Connected Stocks*

Miguel Antón
IESE Business School

Christopher Polk
London School of Economics

First Draft: May 2008
This draft: May 2013

---

* Antón: Department of Finance, IESE Business School, Av. Pearson 21, 08034 Barcelona, Spain. Email: manton@iese.edu. Polk: Department of Finance, London School of Economics, London WC2A 2AE, UK. Email: c.polk@lse.ac.uk. We are grateful to Ken French, David Hsieh, and Bob Shiller for providing us with some of the data used in this study. We are deeply indebted to two anonymous referees, an Associate Editor and the Editor, Cam Harvey, for numerous comments that significantly improved the paper. We are also grateful for comments from Andriy Bodnaruk, John Campbell, Randy Cohen, Jonathan Cohn, Owen Lamont, Augustin Landier, Dong Lou, Narayan Naik, Belén Nieto, Jeremy Stein, Dimitri Vayanos, Tuomo Vuolteenaho, and Paul Woolley, as well as from conference participants at the Summer 2008 LSE lunchtime workshop, Spring 2009 Harvard PhD brownbag lunch, 2010 HEC 2nd Annual Hedge Fund Conference, 2010 Paul Woolley Research Initiative Workshop, 2010 Spanish Finance Forum, 2010 WFA, 2010 EFA, 2010 Yale University Whitebox Graduate Student Conference, and 2010 Paul Woolley Conference, and seminar participants at Bristol University, IESE-ESADE, Imperial College, London School of Economics, and the University of Amsterdam. Financial support from the Paul Woolley Centre at the LSE is gratefully acknowledged. Antón also gratefully acknowledges support from the Fundación Ramón Areces.

Connected Stocks

Abstract

We connect stocks through the active mutual fund owners they have in common. We show that the degree of shared ownership forecasts cross-sectional variation in return correlation, controlling for exposure to systematic return factors, style and sector similarity, and many other pair characteristics. We argue that shared ownership causes this excess comovement based on evidence from a natural experiment—the 2003 mutual fund trading scandal. These results motivate a novel cross-stock reversal trading strategy exploiting information in ownership connections. We show that long/short hedge fund index returns covary negatively with this strategy, suggesting these funds may exacerbate this excess comovement.

JEL Classification: G12, G14.

Barberis and Shleifer (2003) and Barberis, Shleifer, and Wurgler (2005) have argued that institutional features may play an important role in the movement of stocks' discount rates, causing returns to comove above and beyond that implied by their fundamentals. In this paper, we propose a new approach to document that type of institutional-based comovement. Based on the mounting evidence since Coval and Stafford (2007) that mutual fund flows result in price pressure, we focus on connecting stocks through active mutual fund ownership. Specifically, we forecast cross-sectional variation in return correlation for stock pairs using the degree of common ownership by active mutual funds. Our bottom-up approach allows us to measure institutional-driven comovement more precisely. For example, we can isolate abnormally-connected firms by controlling for a host of pair characteristics.

We find that cross-sectional variation in common ownership predicts higher four-factor abnormal return correlation, controlling for similarity in industry, size, book-to-market ratio, and momentum characteristics as well as the degree of common analyst coverage. Our analysis also explores cross-sectional variation in the strength of the effect. In particular we document that common ownership has a stronger effect on subsequent correlation when the common owners are experiencing strong net flows either into or out of their funds. The flow effect is particularly strong when the stocks in the pair have relatively low float.

The effect we document is economically as well as statistically significant even though we restrict our analysis to big stocks, i.e. stocks with market capitalization above the NYSE median value. Indeed, when we take into account cross-sectional variation in the strength of the effect, we find that the predicted variation in the four-factor residual correlation on average ranges from -0.046 to 0.029.

The fact that fund ownership is endogenous is a key concern. Perhaps fund owners merely invest in stocks that have common fundamentals and thus naturally comove. Indeed, many, if not all, papers arguing that fund ownership causes comovement are vulnerable to this criticism.[1] We are the first to address this critique based on evidence from a natural experiment, the 2003 mutual trading scandal.[2] That unexpected event resulted in significant outflows to implicated funds which we show in turn results in exogenous variation in common ownership. We document that this exogenous variation strongly forecasts variation in abnormal return correlation. Thus, we provide novel evidence of a causal relation between stock connectedness and comovement.

Our analysis also explores cross-sectional variation in the strength of the effect. In particular we document that common ownership has a stronger effect on subsequent correlation when the common owners are experiencing strong net flows either into or out of their funds. The flow effect is particularly strong when the stocks in the pair have relatively low float.

Previous and current research looks at related questions: Is there information in institutional holdings about future returns? Or more particularly, does variation in assets under management result in price pressure? Most of these studies are concerned with cross-sectional and time series predictability of abnormal returns. Any implications for comovement are secondary, if examined at all. We begin by measuring comovement and then we turn to the implications for predictability of returns at the end of the analysis. In particular, we use the return on stocks that are connected to a particular firm to predict cross-sectional variation in average returns. Specifically, we define a stock's connected return as the return on a portfolio of all the stocks in our sample abnormally connected to the stock through common ownership at that point in time. By using the actual connected return as a measure of the price impact from mutual fund trading, we avoid having to measure the impact of flows on stock returns.[3]

We use the return on a stock's connected portfolio as a confirming signal that the stock has been temporarily pushed away from true value by mutual fund trading. Our cross-stock-reversal trading strategy generates significant abnormal returns of more than 9% per year, controlling for market, size, value, momentum, and the own-stock, short-term reversal factors.[4]

Finally, we use our connected return strategy to explain hedge fund index returns in standard performance attribution regressions. We show that the long-short hedge fund index loads negatively on our trading strategy and particularly so when the VIX index is increasing. In fact, the exposure of the value-weight long-short hedge fund index is more negative than the corresponding exposure of a value-weight portfolio of the active mutual funds in our sample. This suggests that the typical hedge fund may be part of the problem (causing the comovement) instead of part of the solution.

In summary, we show that understanding connectedness is a simple way to identify institutional-based stock comovement and its link to cross-stock reversal patterns. The rest of the paper is organized as follows. In Section 2, we summarize the literature. In Section 3, we describe our methodology and data sources. Section 4 presents our results. Section 5 concludes.

I. Literature Review

Our work has an intellectual link to papers studying whether crisis periods result in sudden increases in correlations across countries, a feature often referred to as contagion.[5] In particular, we follow Bekaert, Harvey, and Ng (2005) and define contagion as "excess correlation, that is, correlation over and above what one would expect from economic fundamentals." Like Bekaert, Harvey, and Ng, we take an asset pricing perspective to measuring economic fundamentals and identify contagion by the correlation of an asset pricing model's residuals. While Bekaert, Harvey, and Ng examine national equity market returns, we look instead to firm-level stock returns.

Our work is directly related to a growing literature on fund flows. That there is a relation between mutual fund flows and past performance (Ippolito, 1992; Chevalier and Ellison, 1997; Sirri and Tufano, 1998) is now well known. A recent paper by Coval and Stafford (2007) documents that extreme flows result in forced trading that temporarily moves prices away from fundamental value as in the general asset fire sales model of Shleifer and Vishny (1992) through the price pressure mechanism of

Scholes (1972). Ellul, Jotikasthira, and Lundblad (2010) and Mitchell, Pedersen, and Pulvino (2007) document broadly similar findings in the bond and convertible bond markets respectively.[6] Unlike these papers which study particular events, our analysis explores the extent to which institutional connections affect second moments more generally.

Recent theoretical work has emphasized the importance of delegated portfolio management and agency frictions to price movements such as these.[7] In particular, Vayanos and Woolley (2013) show that fund flows can generate comovement and lead-lag effects of the type we document. Their model provides strong theoretical motivation for our empirical analysis. More generally, beginning with Shleifer and Vishny (1997), researchers have studied the role of funding in arbitrage activity and the extent to which arbitrageurs should be expected to demand or provide liquidity.[8] On a related issue, Sadka (2010) shows that the typical hedge fund loads on a liquidity risk factor and that sensitivity to that liquidity risk is priced in the cross section of hedge fund returns. Measuring the extent to which hedge funds' performance can be attributed to a trading strategy that exploits temporary price dislocations resulting from institutional-driven comovement follows naturally from that theory and empirical evidence.

Two contemporaneous papers analyze issues related to comovement and institutional ownership.[9] First, Lou (2012) argues that flow-driven demand shocks more generally affect prices than just in the extreme fire-sale situations of Coval and Stafford. Lou shows that stocks with high expected flows from mutual funds comove. Of course, one might expect that mutual fund flows are to similar stocks that might naturally covary, so endogeneity is an issue.

Lou also examines the predictability of returns linked to mutual fund flows. Unlike Lou (or Coval and Stafford for that matter), we avoid having to measure the impact of flows on stock returns and instead use the actual connected return as a signal of the strength of the contagion effect resulting from ownership-based connections in the stock market.[10] Moreover, whereas Lou's focus is on momentum effects, we instead examine the way the presence of institutional connectedness generates cross-stock reversal patterns in returns.

Second, Greenwood and Thesmar (2011) argue that active mutual fund owners of stocks can have correlated trading needs and thus the stocks that they hold can comove, even if there are no overlapping holdings. Greenwood and Thesmar show that these correlated trading needs predict future price volatility and cross-sectional variation in comovement. Greenwood and Thesmar acknowledge that a concern with their interpretation is that funds with correlated flows might simply choose to invest in similar stocks whose fundamentals comove.

Researchers have adopted our techniques to study comovement in other markets. For example, Jotikasthira, Lundblad, and Ramadorai (2012) use our method to argue that investor flows into global funds affect equity prices, emerging market return correlations, and the correlation between emerging markets and developed markets. Bartram, Griffin, and Ng (2012) study connected stocks using international holdings data. They develop a similar measure of ownership linkages and show that return shocks linked to ownership connections as are important in explaining firm-level stock returns as industry and country returns.

Finally, like us, Chen, Hanson, Hong, and Stein (2008) explore whether hedge funds take advantage of the mutual fund flow-forced trading that Coval and Stafford document. They argue that hedge funds take advantage of that opportunity as average returns of long-short hedge funds are higher in months when the number of mutual funds in distress is large. In particular, Chen, Hanson, Hong, and Stein suggest that this evidence is consistent with hedge funds front-running the trades of distressed mutual funds. Our findings are consistent with their results but further show that the typical hedge fund apparently winds up on the wrong side of the price dislocation that we study.

