedersen (2005)).

To identify common portfolio holdings as a channel of contagion, I hypothesize that the fund flows of a manager's connected neighbors predict portfolio returns through the buying and selling of the commonly held assets. Formally,

Hypothesis 1: *The fund flows of neighbors connected by common asset holdings positively predict a manager's portfolio return.*

To test this hypothesis, I compute a measure of connected-neighbor fund flows weighted by portfolio similarity, and then estimate its impact on portfolio returns. To determine my baseline and control variables, I draw from the existing literature known predictor variables of mutual fund returns. I include the market return which Carhart (1997) shows to be an important predictor of mutual fund returns, as well as past flows to account for the flow-performance relationship established in Chevalier and Ellison (1997). Since contemporaneous fund flows and portfolio returns may suffer from endogeneity, I instrument peer flows in a GMM framework, discussed in detail in Section II.C.

If connected-neighbor's flows positively predict a manager's portfolio returns, the next logical step is to consider the effect on that manager's fund flows, since returns affect future fund flows. Chevalier and Ellison (1997) identify a performance-to-flow relationship such that positive past returns predict future inflows and poor past returns predict future outflows. While Chevalier and Ellison measure these effects through lagged returns, these outflows could be contemporaneous since a sophisticated manager, seeing his poor returns, may sell in anticipation of future outflows. Until now, investigating this relationship has been challenging due to the endogeneity problem between contemporaneous flows and returns, a problem I solve with my instrumental variables specification. This connection between the capital flows of neighboring managers suggests two related hypotheses:

HYPOTHESIS 2: *The fund flows of neighbors connected by common asset holdings positively predict a manager's own fund flows.*

HYPOTHESIS 3: *Spillover effects from each manager onto each other manager are nonzero.*

While I could test Hypothesis 2 with lagged connected-neighbor fund flows in a simple panel framework, that same specification would only provide indirect support for Hypothesis 3. To test both hypotheses, I employ a network specification which allows a contemporaneous equilibrium estimation of spillover effects across a network of connected agents. In this network specification, I include other common predictors of capital flows such as past returns (e.g.,Chevalier and Ellison (1997)), past flows (e.g., Coval and Stafford (2007)), total net assets, and fund category average flows (e.g., Sirri and Tufano (1998)). I include my measure of connected-neighbor fund flows as a predictor variable, instrumented by the two-step neighbor fund flows. If the coefficient on this measure of peer's capital flows is positive and significant, this confirms Hypothesis 2: capital is contagious through interconnected portfolios.

While a positive and significant relationship establishes the existence of a contagion process, obtaining evidence for Hypothesis 3 requires interpreting the resulting coefficient estimate. Indeed, the richness of information available from this network specification constitutes a primary advantage over a standard linear regression model. This specification behaves like an autoregression, but in the cross-section: fund flows at time $t$ show up both as dependent and independent variables, and as such the estimated coefficient on connected-neighbor flows affects all other coefficient estimates in steady-state, similar to an temporal autoregression framework.[6] When the model is rearranged such that flows are only the dependent variable, the coefficient on each independent variable becomes a matrix specifying the effect each portfolio manager has on each other manager in equilibrium.[7] This compares to the scalar coefficient estimating the average effect in most other specifications. The average of the off-diagonals of these matrix coefficients measures the spillover effects, while the average of the diagonal in excess of the non-networked linear coefficient measures feedback effects. Nonzero off diagonals in this matrix coefficient provides evidence of Hypothesis 3.

To test these hypotheses, I need to more fully specify the connection between portfolio managers and how I measure the neighbor's capital flows and estimate my network specification. This is the topic of the next section.

## II. Network Methodology

My network relationship derives from the connections among portfolio managers due to common asset holdings but there are many concepts of interconnection in financial markets. Allen, Babus, and Carletti (2010) and Zawadowski (2011) model connections among financial intermediaries in the interbank market and Babus (2010) does the same for OTC markets. Their analysis focuses on coun-

[6]Specifically, this model is a Spatial Auto-Regression (SAR), which is popular in spatial econometrics.

[7]I develop this more rigorously in Section IV.B.

terparty relationships in a game-theoretic framework in which relationships are typically known and intentionally created by each market participant. My measure of interconnection attempts to identify crowded trades in financial markets as a separate source of connectedness.

Others have studied the effect of common owners on financial assets. Kyle and Xiong (2001) model convergence traders spanning disparate markets inducing comovement in the assets they hold, and more recently Anton and Polk (2010) measure stock comovement as it relates to the number of common owners. Coval and Stafford (2007) and Jotikasthira, Lundblad, and Ramadorai (2011) show that funding pressure on owners affects the assets they hold, inducing price drops in those assets in U.S. and international settings, respectively. My innovation is to consider what happens to other managers holding the same assets with no funding pressure of their own, or spillover effects.

To describe my methodology in more detail, I describe my Data in Section A. I then develop my portfolio similarity measure in detail in Section B before proceeding to descriptions of my GMM estimation approach, network instrument, and full specification in Section C.

### *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.[8] Flows of funds are a simple dollar value per fund, per month or quarter. Note that my

[8] Elton, Gruber, Blake, Krasny, and Ozelge (2009) perform a thorough comparison with the more commonly used data from Thomson Reuters. They also highlight the importance of monthly observations of holdings since more frequent observations capture round trip within-quarter trades otherwise missed, but there is only a subset of funds for which monthly data is available. They also note that this appears to be a representative subsample, and so is unbiased for inference in many areas. But for my purposes, since I am investigating more aggregate effects, I require the entire population and so focus on quarterly observations.

data includes reported values for both fund flows and portfolio returns, whereas other studies typically compute fund flows from returns and changes in total net assets. Because this data includes many bond funds and I want to be as inclusive as possible, I keep any fund with nonzero equity position. I combine this data with CRSP by CUSIP when necessary to obtain stock characteristics.

Importantly, this data contains the entire portfolio holdings of each open ended fund. This means I have quarterly observations of each fund's cash holdings as well rather than the less frequent annual measures reported in the CRSP Mutual Fund database. In what follows, *Flow* is always fund flow divided by total net assets as in Coval and Stafford (2007) and *Size* is the log of total net assets. *Cash* is defined as currency, treasuries, and other cash-like holdings, also divided by total net assets. I compute a fund-level *Amihud* measure which is the portfolio weighted sum of each equity holding's individual Amihud measure over the previous quarter. Summary statistics of these measures as well as peer measures are available in Table I.

### *B. Portfolio Similarity Measure – The Network*

My data represents a set of portfolio managers with detailed holdings data through time, but for simplicity, I drop the $t$ index for this exposition and compute these measures for each $t$. I construct the similarity between two portfolios, $s_{ij}$ 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] Specifically,

$$s_{ij} = \frac{s_i \cdot s_j}{|s_i|\,|s_j|} \tag{1}$$

where for each manager $i$, the Euclidean norm is defined across $M$ securities as

$$|s_i| = \sqrt{\sum_{m=1}^{M} s_{im}^2} \tag{2}$$

Deriving this same measure in matrix form, let $H$ be the 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 measure is then

$$S = \frac{H^T H}{|H| \cdot |H|} \tag{3}$$

in which each $s_{ij}$ already defined above is an element of symmetric similarity matrix $S$. 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

$$|H_j| = \sqrt{\sum_{m=1}^{M} h_{jm}^2} \tag{4}$$

Figure 1 plots percentiles of the distribution of this portfolio distance measure through time.

To construct *Peer Flow* for each manager $i$, I compute a weight vector which

[9] Note that this similarity measure is the same as the cosine of the angle between the two vectors in security space. An alternative choice, popular in social network analysis, is the Euclidean distance, which is the square root of the dot product. While not literally a linear transformation, it is similar to the normalized dot product. The Euclidean distance has two main downsides: first, the maximum is not clearly defined and may vary from network to network, and second, it needs to be inverted since higher values indicate greater distance – it is strictly a *dissimilarity* measure. A typical transform is $\left(\max_{ij} d_{ij}\right) - dij$ for each distance $d_{ij}$ to make it a similarity measure so higher weights go to closer peers.

is each similarity measure $s_{ij}$ divided by the sum over all similarities, setting self-similarity $s_{ii}$ to 0. 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

$$PeerWeight_{ij} = \frac{s_{ij}}{\sum_{k} s_{ik}}, k \neq i \tag{5}$$

and *Peer Flow* is thus

$$PeerFlow_i = \sum_{k} PeerWeight_{ik} Flow_k \tag{6}$$

For example, consider a portfolio manager with three neighbors at distances 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 $(.25 * .01) + (.5 * .05) + (.25 * .10) = 0.0525$.

