3-2011

Pairwise correlations

Tarun CHORDIA

Amit GOYAL

Qing TONG
Singapore Management University, qingtong@smu.edu.sg

Pairwise Correlations

Tarun Chordia **Amit Goyal** **Qing Tong***

April 2011

Abstract

Pairwise stock correlations increase by 27% on average when stock returns are negative. It is trading activity in small stocks that leads to higher correlations when returns are negative. We provide evidence consistent with the hypothesis that co-ordinated selling by retail investors drives this asymmetry in correlations. The co-ordinated selling activity by retail investors is triggered by negative market returns.

*Tarun Chordia is from Goizueta Business School, Emory University, email: Tarun_Chordia@bus.emory.edu, Tel: 404-727-1620; Amit Goyal is from Swisss Finance Institute at the University of Lausanne, email: Amit.Goyal@unil.ch, Tel: +41-21-692-3676; and Qing Tong is from Lee Kong Chian School of Business, Singapore Management University, email: qingtong@smu.edu.sg, Tel: +65-6828-0763. We thank Clifton Green, Narayanan Jayaraman, Narasimhan Jegadeesh, Ernst Maug, Stefan Ruenzi, Richard Sias, Jay Shanken and seminar participants at Emory University, Georgia Institute of Technology, University of Mannheim and Virginia Tech for helpful comments. All errors are our own.

1 Introduction

Correlations have been much studied. A number of studies have documented that correlations between asset returns are higher when prices fall than when prices rise. Longin and Solnik (2001) document that international markets (UK, France, Germany and Japan) have higher correlation with the US market when prices fall in the US market. Ang and Chen (2002) find strong evidence of asymmetric correlations between the US market and stock portfolios formed by sorting on size, book-to-market ratio, past returns and industry. Bae, Karolyi, and Stulz (2003) study contagion in emerging markets. This asymmetry in correlations has important implications for portfolio allocation because diversification may be compromised if asset values decline simultaneously. Ang and Bekaert (2000) model the dynamic portfolio choice problem of a US investor with time-varying investment opportunities as a regime switching process where correlations and volatilities increase in bad times. Hong, Tu and Zhou (2007) show that incorporating the asymmetry in correlations into portfolio decisions can add substantial economic value.

We study correlations between individual stocks. The first contribution of this paper is to study the properties of pairwise correlations at the individual stock level. Prior work (with the exception of Andersen, Bollerslev, Diebold and Ebens (2001) who study daily correlations for a short sample of the thirty stocks in the Dow Jones Industrial Average) has focused on portfolios. We, instead, work directly with individual pairwise correlations which are the building blocks for portfolio correlations.

We study correlations at a monthly frequency for all common stocks listed on the New York Stock Exchange (NYSE). Using daily returns, we calculate monthly pairwise return correlations from January 1963 through December 2008. Market capitalization, beta, book-to-market ratio, turnover, volatility and past returns have a significant impact on the pairwise correlations in the cross-section. Moreover, firms that are in the S&P 500 index, firms in the same industry and firms that are close in terms of stock prices exhibit higher correlations. These facts have important implications for portfolio diversification purposes.

We also document the asymmetry in correlations at the individual stock level similar to the existing evidence of asymmetry at portfolio level. In the time-series, we find a strong asymmetric correlation effect at the individual stock level: negative returns have a statistically significant impact on stock correlations. Stock correlations increase by 27% on average when returns are negative whereas positive returns have no significant impact on correlations.

The second and main contribution of this paper is to search for an explanation for the asymmetry in correlations at the individual stock level. Prior research has focused on statistical tests that can identify asymmetric correlations and on the question of whether particular models can account for the correlation asymmetries in the data (see, for example, Longin and Solnik (2001), Ang and Chen (2002), Bae, Karolyi, and Stulz (2003) and Hong, Tu and Zhou (2007)). In this article, our main goal is to explain the asymmetric correlation phenomenon from a trading perspective. We use turnover (monthly share trading volume divided by shares outstanding) to proxy for trading activity. Our empirical evidence shows that the asymmetric correlation effect is attributable to trading activity when returns are negative. Specifically, after we interact turnover with returns we find that the positive coefficient in the regression of correlation on negative returns is entirely captured by the interaction of return with the trading activity. In other words, it is an increase in trading volume when returns are negative that leads to an increase in correlations.