II. Data and Methodology

A. Data and Sample

Stock returns, trading volume, and other relevant market and accounting data come from the intersection of the *CRSP* daily and monthly files and *COMPUSTAT*. We restrict our analysis to common stocks (share codes 10 and 11) from NYSE, AMEX and NASDAQ whose market capitalizations are above the NYSE median market cap (i.e., "big" stocks).[11] We choose these screening criteria because common ownership by active managers is not pervasive: small stocks, especially in the beginning of the sample, have little institutional ownership. Limiting the data in this way also keeps the sample relatively homogeneous and ensures that the patterns we find are not just due to small or microcap stocks (Fama and French, 2008).

The data on mutual fund holdings come from the merge between the *CDA/Spectrum* database provided by *Thomson Reuters* and the *CRSP Mutual Fund* database. We use the *Mutual Fund Links* dataset created by Russ Wermers and offered by Wharton Research Data Services. As our focus is on US active mutual funds, we remove index, tax-managed funds and international funds by applying standard screening criteria used in the literature.[12] We obtain data on analyst forecasts from the *Institutional Brokers Estimate System (I/B/E/S)* database. Our sample covers the period 1980 to 2008.

B. Measuring Common Ownership

At each quarter-end, we measure common ownership as the total value of stock held by the F common funds of the two stocks, scaled by the total market capitalization of the two stocks, and label this variable $FCAP_{ij,t}$. Thus, $FCAP_{ij,t} = \frac{\sum_{f=1}^{F}(S_{i,t}^{f}P_{i,t}+S_{j,t}^{f}P_{j,t})}{S_{i,t}P_{i,t}+S_{j,t}P_{j,t}}$, where $S_{i,t}^{f}$ is the number of shares held by fund f at time t trading at price $P_{i,t}$ with total shares outstanding of $S_{i,t}$, and similarly for stock j. We define common funds as those funds that held both stocks *i* and *j* in their portfolios at the end of quarter *t*. For each cross section, we calculate the normalized (to have zero mean and unit standard deviation) rank-transformed $FCAP_{ij,t}$, which we denote as $FCAP_{ij,t}^{*}$.

We primarily identify our findings from cross-sectional variation in common ownership. Indeed, by rank transforming *FCAP*, we hope to create a variable that is comparable across cross sections. On the one hand, this choice means that our analysis is not obviously going to be dominated by the later cross sections where there are many more mutual funds. On the other hand, time series variation in the number of mutual funds may help us identify the link between common ownership and excess comovement. Therefore, we also explore time-series variation in our effect by examining trends and sub-period variation in the series of Fama-MacBeth coefficients. Additionally, we show that our finding of a relation between common ownership and comovement is robust to not rank-transforming *FCAP*.

C. Modeling Cross-Sectional Variation in Comovement

We estimate cross-sectional regressions forecasting the within-month realized correlation ($\rho_{ij,t+1}$) of each stock pair's daily four-factor abnormal returns with $FCAP_{ij,t}$ and a host of pair characteristics that we use as controls:

$$\rho_{ij,t+1} = a + b_f * FCAP^*_{ij,t} + \sum_{k=1}^{n} b_k * CONTROL_{ij,k} + \varepsilon_{ij,t+1} \qquad (1)$$

Rather than pool the data, we estimate these regressions monthly and report the time-series average as in Fama and MacBeth (1973) as a robust way to avoid any issues with cross-correlation in the residuals. We then calculate Newey-West standard errors of the Fama-MacBeth estimates that take into account autocorrelation in the time series of cross-sectional estimates out to four lags.[13]

D. Controls

Our ultimate goal is to show that common ownership causes stocks to become more correlated. However, being able to forecast differences in comovement with common ownership may not be surprising if the predictability simply reflects the fact that fund managers choose to hold stocks that are similar. Those stocks would be expected to comove regardless of who owns the stock. The prototypical example is industry classification; we expect firms in similar industries to covary more, all else equal. Certainly fund managers may also tend to invest in a particular sector. To capture that similarity, we measure industry similarity as the number of consecutive SIC digits (*NUMSIC*), beginning with the first digit, that are equal for a given pair.

Similarly, managers may follow investment styles and/or strategies such as growth or value, small cap or large cap, momentum or reversal. Since previous research by Fama and French (1993) and Carhart (1997) has documented the link between these characteristics and sensitivity to common return factors, we expect higher correlation between two stocks if they have a greater similarity in the aforementioned characteristics. By forecasting the correlation of Fama-French-Carhart residuals, we hope that a good portion of the variation has already been removed.[14]

However, since style characteristics do not line up perfectly with regression loadings (Daniel and Titman, 1997), we also control for similarity in characteristics directly. To measure this similarity each quarter, we first calculate every stock's percentile ranking on a particular firm characteristic. Our measures of similarity, *SAMESIZE*, *SAMEBEME*, and *SAMEMOM*, are then just the negative of the absolute difference in percentile ranking for a particular characteristic across a pair.

Presumably analysts specialize in covering similar stocks. As a consequence, we create measures of common analyst coverage. Specifically, we measure the number of analysts $A_{ij,t}$ that issued at least one annual earnings forecast for both stocks *i* and *j* during the twelve month period preceding *t*.[15]

Chen, Chen, and Li (2012) documents that variables other than similarity in the aforementioned style characteristics forecast cross-sectional variation in pair-wise correlations. Their variables include various measures of past correlation: the past five-year monthly return correlation, $RETCORR_{ij,t}$; the past profitability correlation, $ROECORR_{ij,t}$; and the past correlation in the two stocks' abnormal trading volume, $VOLCORR_{ij,t}$. Their variables also include differences in key firm characteristics: the absolute value of the difference in five-year log sales growth rates, $DIFFGROWTH_{ij,t}$; the absolute difference in financial leverage ratios (defined as long-term debt / total assets), $DIFFLEV_{ij,t}$; and the absolute value of the difference in the two stocks' log share prices, $DIFFPRICE_{ij,t}$. Finally, Chen, Chen, and Li include dummies capturing important similarity between the two stocks in a pair. These controls include a dummy variable if the two firms are located in the same state, $DSTATE_{ij,t}$; a dummy variable if the two stocks belong to the S&P 500 index, $DINDEX_{ij,t}$; and a dummy variable if the two stocks are on the same stock exchange, $DLISTING_{ij,t}$. We include all of these variables as controls in many of our specifications.

As with our common ownership measure, we update these controls at each quarter-end, and, except for the dummy variables of Chen, Chen, and Li (2012), we do not use these variables directly but instead work with normalized rank transforms, which we continue to denote with an asterisk superscript. Finally, as institutional ownership is correlated with size, we also create very general size controls based on the normalized rank transform of the percentile market capitalization of the two stocks, *SIZE1* and *SIZE2* (where we label the larger stock in the pair as the first stock), and the interaction between these normalized size rankings.

III. Results

A. Forecasting Comovement

Table I Panel A confirms the well-known marked increase in active funds over this period. Our choice of restricting the analysis to big stocks results in a relatively stable number of stocks through time (approximately 800 to 900 stocks) and a large cross section of stock pairs to analyze (roughly from 300,000 to 400,000 pairs). Table I Panel B documents that the typical active manager holds 58.5 big stocks. As a consequence, a big stock has on average 68.6 active mutual fund owners. Of course, because of the growth of funds over this period, these full-sample numbers mask a strong trend in the number of funds holding a stock. In the early part of the sample (1980-1989), the median number of funds holding a typical big stock is 10. In the later part of the sample (2000-2008), that median number increases to 114.

<<Insert Table I about here>>

Our specific interest is how these numbers translate to the number of common owners for a pair of stocks. We report snapshots of the distribution of common owners in Table I Panel C. The sharing of active fund ownership with another stock is relatively common as more than 75% of all stock pairs have a common active fund owner. A typical pair in our sample has roughly nine funds in common which results in .77% of the combined market capitalization being owned by common funds. One expects from the increase in funds over this period that the number of ownership-based connections among big stocks has also increased dramatically, and we find this to be the case. In 1990, the median number of ownership connections is 1, owning .05% of the combined market equity. In 2008, the median number of ownership connections is 19, owning 0.88% of the combined market equity. Our use of rank-transformed variables in the main analysis is to make certain that these sorts of trends do not drive our results.

Table II Panel A reports the result from forecasting cross-sectional variation in four-factor residual correlation for the whole sample.[16] In the first column, specification (1), we estimate a simplified version of equation (1) with only common ownership, *FCAP**, as a forecasting variable. That variable is highly statistically significant, with a coefficient of 0.00395 and a *t*-statistic of 13.43.[17] The effect we find is economically significant as well, with fitted values ranging from an average minimum of -0.0007 to an average maximum 0.0124 around an average abnormal correlation of 0.0053.[18]

<<Insert Table II about here>>

In the second column, specification (2), we report the result of regressions incorporating our controls for style/sector similarity and common coverage. Recall that these control variables are normalized to have a standard deviation of one and transformed so that higher values indicate greater style similarity. The number of common analysts, *A**, forecasts abnormal return correlation with a statistically significant coefficient of 0.01437 (*t*-statistic of 11.92). This finding is consistent with analysts choosing to cover stocks that are similar. We also find a strong sector effect as the coefficient on *NUMSIC** is 0.00745 with a *t*-statistic of 12.39.

Turning to the style-based controls, we find that stocks with similar momentum characteristics covary; the coefficient on *SAMEMOM** is 0.00228 (*t*-statistic of 8.60). Similarity in book-to-market has a much smaller effect as the coefficient on *SAMEBM** is 0.00031 (*t*-statistic of 2.68). Of course, one should take into account the fact that we are already controlling for exposure to the book-to-market factor, *HML*, by forecasting the correlation of four-factor residuals rather than the correlation of raw returns.