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 $t$. Note that I also compute other peer variables such as peer return, peer size (total net assets), and peer cash (divided by total net assets) in the same way.[10]

### *C. Network Structure as Instrument*

Since cross-sectional fund flows and returns of each portfolio manager at time $t$ are endogenous, I employ an instrument to identify influence rather than just

[10] This notion of portfolio distance is intuitively and mathematically similar to that of social distance as in Conley and Topa (2002).

correlation.[11] Without instrumentation, a correlation between two portfolio manager's fund flows is not sufficient evidence of one's influence on the other.

Following Bramoullé, Djebbari, and Fortin (2009), I employ a network-structure based instrument to address this endogeneity based on "intransitive triads" which are often present in a network. 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 relationship with B. In network terminology, A and C are *Two-Step* neighbors, so my instrument is *TwoStepPeerFlow*.

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 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.

However, not all two-step neighbors form intransitive triads. Additionally, while two portfolio managers may not be directly connected, they both maintain some set relationship to market-wide movements. 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 the instruments are not collinear with the endogenous variable. To further test the validity of my instruments, I compute various tests of weak instruments as well as Hansen's J test of overidentification in all specifications. All reported GMM specifications have results consistent with

[11] Since the diagonal of weighting matrix $W$ is set to zero, $Flow_i$ is never on both sides of the same specification, so there is no mechanical collinearity, only endogeneity.

strong instruments and no correlation of instruments with the second stage residual, thereby indicating a valid specification.

Mathematically, two-step neighbors are computed as $B = S^2$, which is 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.[12] In summation notation, the equivalent product is

$$b_{ij} = \sum_{q=1}^{N} s_{iq} s_{qj},\ q \neq i, j \tag{7}$$

with the diagonal of $B$ also set to zero such that a manager cannot be his own two-step neighbor.[13] 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

$$\widetilde{w}_{ji} = \frac{b_{ji}}{\sum_k b_{jk}} \tag{8}$$

$$TwoStepPeerFlow_j = \sum_k \widetilde{w}_{jk} Flow_k \tag{9}$$

To ensure overidentification, I include not just $TwoStepPeerFlow$ but also $TwoStepPeerFlow^2$ as excluded instruments, which is standard in an IV specification.

To test my first hypothesis, I instrument for peer fund flows as described above,

[12]A nonzero diagonal indicates a 'self-loop.' So, 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 is just a one-step neighbor.

[13]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.

but place portfolio returns as my dependent variable. Specifically, I estimate:

$$PeerFlow_{it} = TwoStepPeerFlow_{it} + TwoStepPeerFlow_{it}^2 \tag{10}$$

$$\begin{aligned} Ret_{it} = {} & \widehat{PeerFlow}_{it} + Flow_{t-p} + Ret_{t-p} \\ & + Size_{it} + Cash_{it} + Amihud_{it} + PeerSize_{it} \\ & + PeerCash_{it} + CategoryAvgFlow_{jt} + MarketReturn_t \end{aligned} \tag{11}$$

with the primary explanatory variable being $MarketReturn$ in a CAPM style framework.[14] If $PeerFlow$ is a positive predictor of portfolio returns, then it seems highly likely that commonly held assets are the channel of influence.

Next, to test my second hypothesis that capital flows are contagious, I incorporate my network measure in addition to common predictor variables in a specification with fund flows as the dependent variable. Coval and Stafford (2007) employ both lagged flows and lagged returns as predictors, and Sirri and Tufano (1998) show that fund category averages and fund size (measured as log of total net assets) are important determinants of flows given investors' non-zero search costs. Since temporary asset price movements may be stronger for illiquid securities, I include a portfolio-wide *Amihud* measure which is simply the weighted average of the Amihud liquidity measure computed for each individual equity holding (Amihud (2002)).[15]

Since fund size is an important predictor of flows, I also include *PeerSize* as a control variable. This control is important in a network specification because if flows primarily go to larger funds (Sirri and Tufano (1998)), then funds who

[14] Carhart (1997) notes that this CAPM specification is gives similar results to his 3 factor model.

[15] I also computed a full portfolio Amihud measure including cash and non-equity, non-cash holdings at the minimum and maximum Amihud measure, respectively, with similar results. Computed portfolio spreads and average daily volumes also gave similar results, available on request

are both large and connected may simply experience correlated flows without any mutual influence.

A portfolio manager's cash holdings provide a vital cushion against unexpected redemptions, and as such they likely influence the prediction of inflows and outflows. Most studies exclude cash holdings because the data is unavailable, not because cash holdings are unimportant. Because I do have this data, I include it for both the manager and connected neightbors (*PeerCash*), since a manager connected to cash-poor neighbors may be more susceptible to flow contagion.

In sum, I estimate the following set of equations in a GMM specification:

$$PeerFlow_{it} = TwoStepPeerFlow_{it} + TwoStepPeerFlow_{it}^2 \tag{12}$$

$$\begin{aligned} Flow_{it} = {} & \widehat{PeerFlow}_{it} + Flow_{t-p} + Ret_{t-p} \\ & + Size_{it} + Cash_{it} + Amihud_{it} + PeerSize_{it} \\ & + PeerCash_{it} + CategoryAvgFlow_{jt} \end{aligned} \tag{13}$$

in which $Fund_i \in Category_j$, 4 time lags are included ($p = 4$) and $\widehat{PeerFlow}_{it}$ is the fitted values from equations (12).[16]

## III. Identification and Estimation of a Network Influence Process

In addition to the more standard identification problems already addressed, there are some unique identification problems associated with network inference, which I now address following the typology in Manski (1993). According to Manski (1993), identifying the endogenous social influence process I have just described requires controlling for two other potential confounding effects: "correlated effects"

[16] Note that the exact specification of equation (12) includes all control variables in equation (13). To use strict GMM terminology, *PeerFlow* is the endogenous regressor, $TwoStepPeerFlow$ and $TwoStepPeerFlow^2$ are excluded instruments, and the rest of equation (13) are included instruments.

and "contextual effects."[17]

"Correlated effects" simultaneously affect both connected managers due to common, time-invariate characteristics. Correlated effects can be conceived as a cointegration relationship where a relatively fixed relation among two neighbors induces a proportional response to exogenous events. For example, two mutual funds, one half the size of the other, may find that on average the smaller fund receives half the capital flows of the large one. Since there may be a similar relationship due to cash holdings, I include *PeerSize* and *PeerCash* to control for these potentially common fund characteristics which may drive correlated flows.

I control for Manski's "contextual" effects by including *CategoryAvgFlows*, which represents the average flow for the Morningstar category to which each open ended fund belongs. Contextual effects can be conceived as a network version of industry effects, in which market-wide trends affect all members of the group equally, but may change across time. For instance, a sector rotation strategy which suggests buying utilities and health care stocks in a declining market represents a wider shift in investor behavior, operating above the level of individual portfolio managers.

A further identification problem may arise due to network density, as noted by Kelejian, Prucha, and Yuzefovich (2006). If I have a very dense or "complete" network such that everyone is equally connected, each network member would have exactly the same *PeerFlow* measure. For example, assume that each portfolio manager is connected to each other manager with a weight of exactly 1. This would make *PeerFlow* equal to the average market-wide flow since the weight on each flow variable would be $\frac{1}{N}$ for every manager and therefore no longer display cross-sectional variation. Given that my weighted density is less than 5%, this is

[17] Bramoullé, Djebbari, and Fortin (2009) also note that these controls are a necessary prerequisite for their instrumentation approach.

unlikely to be a problem. As a further robustness check, I have run my results thresholding my network at the $80^{\text{th}}$ percentile, thus obtaining an unweighted density of 10% with no material change in results.[18]

Finally, I estimate this set of equations by Generalized Method of Moments, whereas most specifications of this type in the spatial econometrics literature estimate this model via Maximum Likelihood. Conley (1999) notes that ML specifications in which spatial dependence is measured with error are misspecified. While this 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, such as I have here, both OLS and GMM estimators are consistent, and thus represents the appropriate estimation approach. Elhorst (2010) includes a short discussion on ML 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.