However, there is a caveat in the above result. Trading activity accounts for the asymmetry in correlations only for small stocks and/or for stocks that have low institutional holdings. For large stocks and/or for stocks that have high institutional holdings, trading activity does not capture the asymmetry in correlations. This dichotomy occurs because a larger fraction of the correlation in returns of small stocks is caused by the unsystematic component of returns. We use the market model to compute the two components of returns, the systematic component and the residuals. Conditional on negative returns, the idiosyncratic component of returns accounts for 17.3% of the correlation in small stocks and only 8.2% in large stocks. This raises the possibility that it is trading by retail investors (who predominantly hold the small stocks) that leads to higher correlations in the residuals of small stocks.

We examine in detail the possibility that trading by small retail investors causes the asymmetry in correlations. Our conjecture is that it is trading by retail investors in the face of declining prices that causes the asymmetry in correlations. Our conjecture is supported by several pieces of evidence using the transactions data. Using the NYSE Trade and Quote (TAQ) dataset from 1993 to 2008, we examine the impact of trade size on correlations between returns and order imbalances.

First, we find that it is not trading per se, but rather small trades, presumably from retail investors,[1] that drive the asymmetry. Moreover, small trades have more impact on the correlations between small firms or firms with low institutional holdings.

[1]Barber, Odean, and Zhu (2009) show that small trades are more likely to come from retail investors.

Second, we show that co-ordinated trading activity is strong for retail investors when prices are going down. We calculate order imbalance correlations for each pair of stocks and group the correlations into different categories based on trade sizes and returns. We find that the order flow correlations are highest when returns are negative and when the trade sizes are small. Moreover, the average pairwise correlation between the order imbalances for small trades (trades less than $10,000) are higher than those for large trades, especially when the returns are negative. We also document that average small (large) order flows are negative (positive) when the returns are negative suggesting that retail investors sell in months when the pairwise returns are negative. The results suggest that herding by small retail investors leads to the asymmetry in correlations.

The question then arises: What causes this herding in small trades? We show that herding is triggered by negative market returns. More specifically, when the first five days of the month experience negative market returns, then there are more sell trades in the small stocks over the remaining days of the month. Moreover, in the small stocks, the correlation in the order flow of small trades is larger than that of the large trades, regardless of how the order flow is measured, by number of trades or by number of shares traded. This is consistent with the hypothesis that small retail investors herd in the face of declining prices possibly triggered by margin calls or stop loss orders and their co-ordinated sell trades leads to the asymmetry in correlations.

A growing literature examines the trading behavior of retail investors. Hvidkjaer (2006) suggests that momentum could partly be driven by the behavior of small traders. Barber, Odean, and Zhu (2009), Hvidkjaer (2008) and Kaniel, Saar, and Titman (2008) study the relation between small trades and stock returns. Kumar and Lee (2006) show that systematic retail trading explains return comovements for stocks with high retail concentration. In line with the studies of the impact of retail investors, our results show that asymmetric correlations could be attributable to the trading activity of retail investors, who seem to simultaneously sell stocks in the face of declining prices.

Our paper is related to a number of studies that examine asymmetries of asset returns. French, Schwert, and Stambaugh (1987), Schwert (1989), Campbell and Hentschel (1992), Bekaert and Wu (2000) and Avramov, Chordia, and Goyal (2006), among others, document asymmetries in volatilities. Other asymmetric phenomena include the asymmetric betas and the asymmetric covariances (see, for instances, Ball and Kothari (1989), Conrad, Gultekin, and Kaul (1991), Cho and Engle (2000), Bekaert and Wu (2000) and Ang, Chen, and Xing (2006)); and asymmetries in higher moments (Harvey and Siddique (2000)). This paper is most closely related to Avramov, Chordia, and Goyal (2006), who study the daily volatility at the individual stock level and provide a trading-based explanation for the asymmetric effect in daily volatility.