The estimates in the second column also include very general size controls because of the strong correlation between institutional ownership and size. These controls include a third-order polynomial in *SAMESIZE** as well as a second-order polynomial in *SIZE1**, *SIZE2**, and the interaction between the two. We report the first-order terms of these polynomials in Table II, leaving the estimates on the higher-order terms for the Internet Appendix. Though these controls are important in describing cross-sectional variation in four-factor residual correlation, common ownership remains quite significant. In particular, the estimated coefficient is somewhat smaller (0.00256 from 0.00395) but remains very statistically significant when we control for style similarity as well as include non-linear controls for size.

In the third specification of Table II Panel A, we further control for a long list of pair characteristics from Chen, Chen, and Li (2012). These estimates are reported in Table AI in the Internet Appendix. All of the Chen, Chen, and Li variables are statistically significant and enter with the intuitive sign. Nevertheless, *FCAP** remains statistically significant.

The fourth and final column of Table II Panel A shows the results when we allow not only *SAMESIZE** but also *SAMEBM** and *SAMEMOM** to enter as a third-order polynomial. This regression also incorporates second-order polynomial functions of the rank-transformed book-to-market and momentum characteristics of each stock in the pair as well as interaction terms. In this regression, our most complete and quite flexible specification, common ownership remains economically and statistically significant.[19]

In Table II Panel B we show that our effect has increased over time. The fact that the effect becomes stronger is consistent with the significant increase in the number of active funds over the last thirty years. We first examine subsample estimates (one for each decade). For the sake of brevity, in each of the three subsamples we only show the coefficients on the *Constant* and *FCAP**, leaving the estimates on the control variables for Table AI in the Internet Appendix. When using all controls, we find that the coefficient on *FCAP** moves from a relatively low (but still statistically significant) 0.00087 in the 1980s, to 0.00197 in the 1990s, and then to a relatively high estimate of 0.00281 in the 2000s.

Table II Panel C shows results from an alternative approach to measuring the evolution of the common ownership effect over time. In particular, we regress the time series of Fama-MacBeth coefficients on a constant and a trend. Again, we only show the estimate for the *Constant* and *FCAP** for brevity, relegating estimates for the remaining controls to the Internet Appendix, Table AI. In each of the four specifications, we find that the trend coefficient in the specification is positive and statistically significant. For example, the estimate for specification (4), which includes all controls, is 0.00002 and very significant, with a *t*-statistic of 3.61.

As a final check of the reasonableness of our method of linking cross-sectional variation in connectedness to excess comovement, we have re-estimated the fourth specification of Table I Panel A using *FCAP* instead of *FCAP** as well as simple counts of the number of common owners of the pair (*F* and *F**). These results can be found in Internet Appendix Table AVI. All of our findings continue to hold. To summarize, the main conclusion from Table II is that common ownership forecasts statistically and economically significant variation in comovement.[20]

B. Mutual Fund Scandal – A Natural Experiment

Though the results in the previous section are suggestive, ultimately they cannot rule out the possibility that mutual fund managers choose to hold similar stocks.

To confirm that the link we have found is causal, we exploit a natural experiment based on the mutual fund scandal that occurred in September 2003. During that month, 25 fund families settled allegations of illegal trading that included market timing and late trading. Implicated families had significant outflows as a consequence of the scandal. Kisin (2011) estimates that funds of implicated families lost 14.1% of their capital within one year and 24.3% within two years. These outflows continued until the end of 2006. In stark contrast, families not implicated increased their capital by nearly 12% over the same period. We argue that capital flow arising from this scandal is an exogenous shock, providing a useful way to eliminate concerns that our results are simply due to the endogenous choice of fund managers.

Our specific instrument is the ratio (*RATIO*) of the total value held by all common "scandal" funds of the two stocks over the total value held by all common funds, as of the time when the scandal broke. To make sure our findings are robust, we use either *RATIO* or a dummy that equals 1 if *RATIO* is above-median and present results for both measures.

In the first stage, we regress *FCAP** on *RATIO*, the level of *FCAP** as of September of 2003 (*FCAP200309**), and the various sets of control variables used in regressions (1) – (4) in Table II. We find that the coefficient on *RATIO* in these four first-stage regressions, reported in Table AII of the Internet Appendix, is always very significant.

Table III presents the results of the second stage regressions where we use the fitted *FCAP** from the first stage to predict cross-sectional variation in four-factor residual correlations. For the sake of brevity, Table III only reports the coefficient on the instrumented *FCAP**. All other coefficient estimates are reported in Table AIII of the Internet Appendix.

<<Insert Table III about here>>

The first row of estimates in Table III correspond to the 2SLS coefficients on instrumented *FCAP** when the continuous variable *RATIO* is the instrumental variable. The coefficient in specification (1), where there are no controls, is 0.04204, which is highly significant with a *t*-statistic of 5.39. This causal effect is economically significant as well, with fitted values ranging from an average minimum of -0.0019 to an average maximum 0.0184 around an average abnormal correlation of 0.0089. As one moves across the column to the right of the table, including more and more controls, the coefficient decreases somewhat as well as the *t*-statistic, but remains statistically significant, even for specification (4), which includes all of our controls.

Similar patterns can be found in the second row of estimates where a dummy for above-median *RATIO* is used as the instrumental variable. The last row of estimates in Table III show the coefficients of the OLS regression estimated in Table II but for the subsample where the natural experiment takes place (from December 2003 to December 2006) and with *FCAP200309** added as an additional control, for purposes of comparison. These OLS results are weaker than the 2SLS, consistent with the endogeneity of common ownership being important.

To summarize, Table III shows that the mutual fund scandal generates exogenous variation in common ownership which causes abnormal return correlation in the following month.

C. When does connectedness matter?

Table II documents that institutional connectedness helps predict cross-sectional variation in comovement and Table III confirms that the relation is causal. The remainder of the analysis will focus on exploring why connecting stocks through common fund ownership matters and understanding the implications of the finding for the cross section of average returns and for performance attribution of hedge funds.

A likely explanation for our findings is price pressure arising from mutual fund flows as in Coval and Stafford (2007) and Lou (2012). Indeed the 2003 mutual fund scandal was chosen as a natural experiment because of the unexpected flows that event produced. To provide additional evidence supporting this interpretation, we exploit cross-sectional heterogeneity in stock pair characteristics. Specifically, in Table IV, we interact *FCAP** with two variables, the total float capitalization of the two stocks in the pair, *PFLOAT**, and the *absolute value* of the total net flow into the common funds (those that held both i and j at time t), *PFLOW*, over the previous quarter.

We follow the literature in defining flows (see Coval and Stafford, 2007). Therefore, we measure the net investment flow of funds into fund k in quarter t as:

$$FLOW_{k,t} = TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}) \tag{2}$$

where $TNA_{i,t}$ is the Total Net Assets of fund i in quarter t and $R_{i,t}$ is the fund return over the period $t - 1$ to t reported by CRSP Mutual Fund Database. TNA is reported quarterly before 1991 and monthly thereafter. To compute quarterly flows after 1990, we simply sum monthly flows.

Then, for each pair, we compute *PFLOW* by measuring the absolute value of the sum of *FLOW* over the K common funds,

$$PFLOW_{i,j,t} = ABS\left(\sum_{k=1}^{K} FLOW_{k,t}\right) \tag{3}$$

We do not rank transform *PFLOW* because of the evidence in Coval and Stafford (2007) that extreme flows are particularly important in understanding price pressure. However, we continue to normalize *PFLOW* to facilitate the economic interpretation of the measured coefficients.

Thus, *PFLOW* measures the total amount of net capital flowing into the common owners. Of course, pairs of stocks may differ in their capacity to accommodate that potential flow. We use the total value of shares available for trading between the two stocks as a proxy for that capacity. Specifically, we measure *PFLOAT* as the total dollar amount of market capitalization available for trading for the pair of stocks in question. *PFLOAT** represents the normalized (to have zero mean and unit standard deviation) rank transform of *PFLOAT*. The analysis is limited to the 1991-2008 subperiod when *PFLOAT* becomes available.

In Table IV we add to the regressions in Table II interaction terms between *FCAP** and the two characteristics described above as well as a triple interaction among *FCAP**, *PFLOAT**, and *PFLOW*.[21] We find that *FCAP** more strongly forecasts comovement for pairs of smaller stocks, as the coefficients for the interaction of *FCAP** with *PFLOAT** are negative, though not statistically significant in all four specifications. Similarly, *FCAP** more strongly forecasts comovement when *PFLOW* is high, i.e., when total flows to common funds are very high or very low. This interaction is always very statistically significant. Finally, the triple interaction reveals that the link between common ownership and future comovement is the strongest when common owners of a pair of low-float stocks experience extreme flows. This interaction is statistically significant in three of the four specifications.

The predicted variation in abnormal return correlation conditional on these two pair characteristics is now of course larger than in Table II. In specification (1), for example, fitted values linked to *FCAP** now range from an average minimum of -0.0462 to an average maximum 0.0291.

These results are consistent with the causal interpretation provided by the natural experiment: prices are more subject to non-fundamental comovement when pressure from net mutual fund trading is high and especially so when the stocks in question have low float and are thus less liquid.

<<Insert Table IV about here>>

The Internet Appendix further analyzes the nature of this comovement using the Campbell (1991) return decomposition, as applied in Campbell, Polk, and

Vuolteenaho (2010).[22] The results presented in Table AXI show that the majority of the variation in comovement linked to connectedness arises from the covariance of cash-flow news and discount-rate news. This finding is consistent with our finding of a causal relation between connectedness and comovement. News about fundamentals in one stock (cash-flow news) results in temporary movements (discount-rate news) in another stock as connected owners sell off positions in both stocks.[23]

D. Connected stock trading strategies

If common ownership causes comovement through price pressure, then we can link ownership-based connectedness to predictable variation in returns. Specifically, if stock *i* experiences a negative shock and connected stock *j*'s price also drops, we conjecture that the stock *i*'s drop is due to price pressure, which we expect to revert. Our trading strategy is thus very simple: we buy (sell) stocks that have gone down (up) if their connected stocks have gone down (up) as well. This strategy uses the connected return as a confirming signal of whether stock *i* is undervalued or overvalued, and we interpret such a strategy as exploiting the price pressure induced by common ownership.[24] Moreover, since one would expect that forced sells are more likely than forced buys, we conjecture that most of the predictable variation will come from the long side of the strategy.