## IV. Results

The baseline fund flow specification is from Sirri and Tufano (1998) and Coval and Stafford (2007). They regress fund flow on lagged flows, lagged returns, fund size, and fund category average flows at time $t$, with fund flow defined as dollar flows normalized by total net assets, the same normalization I apply throughout. When I run this specification in a pooled OLS and Fama-MacBeth framework, I get results qualitatively similar to Coval and Stafford (2007) and others who have investigated this relationship such as Lou (2010) and Ferreira, Keswani, Miguel,

[18] Weighted density is the sum of all network connections in the network divided by the sum of all possible network connections set to 1, $N^2$. Unweighted density is the same, but sets any weighted network link to 1 first.

and Ramos (2011). However, I find it necessary to include both time and firm fixed effects and further cluster my standard errors in both time and portfolio manager dimensions.

When I run both the Breusch-Pagan test and an F test on RSS of regressions with and without time and firm fixed effects, I find 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 (2008)). Clustering standard errors in both time and manager dimensions produces large changes in standard errors indicating that this is a necessary step (Petersen (2009)). I maintain this specification design throughout. Results from these tests as well as a table comparing the varying differences in specification are available upon request.[19] Including time fixed effects also controls for market wide events affecting all funds, and fund fixed effects control for fund or fund manager time-invariant attributes.

*A. Regression Results*

Our econometrics established, I turn to Table II which contains the results from the first stage of the instrumental variables regression. The $R^2$ of the *Peer Flow* regression is 0.83, indicating the excellent fit necessary in a first stage regression.

Next, I begin by regressing $Return$ on my networked and instrumented $PeerFlow$ variable as evidence that portfolio overlaps are driving a contagion effect, rather than a correlated flow process. As shown in Table III, there is a positive and significant coefficient on *PeerFlow* which simultaneously increases the $R^2$ from 0.14 to 0.17 and reduces the magnitude of both *Market Return* from 0.90 to 0.71 and *Category Avg Flow* from 0.39 to 0.23, with all changes statistically significant. That the fund flows from neighboring portfolio managers positively predict returns is

[19] 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.

solid evidence that portfolio interconnections are the channel for this influence.

My main specification is in Table IV. Here, $Flow$ is the dependent variable with $PeerFlow$ as independent variable alongside other control variables. Again, $PeerFlow$ enters in positively and significantly with slight decreases in other predictor variables, indicating a flow contagion process. However, since $Flow$ enters into the specification both as dependent and independent variable, I must transform the equation similar to an autoregression specification to fully interpret this coefficient.

*B. Network Coefficient Interpretation*

To interpret the coefficient on Model 2 in Table IV, I begin by rewriting my specification in Equation 13 in matrix form, without the instrumentation:[20]

$$F_t = \rho_s W_t F_t + \rho_t F_{t-1} + X_t \beta + \epsilon \tag{14}$$

in which $F_t$ is the $N \times 1$ vector of fund flows at time $t$. $W_t$ is a row-stochastic transformation of $N \times N$ portfolio similarity matrix $S$ at time $t$, such that $PeerFlow_t = W_t \cdot F_t$. $X_t$ represents all other control and explanatory variables for simplicity.

Next, I group together all terms involving $F_t$, also setting $F_t = F_{t-1}$ to incorporate a steady-state process.[21] Since flows are not persistent, this is a trivial

[20] For this analysis, I simply use the endogenous *PeerFlow* rather than the predicted value from the first stage regression, which simplifies the exposition and likely is a good approximation since the $R^2$ of the first stage regression is 0.83. However, I still use the coefficient estimates from the instrumented specification.

[21] Note that in my specification, I have 4 Flow lags, so $F_t = F_{t-p}$ for $p = 1, 2, 3, 4$ and $\rho_t$ is the sum of the 4 coefficients. I do the same for the coefficient on Return lags.

simplification. The result is

$$((1-\rho_t) I_N - \rho_s W_t) F_t = X_t\beta + \epsilon \tag{15}$$

$$F_t = X_{1t}\tilde{\beta}_1 + X_{2t}\tilde{\beta}_2 + \ldots + X_{Pt}\tilde{\beta}_P + \epsilon \tag{16}$$

for each $p = 1 \ldots P$ explanatory variables. Each actual estimated coefficient is

$$\tilde{\beta}_{p,N\times N} = ((1-\rho_t) I_N - \rho_s W_t)^{-1} \beta_p \tag{17}$$

which is an $N \times N$ matrix. Without my network specification, the comparable coefficient would be the scalar coefficient estimate times an $N \times N$ identity matrix.

In equation 17, $\beta_{Return}$ is the sum of the return coefficients from Model 2 in Table IV since in steady-state, $t = t - p \; \forall p$. Since $PeerCash = W \cdot Cash$, $\beta_{Cash}$ is the sum of the coefficient on *Cash* times the identity matrix plus the coefficient on *PeerCash* times $W$. Mathematically, if $\beta_C$ is the regression coefficient on *Cash* and $\beta_{PC}$ is the regression coefficient on *PeerCash*, then the overall effect of cash, $\beta_{Cash}$ is

$$\beta_{Cash} = \beta_C \cdot I + \beta_{PC} \cdot W \tag{18}$$

$\beta_{CategoryAvgFlow}$ is simply the corresponding estimated coefficient from Model 2 in Table IV.

To interpret this network coefficient, I divide it into feedback effects, represented by the diagonal, and spillover effects which reside on the off-diagonal. The results for important explanatory variables are in Table V. The first column is the scalar coefficient estimate, $\beta$, without the network transformation. Next are the incremental feedback effects, computed as the average of the diagonal less the scalar coefficient. Finally, spillover effects are computed as the average of all off-diagonal entries in the network coefficient.

Table V shows how my network specification accounts for feedback and spillover effects, increasing estimate by up to 76%. Specifically, returns and category average flows show effects that are 52% greater than non-networked effects, and networked cash holdings effects are 76% greater.

To illustrate spillover, I simulate a shock to approximately 40% of the fund managers in the sample and measure the impact to the other 60%, which is assumed to be zero in a non-networked specification. I shock *Cash* by one standard deviation, simulating an unexpected redemption, and I shock *Returns* by one standard deviation, simulating an unexpected market movement.[22] The results are illustrated in Figure 2 and Figure 3. Note that these spillover effects are as large as 0.01, which is the mean value of flow and approximately 10% of the standard deviation, available in Table I.

To more fully identify capital flow contagion as a unique phenomena, I perform several robustness checks. I re-run my main specification removing all sector funds from the dataset, and find the result strengthened – the coefficient is larger and estimated with more precision.[23,24]

Results with and without sector funds are presented in Tables VI and VII. In Table VI, the contagion process in Model 2 without sector funds is almost 25% greater than the baseline including them (0.50 compared to 0.41) whereas sector funds alone show no significance. *Fund Category Avg* is also smaller without sector funds, at 0.62 vs 0.73 in the baseline result. Among sector funds only, this same

[22]Since managers are connected by assets, for this to be an isolated shock, it could be to non-equity holdings or other non-connected holdings.

[23]Sector funds are those labeled Technology, Utilities, Financials, etc. corresponding to equities held in a specific industry.

[24]Note that this is a simple division of my sample which only considers the portfolio managers who are impacted by peer flows, not a full network subset. Subsetting a network specification is non-trivial in general since there are many connections among and between any chosen grouping of portfolio managers such that any subset arbitrarily cuts some of those ties and keeps others.

control is 0.82, indicating that *Fund Category Avg* is a primary driver of sector fund flows. In Table VII, there is very little difference between the models with and without sector funds, displayed in Models 1 and 2. *Fund Category Avg* drops from 0.23 to 0.13, indicating that while *PeerFlow* and *Fund Category Avg* overlap somewhat among sector funds, they are much less related in the broader sample.

Since financial crises induce correlations across disparate asset groups, it is possible that my result is simply arising from the recent financial crisis. Accrodingly, I re-run my specification omitting the financial crisis, stopping my analysis in the second quarter of 2007 and 2008, respectively, with results presented in Tables VIII and IX. Interestingly, the flow contagion effect is stronger when omitting the financial crisis. This can be seen in Model 2 of Table VIII, in which the *PeerFlow* coefficient rises moderately (though without statistical significance) from 0.41 to 0.44. Table IX presents the results for returns, again showing no marked difference.