The remainder of this paper is organized as follows. Section 2 describes the return correlation data and presents the summary statistics. Section 3 provides the main results. We discuss the sources of asymmetry in correlation in Section 4 and conclude in Section 5.

2 Data

Each month from January 1963 to December 2008, we use daily stock returns from CRSP to compute the monthly pairwise return correlations. We assume a zero mean expected return at the daily frequency to calculate the correlations. We require at least 15 pairwise returns each month and exclude daily returns exceeding 500%. We also exclude stocks priced less than \$5.[2] The sample spans the period January 1963 through December 2008. To keep the sample size manageable, we calculate return correlations only amongst common stocks listed on the New York Stock Exchange (NYSE). The total number of different stocks on the NYSE over the 552-month sample period is 4,792. The average number of stocks is 1,432 per month. This process yields over 0.56 billion pairwise monthly correlations over the sample period.

2.1 Summary statistics

Table 1 reports the summary statistics (means, medians, and standad deviations) of the monthly correlations. Panels A, B, and C provide these statistics for all time periods, expansionary periods, and recessionary periods as defined by the National Bureau of Economic Research (NBER), respectively. Panels D and E report these statistics for subsamples when both returns are positive and negative, respectively.

We also examine the relation between correlation and firms size by dividing the stocks into large size group (larger than the median) and small size group (smaller than the median) based on firm market capitalization at the end of previous month. The second column in Table 1 provides the summary results for all stock pairwise correlations without breaking the stocks down into small size and large size group. The third column presents the summary results for pairwise correlations between small stocks (when both stocks belong to the small size group, Small/Small). The fourth column depicts the summary results for pairwise correlations between large stocks (when both stocks in the

[2] We note that our results are essentially the same when (a) we use the daily sample mean as a measure of expected return in correlation calculation, and/or (b) we do not exclude low priced stocks.

large size group, Large/Large). The last column has the summary results for stocks in different groups, Small/Large.

During our sample period, the average pairwise correlation is 0.129, the median is 0.128 and the standard deviation is 0.234. Average correlation is higher during recessions than that during expansions. Also, the average correlation within the large size group is substantially larger that within the small size group. For instance, the average overall correlation is 0.172 for Large/Large category, compared to 0.095 for Small/Small category. The table also shows that the average standard deviation of correlations is about 0.23 (0.24) for small (large) firms.

We plot the equal- and value-weighted correlations in Figure 1, to see the evolution of correlations. The value-weighted correlations at time t is defined as:

$$\frac{\sum_{i<j} \mathrm{Cor}_{ijt} \times \mathrm{Size}_{it-1} \times \mathrm{Size}_{jt-1}}{\sum_{i<j} \mathrm{Size}_{it-1} \times \mathrm{Size}_{jt-1}} \, ,$$

where Cor_{ijt} is the correlation between returns of stocks i and j in month t. Since large firms tend to have higher correlations than smaller firms, value-weighted correlations are higher than the equally weighted correlations in Figure 1. The figure also shows that correlations are high during recessions in general (NBER-dated recessions are highlighted in the graph). Finally, the figure shows the huge spike in correlation during the market crash of October 1987.

2.2 Cross-sectional determinants of correlations

The summary statistics point to market capitalization as being an important determinant of correlation. In this section, we expand on this by studying other cross-sectional determinants of the pairwise correlations. Our approach is to run the following Fama and MacBeth (1973) cross-sectional regressions of correlations on firm characteristics as follows:

$$\mathrm{Cor}_{ijt} \;=\; \gamma_{0t} + \gamma_{1t}\left(\frac{X_{it} + X_{jt}}{2}\right) + u_{ijt} \, , \tag{1}$$

where Cor_{ijt} is the correlation of firms i and j in month t, and X's are firm characteristics. Table 2 provides the results of these cross-sectional regressions. We report the time-series average of the monthly regression coefficients, and their t-statistics, corrected for autocorrelation in the time-series estimates.