Since our ownership data is updated quarterly, we look for these patterns using past three-month returns. By using the actual connected return as a measure of the price impact from mutual fund trading, we avoid having to measuring the impact of flows on stock returns. In particular, each month we sort our subset of big stocks into quintiles based on their past three-month return. We independently sort these stocks into quintiles based on the past three-month return on the portfolio of stocks that they are connected to through common ownership($r_{iC,t}$). We exploit the results of Table II to construct the connected portfolio and isolate the incremental effect of common ownership as best as possible. Specifically, we first measure the degree of abnormal connections (*FCAPO**) by orthogonalizing *FCAP** each period to the full set of controls of regression (4) in Table II.[25] We then use $FCAPO^*_{ij,t}$ to generate the weights determining $r_{iC,t}$. Specifically, we define

$$FCAP^{**}_{ij,t} = rank(FCAPO^*_{ij,t}) \text{ if } FCAP_{ij,t} > 0 \tag{4}$$

$$FCAP^{**}_{ij,t} = 0 \text{ if } FCAP_{ij,t} = 0 \tag{5}$$

The return on the portfolio is then

$$r_{iC,t} = \frac{\sum_{j=1}^{J} FCAP^{**}_{ij,t-1} r_{j,t}}{\sum_{j=1}^{J} FCAP^{**}_{ij,t-1}} \tag{6}$$

We first examine the buy-and-hold abnormal returns on an equal-weight portfolio of stocks that are in the low own-return and low connected-return portfolio (which we dub the *low* portfolio) and an equal-weight portfolio of stocks that are in the high own-return and high connected-return portfolio (the *high* portfolio). Specifically, we generate the cumulative buy-and-hold abnormal return by regressing the $t+1, t+2, \ldots, t+12$ returns on the corresponding returns on five factors: the four factors of Fama and French (1993) and Carhart (1997), augmented with the short-term reversal factor (*STR*) that makes a size-stratified bet on the one-month reversal effect.[26] We include *STR* as we are sorting stocks on their past return, and we want to make sure that our connected stock strategy is distinct from that effect.[27]

Figure 1 graphs the cumulative abnormal returns on the *low* portfolio, the *high* portfolio, and a portfolio that is long *low* and short *high*. The patterns shown in the graph are consistent with stocks being pushed away from fundamental value by mutual-fund trading, with the connected return being a useful measure of the extent of that temporary misvaluation. These patterns last for the next six months. Thus, compared to the standard short-term reversal effect, the misvaluation is larger but takes more time to revert. Consistent with the forced-selling / price-pressure of Coval and Stafford (2007), we find that more of the effect comes from the *low* portfolio than the *high* portfolio.

<<Insert Figure 1 about here>>

As a consequence, we evaluate the average returns on "composite" portfolios that take these predictable patterns in the cross section of average returns into account. To ensure our findings are not due to the one-month reversal effect, we first skip a month after the sort and then hold the stocks in question for five months, following the methodology of Jegadeesh and Titman (1993). Table V reports the five-factor alphas on these 25 composite portfolios, each of which is an equal-weight average of the corresponding simple strategies initiated one to five months prior.

There are two general patterns in Table V that are consistent with our initial conclusions concerning Figure 1. Alphas on the composite portfolios generally increase as one moves from high to low connected return within an own-return quintile. Within each connected return quintile, the alphas on these composite portfolios generally increase as one moves from high to low own return. These patterns confirm the information in the connected return about a stock's subsequent return; as a consequence, we design a long-short trading strategy to exploit that information. Our connected-stock strategy, which we denote by *CS*, buys the composite *low* portfolio and sells the composite *high* portfolio. The five-factor alpha for *CS* is an impressive 76 basis points per month (*t*-statistic of 4.96).[28] Consistent with our forced-trading story, more than 71% of the alpha comes from the composite *low* portfolio.

<<Insert Table V about here>>

For comparison, Table V also reports the average five-factor alpha on a strategy that instead buys the average (across the own-return quintiles) low connected return composite portfolio and sells the average (across the own-return quintiles) high connected return portfolio. This strategy earns 36 basis points per month (*t*-statistic of 4.13).[29] Since such a strategy ignores the information in the interaction between a stock's own return and its connected return, the strong performance further confirms that our finding of a successful connected stock trading strategy is distinct from the short-term reversal effect.

Table VI reports the regression loadings from the performance attribution for the CS trading strategy as well as shows the effect of including additional variables in the attribution regression. We first include the liquidity factor of Pastor and Stambaugh (2003) to measure how *CS* covaries with this factor.[30] We find that *CS* does covary positively with their factor though the coefficient is only marginally significant. We then include a trend in the regression; we find no evidence that the alpha of our connected stocks trading strategy has been trending over time. Finally, we include dummy variables for each quarter. We find that our effect is strongest in the fourth quarter of the year but that alphas in each of the four quarters remain economically and statistically significant.

<<Insert Table IV about here>>

E. Hedge Fund Index attribution

Our last analysis uses our connected stocks trading strategy to analyze the performance of returns on two *CSFB/Tremont Hedge Fund Indexes*. These indexes have been used in a number of studies including Asness, Krail, and Liew (2001); Agarwal and Naik (2004); Getmansky, Lo, and Makarov (2003); and Bondarenko (2004). The first index we study is the index of all hedge funds. As *CFSB* weights hedge fund returns by assets under management and captures more than 85% of all assets under management in this investing space, this index gives a good representation of the extent to which the general health of the hedge fund industry can be linked to our connected stock strategy.[31] We also examine the performance of the Long/Short component of the CSFB index to measure the extent to which funds that specifically invest in equities are exposed to the connected stocks factor.

Table VII reports the results of this analysis. We find that hedge funds in general and long/short managers in particular load negatively on the connected stocks trading strategy. The coefficient in the first column of Table VII estimates a regression of the overall hedge fund index excess return on the return on our connected strategy and the four factors of Fama and French (1993) and Carhart (1997), augmented with the short-term reversal factor, *STR*. We include the interaction of the connected strategy with the lagged value of the change in the VIX index (ΔVIX) to capture any time-variation in the sensitivity to *CS*. Both theory and evidence suggest that an increase in the VIX is a good measure of turbulent times for the hedge fund industry.[32] We demean and standardize ΔVIX so that the coefficient on *CS* reflects the return loading for typical values of ΔVIX and the interaction coefficient indicates how a one-standard-deviation move affects that loading.

<<Insert Table VII about here>>

The coefficients on *CS* and *CS* interacted with ΔVIX for the all hedge fund index are negative but statistically insignificant in a six-factor regression that adds *CS* to the five factors of Table VI. The second column of the Table instead attributes the performance of the hedge fund index to the connected strategy, its interaction with ΔVIX, and the eight hedge fund factors of Fung and Hsieh (2001, 2004).[33] Though hedge funds in the aggregate load on these eight factors to various degrees, our connected stocks factor is important in describing the returns on hedge funds. The coefficient is now more economically and statistically significant; the point estimate for *CS* is -0.1306 and has an associated *t*-statistic of -5.38. The interaction term remains statistically insignificant. This result suggests that our trading strategy is a useful tool to measure the state of the hedge fund industry as the typical hedge fund is negatively exposed to our factor, regardless of whether the VIX is increasing or decreasing.

Perhaps more interesting results are in columns 3 and 4 of Table VII. In column 3, we measure the degree to which the Long/Short subset of hedge funds covaries with our connected return trading strategy in the presence of the Fama-French/Carhart factors and the short-term reversal factor. In column 4, we use the Fung and Hsieh factors as controls instead. In both cases, we find that the returns on this subset of hedge funds strongly and negatively covary with our connected return factor and particularly so in turbulent times (when VIX is increasing).

Specifically, we find that the average sensitivity to *CS* for long-short hedge funds is more than 70% larger in absolute value than the corresponding estimate for all hedge funds. The *t*-statistics are correspondingly larger and indicate strong statistical significance (a *t*-statistic of -2.88 when controlling for the five equity factors and a *t*-statistic of -9.25 when controlling for the eight Fung and Hsieh factors). The magnitude of the interaction coefficients is more than twice as large for the long-short fund subset compared to all hedge funds and *t*-statistics in both attribution regressions are -3.29 and -2.49 respectively). These findings are comforting as one would expect this subset of hedge funds to be more exposed to our factor on average and particularly so when the VIX is increasing.

For the sake of comparison, we also estimate the loading of a value-weight portfolio consisting of all of the active mutual funds in our sample over the same time period. This portfolio has a smaller (in absolute value) sensitivity to the connected strategy as the estimate is -0.0225 with an associated *t*-statistic of -1.78. Though we do not observe complete holdings data for all hedge funds and therefore cannot see the exact positions of these long/short hedge funds, these results suggest that these hedge funds do not take full advantage of the opportunities that price pressure from mutual fund flows provide. Indeed, these results are consistent with hedge funds exacerbating rather than mitigating the price pressure patterns documented in this paper.[34]

Figure 2 provides evidence on why it is not surprising that the typical hedge fund loads negatively on our connected strategy. This figure plots both the loadings of the two hedge fund indexes on the connected strategy as well as the cumulative abnormal return on the connected strategy in event time, where the event is the forming of the connected stock trading strategy. One reasonable interpretation of this figure is that hedge funds follow a momentum strategy that effectively front-runs mutual fund flows. However, the typical hedge fund is unable to exit its positions in time and therefore exacerbates the price dislocation they help initiate.

<<Insert Figure 2 about here>>

Of course, this finding begs the question as to whether it is value-adding for long-short hedge funds to follow a front-running strategy. In Table VII, the unconditional loading on *CS* is -0.0989 while the corresponding loading on *MOM* is 0.1353. Moreover, the loading on CS decreases by 0.0140 for every unit change in the normalized ΔVIX. Since we find that MOM has an alpha (controlling for the other factors in Table VII, including *CS*) of 71 basis points, the total incremental abnormal return for this joint exposure to *MOM* and *CS* is small but remains positive unless the normalized change in VIX is nearly 2. Of course, this analysis ignores any potential time-variation in the sensitivity of hedge funds to *MOM.* Thus, the extent to which hedge funds can effectively front-run the anticipated forced trades of mutual funds remains an interesting question for future research.[35]

IV. Conclusion

We show that stocks are connected through the mutual fund owners they have in common. In particular, pairs of stocks that are connected in this fashion covary more together, controlling for exposure to systematic return factors, style and sector similarity, and many other pair characteristics. Evidence from a natural experiment—the 2003 mutual fund trading scandal—confirms that this relation is causal. Consistent with these findings, we further show that the general link between shared ownership and comovement is stronger for owners who are experiencing extreme flows in low float stocks.