## V. Crowded Trades and Network Persistence

Having provided evidence that portfolio interconnections may induce capital flow contagion, I proceed to investigate the nature of these connections. If these portfolio connections are relatively persistent, then this static set of connections may be more easily identified from public holdings disclosures by both market participants and regulators alike. On the other hand more transient portfolio interconnections may make capital flow contagion effects much harder to detect ex ante.

Between the two, transient or hard-to-observe portfolio interconnections pose the greater risk to portfolio managers and regulators alike since a hidden contagion process is more likely to generate unexpected negative shocks. These transient portfolio interconnections may arise due to so-called “crowded trades”, which occur when portfolio managers take concentrated or overweighed positions in a small set of stocks.[25] Due to lags in mandatory disclosures, crowded trades may not be detectable to market participants until many months after the trades are established. Thus, with no knowledge of network connections, negative flow shocks across portfolio connections will be unanticipated and likely produce greater negative consequences than shocks which are at least partially anticipated.

Table X presents the results of an autoregression on my network measure, similar to the main specification in Anton and Polk (2010). This specification takes the $N \times N$ network of relationships between all of the portfolio managers at time $t$ and puts them in a $N \times 1$ vector as the dependent variable. Then the same network of relationships at $t-1$, $t-2$, enter as independent variables, vectorized. I then run this regression for each time $t$ and summarize the coefficients across time in a Fama-MacBeth framework.

The marginal effects of the lags diminish to be statistically insignificant after 3 lags, but still show some autoregressive properties. The network distance correlation lagged one quarter is 0.41, which indicates some short-term persistence. To estimate the correlation two quarters previous, I compute $0.41^2 + 0.25 = 0.42$, showing that the persistence extends to the previous six months. But the correlation between the network distance measure and that 3 quarters past is $0.41^3 + 0.25^2 + 0.765 = 0.21$, a significant drop off, and then one year past is $0.41^4 + 0.25^3 + 0.08^2 + 0.0 = 0.05$ if I treat the insignificant $4^{\text{th}}$ lag as 0, or 0.13 if I retain it. After two years, retaining the first four coefficients, the correlation is $0.41^8 + 0.25^7 + 0.08^6 + 0.08^5 = 0.0009$, which is very close to 0.[26] Since portfolio

[25] Crowded trades are also related to the herding literature. Sias (2004) summarizes the broad classifications motivating herding. Rationally, managers herd on correlated private information (Froot, Scharfstein, and Stein (1992)). Other explanations include reputation-based herding (Scharfstein and Stein (1990)), and fads (Barberis and Shleifer (2003)).

[26] This analysis of time-series coefficients comes from Hamilton (1994), Chapter 1.

objectives likely persist greater than two years, this suggests that there is some transience to my measure of interconnectedness and thus that crowded trades or herding among institutional managers plays a role in capital flow contagion.

Finally, to further investigate the nature of portfolio connectivity, I show that my normalized dot product distance measure increases in two dimensions. First, it is increasing in portfolio overlap, which is its primary purpose. As the percentage of portfolio overlap increases, the distance between two managers in security space decreases (they are more similar in security space). But, perhaps less intuitively, my portfolio distance measure is also increasing in the concentration of those holdings. This is illustrated in Figure 4. Holding total portfolio overlap constant, a single concentrated position gives twice as much similarity as two overlapping holdings of equal proportion. This property of my portfolio distance measure indicates that concentrated positions give rise to more interconnectedness. Accordingly, crowding or overweighting in a specific set of securities may induce more connectedness among those managers than they may realize.

## VI. Conclusion

In the wake of the recent financial crisis, the interconnection of market participants has become an important new area of research. Employing a novel, network-based specification, I show that interconnected intermediaries exhibit contagious capital flows, exposing them to feedback effects and spillover effects which result in estimates 50-75% greater than non-networked coefficients. To incorporate these network connections simultaneously, I contemporaneously estimate the influence of each portfolio manager's capital flows on each other manager by exploiting the network structure as an instrument.

I also have shown some evidence that that these contagious flows are the result of crowded trades – short-term, popular market positions – since portfolio connection exhibits only a small amount of short-term persistence. Furthermore, I have illustrated how distances between portfolios in security space emphasize concentrated positions, such that active managers overweighting portions of their portfolio may unintentionally increase their dependence on similar neighbors.

While my analysis focuses on the equity holdings of open ended funds, it also has implications for collateralized financing. Financial intermediaries who rely on collateralized (wholesale) financing to fund their investments are growing in market share (Adrian and Shin (2010)). It may be that my results imply a broader "collateral contagion" effect which could have played a role in recent runs on repo financing (Gorton and Metrick (2011)). Since even interbank lending is becoming more collateralized, Allen and Gale (2000)'s canonical model of interbank financial contagion may be further amplified by connected collateral.[27]

This work also provides motivation for the collection of more detailed holdings data from market participants, since the results described herein can be characterized as a negative network externality which may merit government regulation. Indeed, Brunnermeier, Hansen, Kashyap, Krishnamurthy, and Lo (2011) recently responded to an AEA/NSF call for proposals on "grand challenge questions" for research in the next ten years by advocating the collection of additional data and developing network models in the pursuit of quantifying systemic financial risk. While immediate public disclosure may have unintended predatory trading effects (Brunnermeier and Pedersen (2005)), confidential disclosure to regulatory bodies and/or delayed public disclosure are likely to be beneficial and could be

[27]In November 2009, the ECB (Heider and Hoerova (2009)) reported that interest rates for collateralized lending in the interbank market since 2007 were significantly lower than unsecured rates, a historical divergence, and a more recent report from the Financial Times indicates that interbank unsecured lending has essentially disappeared. (http://ftalphaville.ft.com/blog/2010/08/16/315556/euribor-has-been-vaporised/)

the purview of the newly formed Office of Financial Research established by the Dodd-Frank Act.

While network methods are becoming more popular in corporate finance (e.g., Hochberg, Ljungqvist, and Lu (2007), Cohen, Frazzini, and Malloy (2008), Ahern and Harford (2010)) and market microstructure (Cohen-Cole, Kirilenko, and Patacchini (2010)), little has been done applying network methods to equity markets. My network approach allows a steady-state analysis of this peer influence process in the cross-section, 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 (e.g. Cohen-Cole, Patacchini, and Zenou (2011)) or stock market volatility (Greenwood and Thesmar (2011)). And 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 Brunnermeier, 2002, Synchronization risk and delayed arbitrage, *Journal of Financial Economics* 66, 341–360.

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

Adrian, Tobias, and Hyun Song Shin, 2010, The changing nature of financial intermediation and the financial crisis of 2007–2009, *Annual Review of Economics* 2, 603–618.

Ahern, Kenneth, and Jarrad Harford, 2010, The importance of industry links in merger waves, *Working Paper.*

Allen, Franklin, Ana Babus, and Elena Carletti, 2010, Financial connections and systemic risk, *NBER Working Paper.*

Allen, Franklin, and Douglas Gale, 2000, Financial contagion, *Journal of Political Economy* 108, 1–33.

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

Anton, Miguel, and Christopher Polk, 2010, Connected stocks, *Working Paper.*

Babus, Ana, 2010, Strategic relationships in over-the-counter markets, *Working Paper.*

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

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

Brunnermeier, Markus, Lars Peter Hansen, Anil Kashyap, Arvind Krishnamurthy, and Andrew W. Lo, 2011, Modeling and measuring systemic risk, *SSRN Working Paper.*

Brunnermeier, Markus, and Lasse Pedersen, 2005, Predatory trading, *The Journal of Finance* 60, 1825–1863.

——— , 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201.

Brunnermeier, Markus, and Yuliy Sannikov, 2011, A macroeconomic model with a financial sector, *Working Paper.*

Carhart, Mark, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

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

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

Cohen-Cole, Ethan, Andrei Kirilenko, and Eleonora Patacchini, 2010, Are networks priced? network topology and systemic risk in high liquidity markets, *Working Paper*.

Cohen-Cole, Ethan, Eleonora Patacchini, and Y 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, and E Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Elhorst, J Paul, 2010, Spatial panel data models, *Handbook of Applied Spatial Analysis* pp. 377–407.

Elton, Edwin, Martin J Gruber, CR Blake, Y Krasny, and SO Ozelge, 2009, The effect of holdings data frequency on conclusions about mutual fund behavior, *Journal of Banking and Finance*.