We examine three different regression specifications. In the first specification, we use only stocks' market betas as the explanatory variable. The motivation for using the betas is that part of the correlation between stocks is due to their exposures to common factors. For example, in the market model, correlations and βs have the following relation:

$$\mathrm{Cor}_{ij} = \frac{\beta_i \beta_j \sigma_m^2 + \sigma_{\epsilon_i,\epsilon_j}}{\sqrt{\beta_i^2 \sigma_m^2 + \sigma_{\epsilon_i}^2}\sqrt{\beta_j^2 \sigma_m^2 + \sigma_{\epsilon_j}^2}} , \tag{2}$$

where σ_m is the market volatility, ϵ's are market model residuals with volatility σ_ϵ, and $\sigma_{\epsilon_i,\epsilon_j}$ is the covariance between market model residuals for stocks i and j. We follow Fama and French (1992) in estimating the individual stock betas. Each month we assign firms to the 10 $\times$ 10 size-β portfolios and calculate the equal-weighted monthly post-ranking portfolio returns for the next month. This results in a time series of post-ranking returns for each of the 100 portfolios. The post-ranking β is calculated by regressing the portfolio return on the market return. The post-ranking βs of the size-β portfolios are then assigned to each stock in the portfolio. Table 2 shows the coefficient on the average β's is positive and statistically significant. In other words, correlations between two stocks increase with the average of their betas. However, the average $\overline{R}^2$ of the regressions is very modest at 1.0%.

We add more characteristics in the second regression specification in Table 2. We use (logarithms of) market capitalization and book-to-market, and indicator variables for the sign of stock returns over the past month and past year. The results show that size and book-to-market have significant positive impact on correlations. That is, correlations tend to be high for large firms and value firms. Moreover, the coefficient on $\mathrm{I}(R_i < 0, R_j < 0)$ is 0.011, indicating firms with negative returns have higher correlations than those with positive returns. Finally, correlation between stock returns are higher for firms that have experienced negative returns over the past year.

In the third specification we add yet more firm characteristics. Specifically, we add (logarithms of) turnover (monthly share trading volume divided by shares outstanding) and monthly return volatility. Our results show that the coefficients on turnover and volatility are significantly positive. Green and Hwang (2009) document that return comovement is related to stock price. They find that stocks with similar prices experience more comovement with each other. We use an indicator variable that is equal to one when $|\mathrm{Price}_i - \mathrm{Price}_j| / \max(\mathrm{Price}_i, \mathrm{Price}_j) < 25\%$ to indicate that the two stocks have similar prices. Consistent with the findings of Green and Hwang, our results show that firms with similar prices have higher correlations. We also partition securities into six groups based on the median NYSE market equity breakpoint and the 30th and 70th book-to-market NYSE percentile breakpoints following Fama and French (1993). We find that stocks within the same size/book-to-market group have higher correlations. This specification also shows that firms exhibit higher correlation when they are in the S&P 500 index, which is consistent with the findings of Barberis, Shleifer and Wurgler (2005), who show that stocks added to the S&P 500 index begin to covary more with other members of the index. Finally, we use Fama-French 17 industry classification to document that firms in the same industry have higher correlations. In terms of the model specification, the average $\overline{R}^2$ increases from 1.0% in the first regression to 9.3% in the third regression.

Thus, from a diversification perspective it is better to invest in small, low beta, low turnover, low volatility growth stocks that have large differences in prices and that do not belong to the same industry or the S&P 500 index or the same size and book-to-market group.

3 Results

Our main objective in the paper is to study asymmetric correlations at the individual stock level. Our model specification is similar to that of Andersen et al. (2001).[3] We run the following time series regressions for each pairwise correlation:

$$\begin{aligned}\mathrm{Cor}_{ijt} &= \delta_0 + \delta_1 I(R_{it} * R_{jt} > 0) + \delta_2 I(R_{it} * R_{jt} > 0) * \mathrm{Turnover}_{ijt} \\ &+ \delta_3 I(R_{it} < 0, R_{jt} < 0) + \delta_4 I(R_{it} < 0, R_{jt} < 0) * \mathrm{Turnover}_{ijt} \\ &+ \delta_5 \mathrm{Turnover}_{ijt} + \delta_6 \mathrm{Cor}_{ijt-1} + \delta_7 \mathrm{Vol}_{ijt} + \delta_8 \mathrm{Term}_{t-1} + \delta_9 \mathrm{Tbl}_{t-1} \\ &+ \delta_{10} \mathrm{Def}_{t-1} + \delta_{11} I(\mathrm{SUE}_{it} > 0, \mathrm{SUE}_{jt} > 0) + \delta_{12} I(\mathrm{SUE}_{it} < 0, \mathrm{SUE}_{jt} < 0) + \eta_{ijt} \,,\end{aligned} \tag{3}$$