These results motivate a novel cross-stock reversal trading strategy that exploits the information in ownership connections to generate annual abnormal returns of more than 9%. As a consequence, we provide a simple way to document the extent to which ownership-based connections result in equity market contagion. In an application, we document that the typical long-short hedge fund covaries negatively with our trading strategy (and more so than the typical mutual fund we initially study), suggesting that hedge funds on average may be part of the cause of the excess covariation and price dislocation that contagion from ownership-based connections generates.

References

Ang, Andrew, Sergiy Gorovyy, and Gregory B. van Inwegen, 2011, Hedge Fund Leverage, *Journal of Financial Economics* 102.

Antón, Miguel, 2011, Cash-flow Driven Covariation, IESE working paper.

Agarwal, Vikas and Narayan Y. Naik, 2004, Risk and Portfolio Decisions involving Hedge Funds, *Review of Financial Studies* 17, 63-98.

Asness, Clifford, Robert Krail, and John Liew, 2001. Do Hedge Funds Hedge?, *Journal of Portfolio Management*, 28, 1, pg. 6.

Bekaert, Geert, Campbell R. Harvey, and Angela Ng, 2005, Market Integration and Contagion, *Journal of Business* 78, 39-79.

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

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283-317.

Bartram, Söhnke, John Griffen, and David Ng, 2012, How Important are Foreign Ownership Linkages for International Stock Returns, University of Texas working paper.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2012, Hedge Fund Stock Trading in the Financial Crisis of 2007-2008, *Review of Financial Studies* 25, 1-54.

Bondarenko, Oleg, 2004, Market Price of Variance Risk and Performance of Hedge Funds, University of Illinois Chicago working paper.

Brunnermeier, Markus and Lasse Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Campbell, John Y., Christopher Polk, and Tuomo Vuolteenaho, 2010, Growth or glamour? Fundamentals and systematic risk in stock returns, *Review of Financial Studies* 23, 305-344.

Carhart, Mark, 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 56-82.

Chen, Huafeng, Shaojun Chen, and Feng Li, 2012, Empirical investigation of an equity pairs trading strategy, UBC working paper.

Chen, Joseph, Samuel Hanson, Harrison Hong, and Jeremy C. Stein, 2008, Do hedge funds profit from mutual-fund distress?, Harvard University working paper.

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

Coval, Josh and Eric Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479-512.

Cremers, 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 and Sheridan Titman, 1997. Evidence on the Characteristics of Cross Sectional Variation in Stock Returns, *The Journal of Finance*, Vol. 52, No. 1, pp. 1-33.

DeMiguel, Victor, Lorenzo Garlappi and Raman Uppal, 2007, Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?, *Review of Financial Studies* 22, 1915-1953.

Duffie, Darrell, 2010, Asset Price Dynamics with Slow-Moving Capital, *Journal of Finance* 65, 1238-1268.

Ellul, Andrew, Pab Jotikasthira, and Christian T. Lundblad, 2010, Regulatory Pressure and Fire Sales in the Corporate Bond Market, *Journal of Financial Economics* 101, 596-620.

Fama, Eugene F. and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F. and Kenneth R. French, 2008, Dissecting Anomalies, *Journal of Finance* 63, 1653-1678.

Fama, Eugene and James D. MacBeth, 1973. Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Forbes, Kristin and Roberto Rigobon, 2001, No contagion, only interdependence: Measuring stock market co-movements, *Journal of Finance* 57, 2223-2261.

Fung, William and David A. Hsieh, 2001, The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers, *Review of Financial Studies* 14, 313-341.

Fung, William and David A. Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 60, 65-80.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529-609.

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

Gromb, Denis and Dimitri Vayanos, 2002, Equilibrium and Welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361-407.

Gromb, Denis and Dimitri Vayanos, 2010, Limits of Arbitrage: State of the Theory, *Annual Review of Financial Economics*, 251-275.

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

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

Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai, 2012, Asset Fire Sales and Purchases and the International Transmission of Funding Shocks, *Journal of Finance* 67 2015-2050.

King, Mervyn, Enrique Sentana, and Sushil Wadhwani, 1994, Volatility and Links between National Stock Markets, *Econometrica* 62, 901-933.

King, Mervyn and Sushil Wadhwani, 1990, Transmission of Volatility Between Stock Markets, *Review of Financial Studies* 3, 5-33.

Kisin, Roni, 2011, The Impact of Mutual Fund Ownership on Corporate Investment: Evidence from a Natural Experiment, Washington University of St Louis Working Paper.

Lou, Dong, 2012, A flow-based explanation for return predictability, forthcoming *Review of Financial Studies*.

Lou, Dong and Christopher Polk, 2013, Comomentum: Inferring arbitrage activity from return correlations, LSE working paper.

Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007. Slow Moving Capital, *American Economic Review*, vol. 97(2), pages 215-220, May.

Pastor, Lubos and Robert Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Rigobon, Roberto, 2002, Contagion: How to measure it?, in "Preventing Currency Crises in Emerging Markets" Editors: Sebastian Edwards and Jeffrey Frankel, The University Chicago Press, Chicago, 269-334.

Rozeff, Michael, 1984, Dividend yields are equity premiums, *Journal of Portfolio Management* 11, 68-75.

Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80, 309-349.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge-fund returns, *Journal of Financial Economics* 98, 54-71.

Scholes, Myron, 1972, The market for corporate securities: Substitution versus price pressure and the effects of information on stock prices, *Journal of Business* 45, 179-211.

Shleifer, Andrei and Robert Vishny, 1992, Liquidation values and debt capacity: a market equilibrium approach, *Journal of Finance* 47, 1343-1366.

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

Singleton, Kenneth J., 2011, Investor Flows and the 2008 Boom/Bust in Oil Prices, Stanford University working paper.

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

Sun, Zheng, 2007, Clustered institutional holdings and stock comovement, NYU working paper.

Vayanos, Dimitri, 2004, Flight to quality, flight to liquidity, and the pricing of risk, London School of Economics working paper.

Vayanos, Dimitri and Paul Woolley, 2013, An institutional theory of momentum and reversal, *Review of Financial Studies* 26, 1087-1145.

Vijh, A., 1994, S&P 500 trading strategies and stock betas, *Review of Financial Studies* 7, 215-251.

Table I: Summary Statistics

This table reports summary statistics for all NYSE-AMEX-NASDAQ stocks that are above the median NYSE market capitalization as of the end of each quarter and the active funds that invest in those stocks, from 1980 to 2008. Panel A lists the total number of stocks, pairs of stocks, and funds at the end of the fourth quarter for the first and the last year of the sample, and for every five years. The number of unique stock pairs is $[n * (n-1)]/2$, where n is the number of stocks. We have 41,374,135 pair-quarters in our sample. Panel B reports summary statistics on the number of stocks held by each fund and the number of funds that hold each stock, for both the full sample and each decade within the sample. Panel C reports the distribution of common fund ownership ($FCAP_{ij,t}$), which measures the total value of stock held by all common funds of the two stocks, scaled by the total market capitalization of the two stocks, as of the end of the quarter. The distribution is shown for the average of the entire sample (ALL), for the first and the last year of the sample, and for every five years.

PANEL A: Number of Stocks, Pairs, and Funds

Year	Stocks	Pairs	Funds
1980	895	400065	171
1985	828	342378	251
1990	809	326836	469
1995	905	409060	1002
2000	902	406351	1780
2005	803	322003	1901
2008	746	277885	1642

PANEL B: Summary Statistics for Stocks and Funds

Variable	Period	Mean	Median	Std	Min	Max
Num stocks per fund	**1980-2008**	**58.5**	**41**	**69.8**	**1**	**1028**
	1980-1989	40.4	33	35.1	1	439
	1990-1999	54.4	40	61.7	1	932
	2000-2008	64.0	44	77.7	1	1028
Num funds per stock	**1980-2008**	**68.6**	**39**	**84.1**	**1**	**829**
	1980-1989	14.1	10	14.2	1	167
	1990-1999	58.0	42	56.3	1	603
	2000-2008	141.8	114	103.1	1	829

PANEL C: The Cross-sectional Distribution of Common Fund Ownership

			Percentiles						
Year	Mean	Std	0%	25%	50%	75%	95%	99%	100%
ALL	0.0077	0.0142	0.0000	0.0000	0.0020	0.0086	0.0356	0.0691	0.8143
1980	0.0016	0.0045	0.0000	0.0000	0.0000	0.0005	0.0099	0.0224	0.0970
1985	0.0018	0.0045	0.0000	0.0000	0.0000	0.0015	0.0103	0.0218	0.1013
1990	0.0035	0.0072	0.0000	0.0000	0.0005	0.0038	0.0168	0.0335	0.1447
1995	0.0080	0.0141	0.0000	0.0004	0.0022	0.0091	0.0355	0.0679	0.2755
2000	0.0119	0.0176	0.0000	0.0012	0.0051	0.0148	0.0485	0.0828	0.2581
2005	0.0152	0.0185	0.0000	0.0031	0.0080	0.0199	0.0545	0.0859	0.2164
2008	0.0156	0.0182	0.0000	0.0037	0.0088	0.0205	0.0531	0.0856	0.2922

Table II: Connected Comovement

This table reports Fama-MacBeth estimates of monthly cross-sectional regressions forecasting the correlation of daily Fama-French-Carhart residuals in month $t+1$ for the sample of stocks defined in Table I. The independent variables are updated quarterly and include our measure of institutional connectedness, the total ownership value held by all common funds of the two stocks scaled by the total market capitalization of the two stocks, $FCAP_{ij,t}$, and a series of controls at time t. We measure the negative of the absolute value of the difference in size, BE/ME and momentum percentile ranking across the two stocks in the pair ($SAMESIZE_{ij,t}$, $SAMEBM_{ij,t}$, and $SAMEMOM_{ij,t}$ respectively). We also measure the number of similar SIC digits beginning with the first digit, $NUMSIC_{ij,t}$, for the two stocks in a pair. Thirty nine other controls are also included, but reported in the Internet Appendix, Table AI, for the sake of brevity. All independent variables, excluding dummy variables, are then rank transformed and normalized to have unit standard deviation, which we denote with an asterisk superscript. We calculate Newey-West standard errors (four lags) of the Fama-MacBeth estimates that take into account autocorrelation in the cross-sectional slopes. We report estimates of regressions using various subsets of these variables in Panel A. Panel B shows the same set of regressions as in Panel A, but for three different subsamples (corresponding to the three decades of the sample: the 1980s, 1990s, and 2000s). In Panel C, we report the results of regressing the time series of Fama-MacBeth *FCAP** coefficients on a constant and a trend. In Panels B and C, only the *FCAP** coefficient and the intercept estimates are shown; the remaining controls are reported in Table AI in the Internet Appendix.