Ferreira, Miguel, Aneel Keswani, Antonio F Miguel, and Sofia B Ramos, 2011, The flow-performance relationship around the world, *Working Paper*.

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

Froot, KA, DS Scharfstein, and Jeremy Stein, 1992, Herd on the street: Informational inefficiencies in a market with short-term speculation, *Journal of Finance* pp. 1461–1484.

Gorton, Gary, and Andrew Metrick, 2011, Securitized banking and the run on repo, *Journal of Financial Economics*.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Working Paper*.

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

Heider, Florian, and Marie Hoerova, 2009, Interbank lending, credit risk premia and collateral, Discussion paper, .

Hochberg, Yael, Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *The Journal of Finance* 62, 251–301.

Jotikasthira, C, Christian Lundblad, and T Ramadorai, 2011, Asset fire sales and purchases and the international transmission of financial shocks, *Working Paper.*

Kelejian, Harry H., and Ingmar R. Prucha, 2002, 2sls and ols in a spatial autoregressive model with equal spatial weights, *Regional Science and Urban Economics* 32, 691–707.

———, and Yevgeny Yuzefovich, 2006, Estimation problems in models with spatial weighting matrices which have blocks of equal elements, *Journal of Regional Science* 46, 507–515.

Kennedy, Peter, 2008, A guide to econometrics, *Blackwell Publishing* p. 623.

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

———, and Wei Xiong, 2001, Contagion as a wealth effect, *Journal of Finance* 56, 1401–1440.

Lou, Dong, 2010, A flow-based explanation for return predictability, *Working Paper.*

Manski, CF, 1993, Identification of endogenous social effects: The reflection problem, *The Review of Economic Studies.*

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

Scharfstein, DS, and Jeremy Stein, 1990, Herd behavior and investment, *American Economic Review* pp. 465–479.

Shleifer, Andrei, and Robert Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

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

Sirri, Erik, and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

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

Zawadowski, Adam, 2011, Entangled financial systems, *Working Paper*.

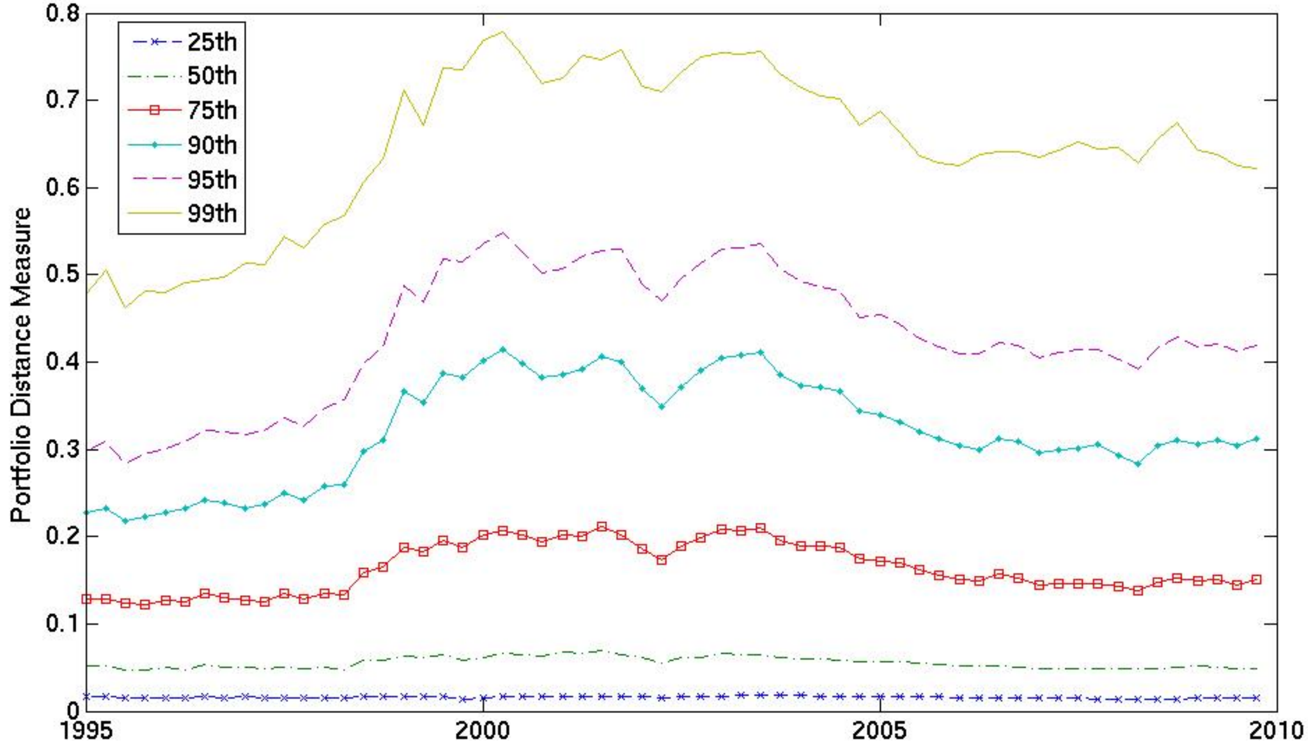


**Figure 1: Percentiles of portfolio distance measure through time.** Each line represents the time series of a specific percentile of the cross-sectional distribution of the normalized dot product portfolio distance measure. The top is the $99^{\text{th}}$ percentile, then the $95^{\text{th}}$, $90^{\text{th}}$, $75^{\text{th}}$, $50^{\text{th}}$, and $25^{\text{th}}$.

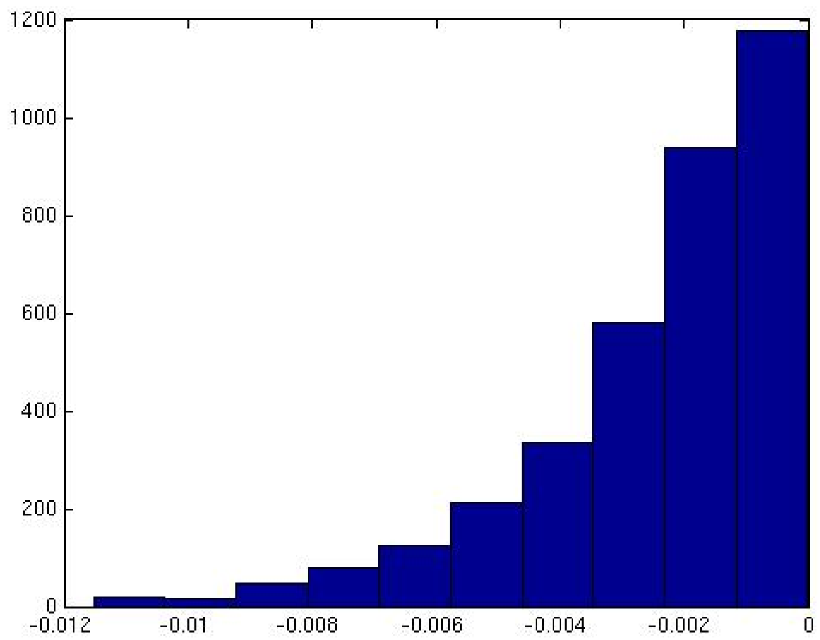


**Figure 2: The effect of a shock to cash holdings within a subset of portfolio managers.** This histogram illustrates the effect of one standard deviation negative shock to cash holdings divided by total net assets, which simulates an unexpected redemption by investors. Shock is applied to approximately 40% of portfolio managers, defined as the most connected managers (top tercile) in any time period. Plotted here is the impact to the 60% of managers *not shocked* and thus is illustrative of spillover effects.