where Cor_{ijt} is the correlation between stocks i and j in month t, R is the return, $\mathrm{Turnover}_{ijt}$ is the average of turnover of the two stocks, and Vol_{ijt} is the average of the two stock volatilities. $I(\cdot)$ is the indicator function. Term, Tbl and Def denote the term spread, the T-bill yield and the default spread, respectively. The term spread is the yield differential between Treasury bonds with more than ten years to maturity and T-bills that mature in three months. The default spread is the yield differential between bonds rated BAA and AAA by Moody's. SUE is the standardized unexpected

[3]Unlike the daily time-series regressions in Andersen et al. (2001), our regressions are at a monthly frequency with one lag of the dependent variable. At the monthly frequency, correlations are not as persistent as at the daily frequency. Hence, we do not use the fractionally integrated regression specification of Andersen et al.

earnings, computed as the most recently announced quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. The impact of positive returns on correlations is δ_1 and the impact of negative returns is $\delta_1 + \delta_3$. Thus, a positive δ_3 shows asymmetric correlation since it gives the additional influence when returns are negative.[4]

We require a minimum of seventy-two observations for each regression, which results in roughly 3 million regressions for the 1963-2008 sample period. Table 3 reports the (cross-sectional) means of the parameter estimates and the t-statistic for the mean estimate. Finally, we also report the percentage of parameter estimates with t-statistics less than -1.96 and greater than 1.96.

The computation of standard error for the mean estimate is a non-trivial exercise as the estimators in equation (3) are cross-sectionally correlated. The regression equation can be written as follows: $y_m = X_m\beta_m + \eta_m$, where $m = 1, \cdots, M$, and M is the total number of regressions. Coefficient estimate from the mth regression is $\widehat{\beta}_m$ and the average coefficient estimate that we report is $\widehat{\widehat{\beta}} = \frac{1}{M}\sum_m \widehat{\beta}_m$. Then, the variance of this average estimate is given by:

$$\text{Var}(\widehat{\widehat{\beta}}) = \frac{1}{M^2}\left[\sum_{m=1}^{M} \text{Var}(\widehat{\beta_m}) + \sum_{m=1}^{M}\sum_{n=1,n\neq m}^{M} \text{Cov}(\widehat{\beta_m}, \widehat{\beta_n})\right],$$

where $\text{Var}(\widehat{\beta_m}) = \frac{(\widehat{\eta'_m\eta_m})}{(T-k)}(X'_mX_m)^{-1}$ and $\text{Cov}(\widehat{\beta_m}, \widehat{\beta_n}) = \frac{(\widehat{\eta'_m\eta_n})}{(T-k)}(X'_mX_m)^{-1}(X'_mX_n)(X'_nX_n)^{-1}$. The computation of the entire covariance matrix $\text{Cov}(\widehat{\beta_m}, \widehat{\beta_n})$ as proposed by Jones, Kaul, and Lipson (1994) is infeasible. Given the three million regressions, the order of magnitude for the computation of the entire covariance matrix is 10^{12}. We simplify the task by using the following identity:

$$\text{Var}(\widehat{\widehat{\beta}}) = \frac{1}{M^2}\left[\sum_{m=1}^{M} \text{Var}(\widehat{\beta_m}) + M*(M-1)\overline{\text{Cov}}(\widehat{\beta_m}, \widehat{\beta_n})\right],$$

where $\overline{\text{Cov}}$ is the average covariance between regression estimates from two different regressions. We randomly draw 1,000 regressions and calculate the average of the pairwise covariances between estimates from these 1,000 regressions. This number is then plugged in the above equation to get the desired variance of the mean estimate. For robustness, we randomly sample twenty-five times and find that the difference in the

[4]Instead of using only the indicator variable for negative returns, we also experiment with using the returns in the above regression. While there is some evidence that correlations are higher with greater decline in returns, most of the impact of asymmetry in correlations is captured by the indicator variable.

computed variance is negligible. Using the average of the twenty-five randomly sampled 1,000 regressions and the single draw of 1,000 regressions results in t-statistics that differ by about 0.01.