PANEL A: Full sample (1980-2008)

	Dependent Variable: Correlation of 4F residuals			
	(1)	(2)	(3)	(4)
Constant	0.00508	0.00512	0.00284	0.00288
	(11.30)	(11.17)	(6.92)	(6.85)
*FCAP**	0.00395	0.00256	0.00168	0.00184
	(13.43)	(11.61)	(8.58)	(9.85)
*A**		0.01437	0.01342	0.01334
		(11.92)	(11.83)	(11.77)
*SAMESIZE**		-0.00365	-0.00396	-0.00402
		(-1.43)	(-1.53)	(-1.54)
*SAMEBM**		0.00031	-0.00024	-0.00001
		(2.68)	(-2.80)	(-0.00)
*SAMEMOM**		0.00228	0.00143	-0.00736
		(8.60)	(6.87)	(-2.36)
*NUMSIC**		0.00745	0.00676	0.00671
		(12.39)	(12.22)	(12.03)
*SIZE1**		0.04683	0.04816	0.04855
		(11.90)	(11.84)	(11.66)
*SIZE2**		0.01012	0.01021	0.01033
		(2.78)	(2.79)	(2.83)
SIZE1 x SIZE2**		-0.06530	-0.06750	-0.06692
		(-12.2)	(-11.8)	(-11.8)
Additional Controls (reported in the Internet Appendix)				
Non-linear size controls	No	Yes	Yes	Yes
Pair characteristic controls	No	No	Yes	Yes
Non-linear style controls	No	No	No	Yes

Table II: Connected Comovement (continued)

PANEL B: Subsample analysis of Panel A

	Dependent Variable: Correlation of four-factor residuals			
	(1)	(2)	(3)	(4)
		First subsample (1980-1990)		
Constant	0.00265	0.00264	0.00158	0.00161
	(15.61)	(15.39)	(5.20)	(5.34)
*FCAP**	0.00164	0.00135	0.00084	0.00087
	(11.81)	(12.59)	(8.20)	(8.77)
		Second subsample (1991-2000)		
Constant	0.00493	0.00500	0.00255	0.00256
	(5.75)	(5.67)	(2.85)	(2.79)
*FCAP**	0.00439	0.00266	0.00182	0.00197
	(12.46)	(12.06)	(9.81)	(11.76)
		Third subsample (2001-2008)		
Constant	0.00811	0.00815	0.00470	0.00478
	(16.88)	(16.57)	(11.06)	(10.34)
*FCAP**	0.00607	0.00385	0.00246	0.00281
	(13.77)	(7.13)	(4.56)	(5.76)

Other controls (including non-linear controls on size, BM, Mom) reported in the Internet Appendix.

PANEL C: Measuring a trend in the time-series of Fama MacBeth coefficients

	Dependent Variable: Corresponding coefficient estimates from Panel A			
	(1)	(2)	(3)	(4)
		Intercept		
Constant	0.00508	0.00512	0.00284	0.00288
	(15.58)	(15.30)	(7.92)	(7.78)
*FCAP**	0.00395	0.00256	0.00168	0.00184
	(20.89)	(13.58)	(9.17)	(11.07)
		Trend		
Constant	0.00008	0.00008	0.00005	0.00005
	(9.53)	(9.47)	(5.62)	(5.34)
*FCAP**	0.00006	0.00003	0.00002	0.00002
	(9.26)	(4.55)	(2.64)	(3.61)

Other controls (including non-linear controls on size, BM, Mom) reported in the Internet Appendix.

Table III: Natural Experiment

This table reports results from a 2SLS instrumental variables regression based on a natural experiment. In September 2003, 25 fund families experienced large outflows of capital as a consequence of a settlement regarding alleged illegal trading. We argue that the outflow of capital because of this scandal is an exogenous shock unrelated to the endogenous investment decisions of fund managers. In the first stage we predict the variable *FCAP** with the ratio (*RATIO*) of the total ownership value held by all common "scandal" funds of the two stocks over the total ownership value held by all common funds, both measured as of the time of the scandal (end of September of 2003). The second stage of the regression uses the fitted *FCAP** to forecast the correlation of daily Fama-French-Carhart residuals in month $t+1$. Four different specifications are shown, corresponding to the specifications shown in Table II. We add to each of those four groups of controls an additional control for the level of connectedness as of the scandal, *FCAP_200309**. The first row of estimates directly uses the IV variable, *RATIO*, in the first stage. The second row of estimates replaces *RATIO* in the first stage with a dummy variable that equals *1* if *RATIO* is above its median value. The final row of estimates report the coefficients on *FCAP** in an OLS regression corresponding to Table II, but for this specific subsample (January 2004 to December 2006) and with *FCAP_200309**, for comparison with the IV analysis. The coefficients for the first stage are not shown in this table and are instead reported entirely in Table AII in the Internet Appendix. Similarly, estimates on all control variables in the second stage in each of the twelve specifications are not shown here but reported in Table AIII in the Internet Appendix.

		Dependent Variable: Correlation of four-factor residuals			
Specification	Instrument	(1)	(2)	(3)	(4)
2SLS	*RATIO*	0.04204	0.04261	0.03345	0.02874
		(5.39)	(3.51)	(2.41)	(2.03)
2SLS	*RATIO*>Median[*RATIO*]	0.04107	0.04007	0.03222	0.02826
		(5.98)	(3.44)	(2.50)	(2.15)
OLS		0.00181	0.00062	-0.00026	0.00041
		(4.43)	(1.47)	(-0.47)	(0.83)
Additional Controls (reported in the Internet Appendix)					
Non-linear size controls		No	Yes	Yes	Yes
Pair characteristic controls		No	No	Yes	Yes
Non-linear style controls		No	No	No	Yes

Table IV: Connected Comovement – Cross-sectional Variation

This table reports Fama-MacBeth estimates of monthly cross-sectional regressions forecasting the correlation of daily Fama-French-Carhart residuals in month $t + 1$ for the sample of stocks defined in Table I. We add interactions between *FCAP** and the total float of the pair (*PFLOAT)* and the *absolute* value of the total flows into the common funds holding the pair (*PFLOW)*. We also include a triple interaction among these three variables. *PFLOAT** is first rank transformed and then normalized to have unit standard deviation, which we denote with an asterisk super-script. We calculate Newey-West standard errors (four lags) of the Fama-MacBeth estimates that take into account autocorrelation in the cross-sectional slopes. The analysis is limited to the 1991-2008 sub-period when *PFLOAT* becomes available. Controls not shown here are reported in Table AV in the Internet Appendix.

	Dep. Variable: Correlation of four-factor residuals			
	(1)	(2)	(3)	(4)
Intercept	0.00426	0.00454	0.00084	0.00101
	(6.74)	(6.51)	(1.19)	(1.51)
*FCAP**	0.00835	0.00428	0.00313	0.00338
	(15.95)	(12.59)	(10.86)	(12.17)
FCAP x PFLOAT**	-0.00038	-0.00016	-0.00043	-0.00038
	(-1.64)	(-0.88)	(-2.44)	(-2.10)
FCAP x PFLOW*	0.00407	0.00244	0.00253	0.00248
	(8.18)	(6.17)	(7.00)	(6.90)
FCAP x PFLOAT* x PFLOW*	-0.00051	-0.00028	-0.00050	-0.00046
	(-2.15)	(-1.39)	(-2.58)	(-2.41)
Additional Controls (reported in the Internet Appendix)				
Non-linear size controls	No	Yes	Yes	Yes
Pair characteristic controls	No	No	Yes	Yes
Non-linear style controls	No	No	No	Yes

Table V: Alphas on Connected Stocks Trading Strategies

This table presents the profitability of a simple trading strategy exploiting stock connectedness. We independently sort stocks into quintiles based on their own return over the last three months and the return on their connected portfolio over the last three months. We first measure the degree of abnormal connections by orthogonalizing FCAP* (which we denote by FCAPO*) each period to the full set of controls in regression (4) of Table II. We then define the connected return as $r_{iC,t} = \sum_{j=1}^{J} FCAP_{ij,t-1}^{**} r_{j,t} / \sum_{j=1}^{J} FCAP_{ij,t-1}^{**}$ where $FCAP_{ij,t}^{**} = rank(FCAPO_{ij,t}^{*})$ if $FCAP_{ij,t} > 0$ and $FCAP_{ij,t}^{**} = 0$ if $FCAP_{ij,t} = 0$. Following Jegadeesh and Titman (1993), each composite portfolio below is an equal-weight average of the corresponding simple strategies initiated one to five months prior. The table reports the five-factor alphas on these 25 composite portfolios. The five factors include the four Fama-French/Carhart factors plus a short-term reversal factor, all downloaded from Ken French's website. We also report the average returns on a connected stocks trading strategy (CS) which buys the low own return / low connected return composite portfolio and sells the high own return / high connected return composite portfolio.