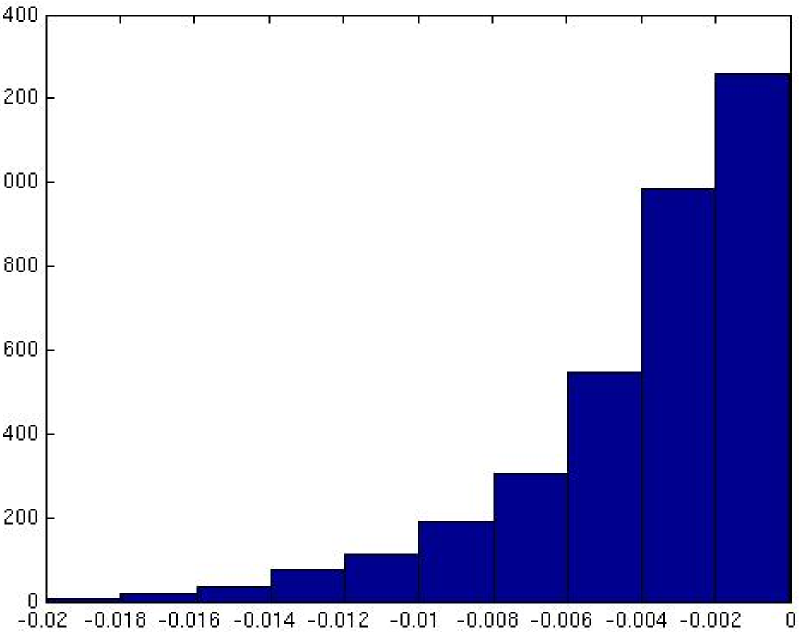


**Figure 3: The effect of a shock to portfolio returns within a subset of portfolio managers.** This histogram illustrates the effect of one standard deviation negative shock to portfolio returns, which simulates an unexpected market movement or impact to non-equity (i.e. non-connected) portion of portfolio. Shock is applied to approximately 40% of portfolio managers, defined as the most connected managers (top tercile) in any time period. Plotted here is the impact to the 60% of managers *not shocked* and thus is illustrative of spillover effects.

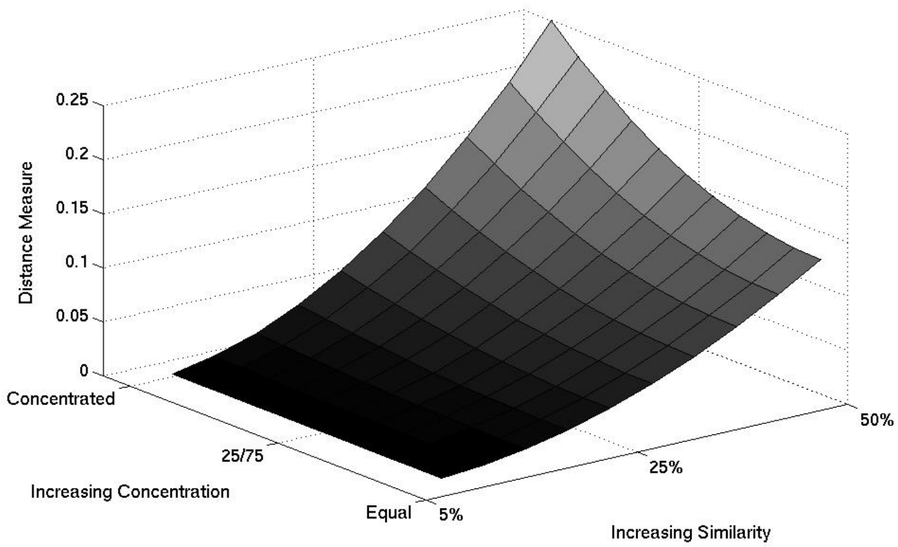


**Figure 4: Two dimensions of the portfolio distance measure.** This demonstrates how interconnectedness as measured by the normalized dot product between two portfolios increases in two different ways. The axis on the right is increasing in portfolio overlap, i.e. the percentage of the portfolio that overlaps. The axis on the left is increasing in the concentration of that position, holding the percentage of overlap constant. More concentrated positions are thus closer in security space, holding overlap constant.

**Table I: Fund Summary Statistics**

Summary statistics for fund data as used in regression specifications. Flow is dollar flows divided by total net assets and Cash is cash holdings divided by total net assets. Size is log of total net assets. Amihud is a portfolio weighted measure of the Amihud values of equity holdings, logged. Fund Category Average is the average Flow across Morningstar categories. Peer variables are weighted by network connections. Data is quarterly from 1995 to 2009.

| Variable Names | N | Mean | Std Dev | Min | Max |
|---|---|---|---|---|---|
| Flow | 147,753 | 0.010 | 0.168 | -1.000 | 0.735 |
| Return | 147,608 | 0.014 | 0.115 | -0.990 | 4.871 |
| Size | 147,753 | 18.930 | 1.992 | 0.693 | 25.988 |
| Cash | 147,753 | 0.047 | 0.071 | -0.003 | 0.535 |
| Amihud | 138,259 | -13.759 | 2.127 | -30.567 | -5.515 |
| Fund Category Avg | 147,753 | 0.010 | 0.045 | -1.000 | 0.735 |
| Peer Flow | 147,745 | 0.007 | 0.027 | -1.000 | 0.735 |
| Peer Return | 147,749 | 0.012 | 0.097 | -0.325 | 0.488 |
| Peer Size | 147,753 | 20.869 | 0.526 | 10.455 | 25.467 |
| Peer Cash | 147,753 | 0.043 | 0.016 | -0.003 | 0.535 |

**Table II: First Stage GMM Regression**

First stage regressions with endogenous regressors as dependent variables. Peer Flow is the weighted average of peer connected flow, and Two Step Peer Flow is the same of their neighbor's neighbors, used as instruments. Flow is dollar flows divided by total net assets and Cash is cash holdings divided by total net assets. Size is log of total net assets. Amihud is a portfolio weighted measure of the Amihud values of equity holdings, logged. Fund Category Average is the average Flow across Morningstar categories. Data is quarterly from 1995 to 2009. Time and Fund Fixed Effects included. T statistics are in parentheses and significance is denoted at the 1, 5, and 10% level.

| | (1)<br>Peer Flow |
|---|---|
| Two Step Peer Flow | 1.3973*** |
| | (58.28) |
| Two Step Peer Flow$^2$ | -1.9472*** |
| | (-4.08) |
| Lag1 Flow | 0.0006** |
| | (2.21) |
| Lag2 Flow | 0.0006** |
| | (2.32) |
| Lag3 Flow | 0.0001 |
| | (0.55) |
| Lag4 Flow | 0.0005** |
| | (2.51) |
| Lag1 Return | 0.0129*** |
| | (4.27) |
| Lag2 Return | 0.0082*** |
| | (3.86) |
| Lag3 Return | 0.0054** |
| | (2.52) |
| Lag4 Return | 0.0038** |
| | (2.06) |
| Fund Size | 0.0002* |
| | (1.84) |
| Cash Pct | 0.0017* |
| | (1.87) |
| Amihud Illiq | 0.0001 |
| | (0.63) |
| Fund Category Avg | 0.0388*** |

Continued on next page...

| | (1) Peer Flow |
|---|---|
| | (9.77) |
| Pr Fund Size | 0.0009 |
| | (1.00) |
| Pr Cash Pct | 0.2930*** |
| | (6.73) |
| Observations | 84882 |
| R Squared | 0.83 |
| Fund clusters | 5,158 |
| Time clusters | 44 |

**Table III: Effect of Peer Flows on Portfolio Returns**

Portfolio return is the dependent variable, provided by Morningstar. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Flow is the fund flow divided by total net assets. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Market return is CRSP value weighted market return, and Category Avg Flow is the average of all reported fund flows by Morningstar category. Flow and return lags 3 and 4 included but not shown. Time and Fund Fixed Effects included. 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.

| | (1)<br>Port Ret | (2)<br>Port Ret |
|---|---|---|
| Peer Flow | | 1.2641*** |
| | | (6.25) |
| Market Return | 0.9038*** | 0.7143*** |
| | (4.91) | (3.86) |
| Fund Category Avg | 0.3896*** | 0.2256*** |
| | (6.44) | (5.98) |
| Pr Fund Size | | 0.0046 |
| | | (0.52) |
| Pr Cash Pct | | -0.6319 |
| | | (-1.30) |
| Lag1 Flow | -0.0030 | -0.0026 |
| | (-0.80) | (-0.71) |
| Lag2 Flow | -0.0000 | 0.0010 |
| | (-0.01) | (0.32) |
| Lag1 Return | -0.0000 | -0.0614 |
| | (-0.00) | (-0.91) |
| Lag2 Return | -0.0517 | -0.0868 |
| | (-0.84) | (-1.47) |
| Fund Size | -0.0016 | -0.0039** |
| | (-0.88) | (-2.48) |
| Cash Pct | 0.0249** | 0.0155 |
| | (1.99) | (1.64) |
| Amihud Illiq | 0.0003 | 0.0016** |

Continued on next page...