Five factor alphas

		Connected portfolio return						
		Low	2	3	4	High	L - H	Avg L - H
	Low	0.0055	0.0039	0.0039	0.0028	0.0002	0.0053	
		(4.58)	(3.73)	(3.79)	(2.56)	(0.15)	(3.57)	
	2	0.0047	0.0041	0.0027	0.0024	0.0008	0.0039	
Own		(4.59)	(4.73)	(3.29)	(2.55)	(0.84)	(3.68)	
Return	3	0.0031	0.0024	0.0014	0.0011	-0.0005	0.0036	0.0036
		(3.33)	(2.87)	(1.58)	(1.25)	(-.56)	(3.56)	(4.13)
	4	0.0019	0.0007	0.0002	-0.0004	-0.0011	0.0030	
		(2.10)	(0.78)	(0.31)	(-.47)	(-1.3)	(3.18)	
	High	0.0003	0.0002	-0.0017	-0.0017	-0.0022	0.0025	
		(0.27)	(0.30)	(-2.0)	(-2.0)	(-2.4)	(1.90)	
	L - H	0.0052	0.0036	0.0056	0.0045	0.0023	**0.0076**	**LL–HH (*CS*)**
		(3.74)	(3.08)	(4.48)	(3.54)	(1.78)	**(4.96)**	

Table VI: The Connected Stocks Trading Strategy and Liquidity Risk

This table measures the loadings of the connected stock trading strategy on a liquidity factor as well as on trend and quarter dummies. We study the connected strategy, *CS*, formed in Table V, which buys the low own return / low connected return composite portfolio and sells the high own return / high connected return composite portfolio. We regress the return on this trading strategy on a constant, the liquidity factor (*PS_INNOV*) from the work of Pastor and Stambaugh (2003), the Fama-French/Carhart factors, a short-term reversal factor (*STR*), a trend, and seasonal (quarterly) dummies over the period June 1980 to December 2008. Column 1 reports loadings of the return on our connected strategy on the five factors used in Table V. Column 2 adds PS_INNOV as an explanatory variable. Columns 3 and 4 include a trend or quarterly seasonal dummies respectively as additional explanatory variables.

Dependent Variable: Connected Stock Trading Strategy Return				
	1	2	3	4
Alpha	0.0076	0.0076	0.0076	0.0147
	(4.96)	(4.99)	(4.94)	(5.08)
PS_INNOV		0.0433		
		(1.73)		
RMRF	-0.0623	-0.0832	-0.0597	-0.0654
	(-1.68)	(-2.14)	(-1.60)	(-1.77)
SMB	-0.2828	-0.2861	-0.2831	-0.2695
	(-5.73)	(-5.81)	(-5.73)	(-5.44)
HML	-0.1502	-0.1595	-0.1469	-0.1343
	(-2.77)	(-2.93)	(-2.69)	(-2.47)
UMD	-0.8565	-0.8545	-0.8561	-0.8584
	(-23.4)	(-23.4)	(-23.3)	(-23.6)
STR	0.0327	0.0368	0.0329	0.0375
	(0.70)	(0.79)	(0.70)	(0.81)
Trend			0.0000	
			(0.61)	
Q1				-0.0071
				(-1.71)
Q2				-0.0132
				(-3.22)
Q3				-0.0086
				(-2.09)
Obs	341	341	341	341
R2	67%	68%	67%	68%

Table VII: Hedge Fund and Mutual Fund Exposure to Connected Strategy

This table measures the exposure of two CSFB hedge fund return indexes (all and long/short) as well as the value-weight average active mutual fund return (net of fees) to the connected strategy described in Table V. We regress fund index returns in excess of the Treasury bill rate on a constant, the connected strategy and either the eight Fung and Hsieh (2001, 2004) hedge fund factors or the Fama-French/Carhart model plus a short-term reversal factor (*STR*). The time period is January 1994 to December 2008. The analysis adds as additional explanatory variables the return on the connected stocks trading strategy (*CS*) of Table V as well as an interaction between *CS* and the lagged, demeaned, normalized change in the VIX (*ΔVIX*).

	HF ALL		HF LONG/SHORT		MF ALL
Alpha	0.0025	0.0030	0.0029	0.0027	-0.0009
	(1.93)	(2.63)	(2.91)	(2.42)	(-2.45)
CS	-0.0490	-0.1306	-0.0989	-0.2245	-0.0225
	(-1.13)	(-5.38)	(-2.88)	(-9.25)	(-1.78)
CS * ΔVIX_{t-1}	-0.0070	-0.0034	-0.0140	-0.0124	-0.0018
	(-1.30)	(-0.67)	(-3.29)	(-2.49)	(-1.18)
RMRF	0.3328		0.4658		0.9518
	(10.85)		(19.17)		(106.5)
SMB	0.0794		0.1345		0.0553
	(2.02)		(4.33)		(4.84)
HML	0.0236		-0.1073		-0.0273
	(0.58)		(-3.35)		(-2.31)
UMD	0.1026		0.1353		0.0057
	(2.28)		(3.80)		(0.44)
STR	-0.0428		-0.0580		-0.0189
	(-1.30)		(-2.23)		(-1.97)
Bond-trend		-0.0182		-0.0041	
		(-2.28)		(-0.51)	
Currency-trend		0.0131		0.0066	
		(2.09)		(1.07)	
Commodity-trend		0.0112		-0.0011	
		(1.26)		(-0.13)	
Equity Market		0.1578		0.3423	
		(4.12)		(8.96)	
Size Spread		0.0499		0.1825	
		(1.39)		(5.10)	
Bond Market		-0.1531		-0.0551	
		(-4.10)		(-1.48)	
Credit Spread		-0.2298		-0.0669	
		(-4.27)		(-1.25)	
Emerging Market		0.0903		0.0929	
		(3.61)		(3.72)	
Obs	182	182	182	182	182
RSquare	52%	59%	82%	75%	99%

Figure 1: Cumulative Alphas for a Connected Stock Strategy

This figure graphs the abnormal buy-and-hold performance of a cross-stock reversal trading strategy that exploits information in ownership connections among big stocks (stocks above the median NYSE market capitalization). These stocks are sorted into 25 portfolios based on independent quintile sorts on their own three-month return and the three-month return on their connected stock portfolio. We define a stock's connected portfolio as those stocks each quarter that have active mutual fund owners in common with the stock in question. The connected portfolio weights are proportional to the degree of abnormal common fund ownership, as defined in the text. The figure plots the buy-and-hold abnormal returns on stocks that are in the low own-return and low connected-return portfolio (which we dub the *low* portfolio), stocks that are in the high own-return and high connected-return portfolio (the *high* portfolio), and the difference between *low* and *high* abnormal returns. Returns are bench-marked against the Fama-French/Carhart four-factor model augmented with the short-term reversal factor.

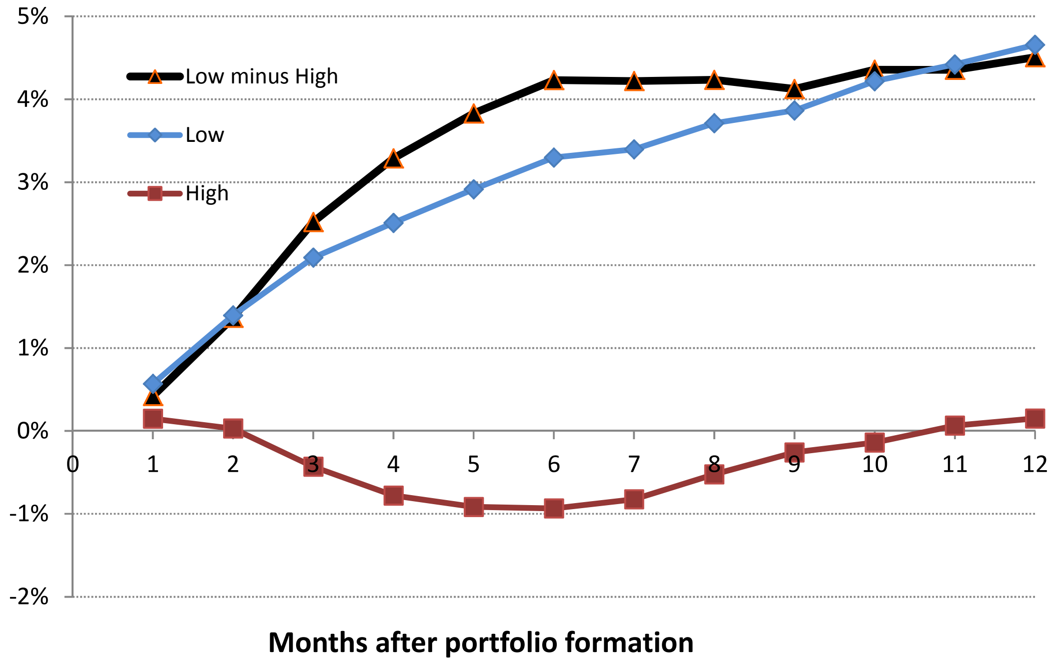
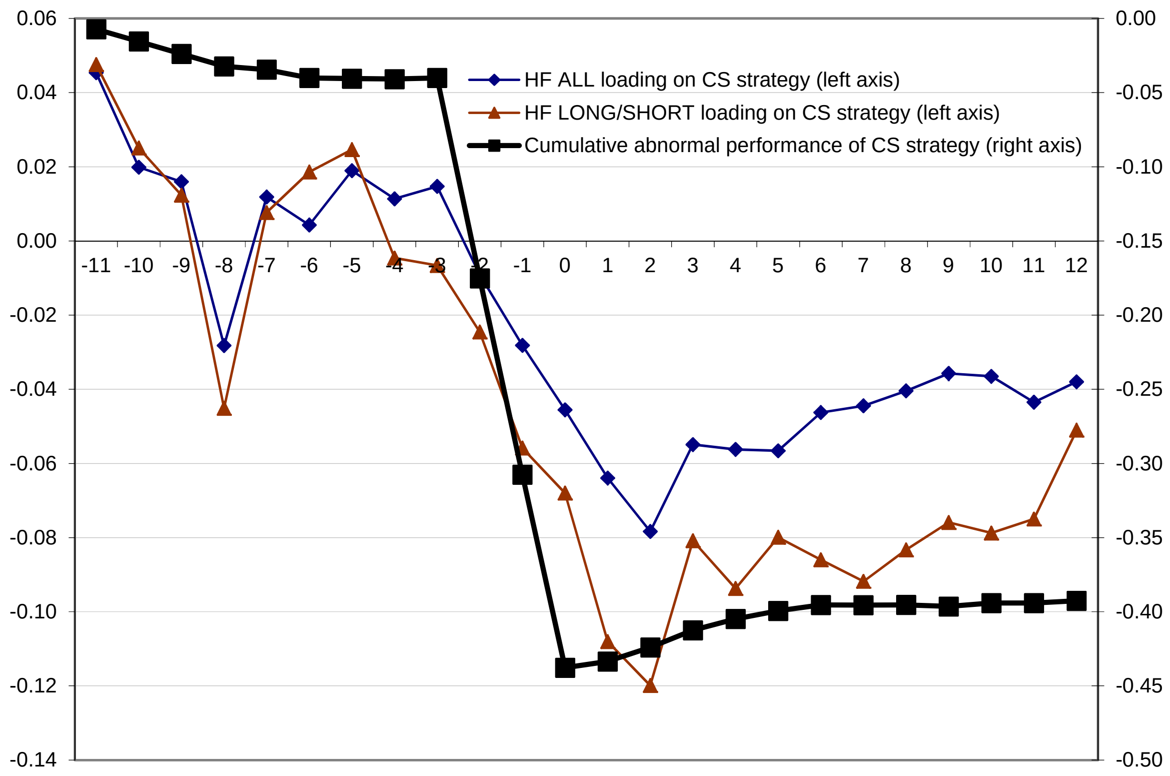


Figure 2: Hedge Fund Sensitivity to the Connected Stocks Strategy Return

This figure plots the loadings of hedge fund index returns on the connected stocks trading strategy, *CS*, in event time as well as the corresponding cumulative five-factor abnormal return on the *CS* strategy.



FOOTNOTES

---

[1] For example, Vijh (1994) and Barberis, Shleifer and Wurgler (2005) find that S&P 500 index additions are followed by an increase in return covariation. However, members of the S&P500 selection committee might implicitly look to include firms in the index that, going forward, are central to the economy. Intuitively, that implicit criterion might be plausibly related to the sensitivity of a firm's fundamentals to the fundamentals of other central firms. Indeed, Anton (2011) provides evidence that this is the case. Similarly, Greenwood and Thesmar (2011) show that their measure of co-fragility (based on funds' correlated flows) forecasts subsequent comovement. However, funds with correlated flows might choose to invest in similar stocks whose fundamentals comove. Indeed, Greenwood and Thesmar (2011) grant that "concerns about endogeneity loom very large" and thus their results are only suggestive.

[2] Kisin (2011) uses this scandal to explore determinants of firm investment.

[3] Researchers must make several assumptions to back out flows from data on funds' total net assets (see, for example, Chevalier and Ellison 1997, Sirri and Tufano 1998, and Lou 2012). These assumptions include that inflows and outflows occur at the end of each period (often as infrequently as quarterly), and that investors reinvest their dividends and capital appreciation distributions in the same fund. Lack of data on flow activity during fund initiation, merger, and liquidation is also an issue. Furthermore, the relation between flows and stock-level price impact is complicated by the fact that funds can absorb capital flows using cash buffers and can initiate, expand, or liquidate individual positions to different degrees based on liquidity costs.

[4] As our focus is on the economics of how connectedness causes excess comovement that, by definition, must eventually revert, we do not analyse whether our trading strategy is costly to implement. Though the strategy is restricted to big stocks and the predictability persists for six months, we grant that the

relatively high turnover of the strategy will make the performance after transaction costs less attractive.

[5] See, for example, King and Wadhwani (1990), King, Sentana, and Wadhwani (1994), Forbes and Rigobon (2001), Rigobon (2002), and Bekaert, Harvey, and Ng (2005).

[6] Similarly, Singleton (2011) examines the influence of investor flows on returns in the crude-oil futures market.

[7] See, for example, the 2010 AFA presidential address, Duffie (2010).

[8] Many researchers have built on the ideas in Shleifer and Vishny (1997), including Gromb and Vayanos (2002), Vayanos (2004), and Brunnermeier and Pedersen (2009). For a recent survey of this literature, see Gromb and Vayanos (2010).

[9] Sun (2007) uses standard clustering techniques to identify subsets of funds that hold similar stocks. Sun shows that the typical stock's return covaries with the equal-weight average return on all of the stocks in the top five fund clusters holding the stock in question. Moreover, Sun shows that this covariance is stronger if the average flow for the top five clusters in question is lower than the tenth percentile of the historical distribution of fund flows for that group of five fund clusters. In contrast, our approach predicts the pair-specific return correlation with the degree of the pair's common ownership, controlling for exposure to systematic return factors, style and sector similarity, and many other pair characteristics. Additionally, Sun does not examine any implications of the covariance she documents for profitable trading strategies.

[10] In a similar fashion, Lou and Polk (2013) propose using the past degree of abnormal return correlation among those stocks that an arbitrageur would speculate on as a measure of arbitrage activity.

Their approach is successful in linking time variation in both the profitability and subsequent reversal of momentum strategy returns to their measure of momentum arbitrage activity.

[11] Choosing this particular cutoff is broadly consistent with the empirical literature. See, for example, Fama and French (2008).

[12] We specifically follow the algorithm described in Cremers and Petajisto (2009).

[13] This choice is consistent with the degree of autocorrelation generally observed in the time series of coefficients. In particular, the partial serial correlation in the time series of regression coefficients associated with FCAP becomes statistically insignificant after four months. Moreover, we have re-estimated our standard errors using lag lengths as long as 36 months, and our results remain statistically significant.

[14] We obtain the Fama and French (1993) and Carhart (1997) daily returns factors from Ken French's website.

[15] We base our measure of common coverage on annual forecasts as quarterly earnings forecasts are not issued as consistently.

[16] One may worry that we miss some portion of excess comovement between returns due to non-synchronicity. Our results become slightly stronger if we include lags when computing the correlation of the daily four-factor residuals.

[17] Of course, a portion of the covariation linked to a pair's exposure to systematic risk factors could be caused by common ownership. If we instead forecast the correlation of raw returns (i.e. not four-factor residuals), the coefficient is 0.02118, more than five times as large.

[18] To calculate the average range in abnormal correlation linked to *FCAP**, we first orthogonalize *FCAP** to the other variables. We then forecast four-factor residual correlations using orthogonalized *FCAP**, saving the minimum and maximum forecast in each cross section. Finally, we average this range over time.

[19] Though the variation linked to *FCAP** may seem modest, note that predicting realized abnormal return correlation is a difficult task. Realized abnormal return correlation is very noisy, with an average standard deviation in each cross section of approximately 0.25. R^2's are less than 2% on average, even in the fourth specification which contains very flexible style controls.

[20] We only examine in-sample forecasts of cross-sectional variation in comovement. Since DeMiguel, Garlappi, and Uppal (2007) argue that 1/N rules perform better out-of-sample than traditional forecasts of the covariance matrix, it would be interesting to explore how our method and our characteristics perform out-of-sample. We leave this question for future research.

[21] We add the direct effect of *PFLOAT** and *PFLOW* as well as the interaction between *PFLOAT** and *PFLOW* to the control variables to ensure that any triple interaction effect we find is not simply because of an omitted variable. For the sake of brevity, the table only reports the coefficients on FCAP and its interactions. All other coefficient estimates are reported in Table AV of the Internet Appendix.

[22] Of course, these conclusions depend on having correctly specified the VAR producing the return decomposition. The Internet Appendix describes the methodology. Table AX presents the VAR estimates.

[23] Alternatively, we could have found that the majority of the variation in comovement linked to connectedness is entirely because of cash-flow news covariance. If so, this finding would be incon-

sistent with institutional ownership causing comovement as stock owners cannot directly affect cash flows.

[24] Chen, Chen, and Li (2012) also measure the cross-sectional variation in pair-wise correlations and show that a large portion of that cross-sectional variation is persistent, yet unexplained by a long list of variables. They do not incorporate holdings data in their forecasts of correlation. Chen, Chen, and Li also develop a trading strategy. However, their focus is on enhancing the profitability of pairs trading strategies that look for divergence in the prices of a pair of similar stocks. The strategy then profits from the correction of the divergence. In contrast, our strategy identifies stocks that temporarily move together and profits from their eventual divergence.

[25] We thank an anonymous referee for this suggestion.

[26] The short-term reversal factor is also from Ken French's website.

[27] We find that including *STR* in our benchmarking has very little effect on our results.

[28] If we exclude *STR* from the factor regression, the alpha is 78 basis points and the corresponding *t*-statistic is 5.10. We have used FCAP* instead of FCAPO* to determine the weights on a stock's connected portfolio. The results we find (reported in the Internet Appendix, Table AVII) are weaker but are still economically large, generating 48 basis points per month with a corresponding *t*-statistic of 2.76. This finding shows the robustness of our result, as well as the usefulness of measuring abnormal common ownership. We have also examined different lengths of the formation period in which we measure the connected and own return; our results are robust to formation periods ranging from one month (45 basis points, *t*-statistic 3.77) to 12 months (32 basis points, *t*-statistic 1.56).

[29] The *FCAP**-weighted version of this strategy earns 30 basis points per month (*t*-statistic of 2.72).

[30] Since this factor is not a traded portfolio, the loading of *CS* on the liquidity factor is the only object of interest in this regression.

[31] Note that the CFSB does not include managed accounts or funds of funds in its indexes.

[32] Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009) predict that increasing volatility tightens the funding constraints of liquidity providers like hedge funds. Indeed, hedge funds face redemptions and lower leverage when the VIX is increasing (Ang, Gorovyy, and van Inwegen 2011 and Ben-David, Franzoni, and Moussawi 2012).

[33] We downloaded three of the Fung and Hsieh (2001) factors from http://faculty.fuqua.duke.edu/~dah7/DataLibrary/TF-FAC.xls.

[34] In the Internet Appendix, Table AVIII repeats the analysis when the strategy is computed using *FCAP** instead of *FCAPO**. We find results that are qualitatively similar.

[35] In somewhat related work, Lou and Polk (2013) find that the typical long-short equity hedge fund is able to time the momentum factor, but this timing ability is decreasing in the size of the fund's assets under management.