| | (1) Port Ret | (2) Port Ret |
|---|---|---|
| | (0.44) | (2.48) |
| Observations | 84804 | 84804 |
| R Squared | 0.14 | 0.17 |
| Fund clusters | 5,152 | 5,152 |
| Time clusters | 44 | 44 |
| Est Method | OLS | GMM |
| Hansen J stat | . | 1.71 |
| J p value | | 0.1906 |
| KP LM Stat | | 32.11 |
| KP LM p value | | 0.0000 |

**Table IV: Effect of Peer Flow on Fund Flows**

Flow ratio is the dependent variable and is the fund flow divided by total net assets. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Category Avg Flow is the average of all reported fund flows by Morningstar category. Flow and return lags 3 and 4 included but not shown. Time and Fund Fixed Effects included. 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.

| | (1) Flow | (2) Flow |
|---|---|---|
| Peer Flow | | 0.4052*** |
| | | (3.21) |
| Fund Category Avg | 0.7840*** | 0.7330*** |
| | (11.31) | (8.91) |
| Pr Fund Size | | -0.0258*** |
| | | (-3.79) |
| Pr Cash Pct | | 0.1064 |
| | | (0.45) |
| Lag1 Flow | 0.0504*** | 0.0494*** |
| | (3.01) | (2.96) |
| Lag2 Flow | 0.0828*** | 0.0862*** |
| | (5.77) | (6.09) |
| Lag1 Return | 0.1615*** | 0.1471*** |
| | (8.85) | (8.78) |
| Lag2 Return | 0.0874*** | 0.0772*** |
| | (3.67) | (3.58) |
| Fund Size | 0.0170*** | 0.0162*** |
| | (6.59) | (6.40) |
| Cash Pct | 0.3046*** | 0.2985*** |
| | (16.94) | (16.81) |
| Amihud Illiq | 0.0027*** | 0.0024*** |
| | (4.37) | (3.76) |
| Observations | 84757 | 84757 |
| R Squared | 0.09 | 0.09 |

Continued on next page...

| | (1) Flow | (2) Flow |
|---|---|---|
| Fund clusters | 5,148 | 5,148 |
| Time clusters | 44 | 44 |
| Est Method | OLS | GMM |
| Hansen J stat | 0.00 | 1.98 |
| J p value | | 0.1596 |
| KP LM Stat | | 32.11 |
| KP LM p value | | 0.0000 |

**Table V:**

**Contagion Effect of Peer Flows on Fund Flows - Long Run Steady State**

Contagion effect based on Model 2 in Table IV, assuming long run and cross-sectional equilibrium (through time and across funds). Coeff Estimate is non-networked estimate, Feedback Effect includes the incremental average spillover effects which circulate back to the same fund, Spillover effect is the average off-diagonal effects among portfolio managers. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product. Category Avg Flow is the average of all reported fund flows by Morningstar category.

| | Coeff Estimate | Feedback Effect | Spillover Effect | Percent Underestimated |
|---|---|---|---|---|
| Return | 0.3194 | 0.0640 | 0.1024 | 52% |
| Cash | 0.2985 | 0.0598 | 0.1659 | 76% |
| Category Mean | 0.7330 | 0.1468 | 0.2350 | 52% |

**Table VI: Results removing sector funds – fund flows**

Fund flow divided by total net assets is the dependent variable, provided by Morningstar. Model 1 is the baseline, taken from Model 2 of Table IV. Model 2 is the same, but with sector funds omitted from the analysis. Model 3 includes only sector funds. Sector funds are mutual funds with an industry-specific category, such as Technology or Health Care. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Market return is CRSP value weighted market return, and Category Avg Flow is the average of all reported fund flows by Morningstar category. Time and Fund Fixed Effects included. 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.

| | (1)<br>Flow | (2)<br>Flow | (3)<br>Flow |
|---|---|---|---|
| Peer Flow | 0.4052*** | 0.5028*** | 0.3706 |
| | (3.21) | (3.61) | (1.38) |
| Lag1 Flow | 0.0494*** | 0.0868*** | -0.1951*** |
| | (2.96) | (5.14) | (-5.11) |
| Lag2 Flow | 0.0862*** | 0.0859*** | -0.0040 |
| | (6.09) | (5.97) | (-0.14) |
| Lag3 Flow | 0.0192** | 0.0193* | -0.0268 |
| | (2.00) | (1.95) | (-1.20) |
| Lag4 Flow | 0.0119 | 0.0197** | -0.0413 |
| | (1.19) | (2.12) | (-1.41) |
| Lag1 Return | 0.1471*** | 0.1920*** | 0.0885*** |
| | (8.78) | (9.60) | (3.51) |
| Lag2 Return | 0.0772*** | 0.0993*** | 0.0240 |
| | (3.58) | (3.81) | (1.35) |
| Lag3 Return | 0.0507*** | 0.0724*** | -0.0003 |
| | (3.10) | (3.92) | (-0.01) |
| Lag4 Return | 0.0444** | 0.0632*** | -0.0019 |
| | (2.42) | (2.86) | (-0.12) |
| Fund Size | 0.0162*** | 0.0137*** | 0.0514*** |
| | (6.40) | (5.90) | (4.82) |
| Cash Pct | 0.2985*** | 0.2846*** | 0.3845*** |

Continued on next page...

| | (1) Flow | (2) Flow | (3) Flow |
|---|---|---|---|
| | (16.81) | (15.06) | (6.14) |
| Amihud Illiq | 0.0024*** | 0.0023*** | 0.0031 |
| | (3.76) | (3.67) | (1.25) |
| Fund Category Avg | 0.7330*** | 0.6238*** | 0.8167*** |
| | (8.91) | (6.44) | (6.17) |
| Pr Fund Size | -0.0258*** | -0.0256*** | -0.0608*** |
| | (-3.79) | (-3.66) | (-2.74) |
| Pr Cash Pct | 0.1064 | 0.0981 | -0.1813 |
| | (0.45) | (0.40) | (-0.38) |
| Observations | 84757 | 76698 | 8059 |
| R Squared | 0.09 | 0.09 | 0.18 |
| Fund clusters | 5,148 | 4,704 | 444 |
| Time clusters | 44 | 44 | 44 |
| Est Method | | | |
| Hansen J stat | 1.98 | 2.30 | 1.79 |
| J p value | 0.1596 | 0.1292 | 0.1808 |
| KP LM Stat | 32.11 | 29.86 | 23.80 |

**Table VII: Results removing sector funds – portfolio returns**

Portfolio return is the dependent variable, provided by Morningstar. Model 1 is the baseline, taken from Model 2 of Table III. Model 2 is the same, but with sector funds omitted from the analysis. Model 3 includes only sector funds. Sector funds are mutual funds with an industry-specific category, such as Technology or Health Care. Data is quarterly from 1998 to 2009, each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Market return is CRSP value weighted market return, and Category Avg Flow is the average of all reported fund flows by Morningstar category. Time and Fund Fixed Effects included. 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.

| | (1)<br>Port Ret | (2)<br>Port Ret | (3)<br>Port Ret |
|---|---|---|---|
| Peer Flow | 1.2641*** | 1.2407*** | 1.9446*** |
| | (6.25) | (6.15) | (4.66) |
| Lag1 Flow | -0.0026 | -0.0028 | 0.0043 |
| | (-0.71) | (-0.95) | (0.39) |
| Lag2 Flow | 0.0010 | 0.0010 | 0.0063 |
| | (0.32) | (0.37) | (0.63) |
| Lag1 Return | -0.0614 | -0.0576 | -0.1003 |
| | (-0.91) | (-0.89) | (-1.17) |
| Lag2 Return | -0.0868 | -0.0685 | -0.1630** |
| | (-1.47) | (-1.26) | (-2.05) |
| Market Return | 0.7143*** | 0.7567*** | 0.6405** |
| | (3.86) | (4.35) | (2.23) |
| Fund Category Avg | 0.2256*** | 0.1327*** | 0.2539*** |
| | (5.98) | (3.33) | (4.01) |
| Amihud Illiq | 0.0016** | 0.0014** | 0.0019 |
| | (2.48) | (2.17) | (0.94) |
| Cash Pct | 0.0155 | 0.0147* | 0.0359 |
| | (1.64) | (1.66) | (1.01) |
| Fund Size | -0.0039** | -0.0045*** | 0.0011 |
| | (-2.48) | (-3.24) | (0.23) |
| Pr Fund Size | 0.0046 | 0.0028 | 0.0395 |

Continued on next page...

| | (1) Port Ret | (2) Port Ret | (3) Port Ret |
|---|---|---|---|
| | (0.52) | (0.35) | (1.12) |
| Pr Cash Pct | -0.6319 | -0.3495 | -2.7786** |
| | (-1.30) | (-0.80) | (-2.50) |
| Observations | 84804 | 76742 | 8062 |
| R Squared | 0.17 | 0.17 | 0.23 |
| Fund clusters | 5,152 | 4,708 | 444 |
| Time clusters | 44 | 44 | 44 |
| Est Method | | | |
| Hansen J stat | 1.71 | 2.04 | 0.78 |
| J p value | 0.1906 | 0.1533 | 0.3769 |
| KP LM Stat | 32.11 | 29.86 | 23.83 |
| KP LM p value | 0.0000 | 0.0000 | 0.0000 |

**Table VIII: Results removing the financial crisis – fund flows**

Fund flow divided by total net assets is the dependent variable, provided by Morningstar. Model 1 is the baseline, taken from Model 2 of Table IV, ranging from 1998 to 2009. Model 2 is the same, but only including quarters from 1998 through the second quarter of 2007. Model 3 extends through the second quarter of 2008. Each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Market return is CRSP value weighted market return, and Category Avg Flow is the average of all reported fund flows by Morningstar category. Time and Fund Fixed Effects included. 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.

| | (1)<br>Flow | (2)<br>Flow | (3)<br>Flow |
|---|---|---|---|
| Peer Flow | 0.4052*** | 0.4393*** | 0.4070*** |
| | (3.21) | (3.17) | (3.15) |
| Lag1 Flow | 0.0494*** | 0.0117 | 0.0358* |
| | (2.96) | (0.53) | (1.81) |
| Lag2 Flow | 0.0862*** | 0.0757*** | 0.0860*** |
| | (6.09) | (4.87) | (5.42) |
| Lag3 Flow | 0.0192** | 0.0060 | 0.0129 |
| | (2.00) | (0.45) | (1.24) |
| Lag4 Flow | 0.0119 | -0.0055 | 0.0003 |
| | (1.19) | (-0.42) | (0.02) |
| Lag1 Return | 0.1471*** | 0.1428*** | 0.1510*** |
| | (8.78) | (6.69) | (7.25) |
| Lag2 Return | 0.0772*** | 0.0868*** | 0.0962*** |
| | (3.58) | (2.84) | (3.23) |
| Lag3 Return | 0.0507*** | 0.0375** | 0.0397** |
| | (3.10) | (2.13) | (2.27) |
| Lag4 Return | 0.0444** | 0.0378* | 0.0399** |
| | (2.42) | (1.85) | (2.00) |
| Fund Size | 0.0162*** | 0.0238*** | 0.0188*** |
| | (6.40) | (4.81) | (5.35) |
| Cash Pct | 0.2985*** | 0.2958*** | 0.3046*** |
| | (16.81) | (13.64) | (15.44) |

Continued on next page...

| | (1) Flow | (2) Flow | (3) Flow |
|---|---|---|---|
| Amihud Illiq | 0.0024*** | 0.0027*** | 0.0022*** |
| | (3.76) | (3.49) | (3.09) |
| Fund Category Avg | 0.7330*** | 0.7165*** | 0.7267*** |
| | (8.91) | (8.46) | (8.89) |
| Pr Fund Size | -0.0258*** | -0.0317*** | -0.0307*** |
| | (-3.79) | (-3.84) | (-4.31) |
| Pr Cash Pct | 0.1064 | -0.0129 | 0.1884 |
| | (0.45) | (-0.06) | (0.77) |
| Observations | 84757 | 59376 | 70554 |
| R Squared | 0.09 | 0.09 | 0.09 |
| Fund clusters | 5,148 | 4,485 | 4,839 |
| Time clusters | 44 | 35 | 39 |
| Est Method | | | |
| Hansen J stat | 1.98 | 0.00 | 0.81 |
| J p value | 0.1596 | 0.9588 | 0.3685 |
| KP LM Stat | 32.11 | 26.73 | 29.62 |
| KP LM p value | 0.0000 | 0.0000 | 0.0000 |

**Table IX: Results removing the financial crisis – portfolio returns**

Portfolio return is the dependent variable, provided by Morningstar. Model 1 is the baseline, taken from Model 2 of Table III, ranging from 1998 to 2009. Model 2 is the same, but only including quarters from 1998 through the second quarter of 2007. Model 3 extends through the second quarter of 2008. Each panel variable is any open ended fund holding a nonzero equity position. Network relation is the normalized dot product, and peer effects are the weighted average of peer characteristics. Fund size is the log of total net assets. Cash Pct is cash holdings divided by total net assets. Amihud is the portfolio weighted sum of equity holdings' Amihud measures computed over the previous quarter. Market return is CRSP value weighted market return, and Category Avg Flow is the average of all reported fund flows by Morningstar category. Time and Fund Fixed Effects included. 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.

| | (1)<br>Port Ret | (2)<br>Port Ret | (3)<br>Port Ret |
|---|---|---|---|
| Peer Flow | 1.2641*** | 1.2432*** | 1.4065*** |
| | (6.25) | (6.90) | (6.67) |
| Lag1 Flow | -0.0026 | -0.0008 | -0.0029 |
| | (-0.71) | (-0.17) | (-0.72) |
| Lag2 Flow | 0.0010 | 0.0003 | -0.0003 |
| | (0.32) | (0.07) | (-0.08) |
| Lag1 Return | -0.0614 | -0.0655 | -0.1129 |
| | (-0.91) | (-0.90) | (-1.58) |
| Lag2 Return | -0.0868 | -0.0284 | -0.0039 |
| | (-1.47) | (-0.44) | (-0.06) |
| Market Return | 0.7143*** | 0.6498*** | 1.0145*** |
| | (3.86) | (3.75) | (9.09) |
| Fund Category Avg | 0.2256*** | 0.2180*** | 0.2416*** |
| | (5.98) | (5.79) | (6.51) |
| Amihud Illiq | 0.0016** | 0.0015** | 0.0020*** |
| | (2.48) | (1.97) | (2.68) |
| Cash Pct | 0.0155 | 0.0043 | 0.0154* |
| | (1.64) | (0.45) | (1.71) |
| Fund Size | -0.0039** | -0.0055*** | -0.0046** |
| | (-2.48) | (-2.83) | (-2.55) |
| Pr Fund Size | 0.0046 | 0.0035 | -0.0013 |
| | (0.52) | (0.31) | (-0.13) |

Continued on next page...

| | (1) Port Ret | (2) Port Ret | (3) Port Ret |
|---|---|---|---|
| Pr Cash Pct | -0.6319 | -0.4411 | -0.9395* |
| | (-1.30) | (-0.93) | (-1.82) |
| Observations | 84804 | 59420 | 70601 |
| R Squared | 0.17 | 0.14 | 0.24 |
| Fund clusters | 5,152 | 4,489 | 4,843 |
| Time clusters | 44 | 35 | 39 |
| Est Method | | | |
| Hansen J stat | 1.71 | 1.11 | 0.03 |
| J p value | 0.1906 | 0.2926 | 0.8643 |
| KP LM Stat | 32.11 | 26.72 | 29.62 |
| KP LM p value | 0.0000 | 0.0000 | 0.0000 |

**Table X: Persistence of Network Distance Relation**

Network relation is the normalized dot product, and is the dependent variable. Results shown from Fama-MacBeth regression of eight lags of network connectivity. Data is quarterly from 1998 to 2009 Significance is denoted at the 1, 5, and 10% level.

| | Coeff Estimate | Std Dev | T statistic |
|---|---|---|---|
| Lag 1 | 0.4138*** | 0.1255 | 3.2972 |
| Lag 2 | 0.2500*** | 0.0830 | 3.0112 |
| Lag 3 | 0.0765* | 0.0454 | 1.6835 |
| Lag 4 | 0.0851 | 0.0535 | 1.5918 |
| Lag 5 | 0.0189 | 0.0395 | 0.4785 |
| Lag 6 | 0.0407 | 0.0495 | 0.8222 |
| Lag 7 | 0.0110 | 0.0392 | 0.2802 |
| Lag 8 | 0.0460 | 0.0488 | 0.9418 |