In the first regression in Table 3 we set δ_2 and δ_4 to zero. The asymmetric correlation effect is strong: the average δ_3 is 0.035 with a t-statistic of 9.32; percentage of δ_3 with t-statistics greater than (less than) 1.96 (-1.96) is 9.21% (0.85%). Thus, individual stock correlations at the monthly frequency increase as returns decline. From Table 1 we know that the mean correlation is 0.129. When both stock returns i and j are negative, the correlation increases by 0.035, an increase of 27% over 0.129. Note that δ_1 is statistically indistinguishable from zero at the 5% level.

The table also presents the results for other variables. Correlation is persistent as evident by the positive coefficient, δ_6, on its lagged value. Surprisingly, turnover has a negative impact on correlations. Unlike the cross-sectional regressions of Table 2, higher average turnover leads to lower correlations. Consistent with the results of Andersen et al. (2001), the average volatility has a positive significant impact on correlations. Since macroeconomic shocks are likely to impact all stocks, we include variables (term spread, default spread and the three month T-bill yield) that proxy for the business cycle. The term spread and the T-bill yields are higher during expansions and the default spread is higher during recessions. The regression coefficients on the term spread and the T-bill yield are negative while that on default spread is significantly positive. Thus, these estimates imply that correlations are high during recessions and low during expansions. We also use the positive and negative measures of the standardized unexpected earnings (SUE) to test whether earnings surprises impact correlations. The coefficient on the dummy variable indicating negative SUE is positive and significant at the 10% level suggesting that, after controlling for the business cycle, a negative earnings surprise in stocks i and j leads to an increase in their correlation.

In the second regression specification δ_2 and δ_4 are not restricted to zero. The average δ_3, which is the coefficient on $I(R_i < 0, R_j < 0)$, is reduced to 0.015 with a t-statistic of 1.50. Moreover, the percentage of δ_3 with t-statistics greater than 1.96 is only 4.32%. Coefficient δ_4, which captures the interaction between turnover and $I(R_i < 0, R_j < 0)$, is statistically significant at 0.475 (t-statistic $= 3.01$). Further, the percentage of δ_4 with t-statistics greater than (less than) 1.96 (-1.96) is 5.90% (1.73%). Insignificant δ_3 along with a positive and significant δ_4 suggests that the asymmetric correlation effect is attributable to the interaction between trading and negative returns. In other words, it is an increase in trading volume when returns are negative that leads to an increase in correlations.

3.1 Asymmetric correlation in subsamples

We next examine the asymmetry in correlations across firms sorted by size and institutional holdings. We calculate market capitalization from CRSP over the sample period 1963-2008 and institutional holdings data is obtained from Thomson Financial for every NYSE firm between 1980 and 2008.[5] Firms are sorted by market capitalization and institutional holdings as follows. Each month stocks are ranked from 1 (lowest) to 10 (highest) based on market capitalization. A firm is defined as a small firm (large firm) if the time-series average ranking of the firm is less than 5 (greater than or equal to 5). Similarly, each quarter, stocks are ranked from 1 (lowest) to 10 (highest) based on institutional holdings. A firm is defined as a small institutional-holding firm (large institutional-holding firm) if the time-series average ranking is less than 5 (greater than or equal to 5).

We estimate the same regression as in equation (3) but average coefficients only across subgroups of firms as defined above. Table 4 presents the results. For brevity we only report the coefficients δ_1 through δ_4. The results for the case when two firms are in different size groups or in different institutional holdings group are similar to those in Table 3 and are omitted. Panel A shows that the coefficient δ_3 on $I(R_i < 0, R_j < 0)$ is statistically significant for small firms and large firms in the regression without interaction terms with turnover. For small firms (large firms), the average δ_3 is 0.034 (0.035) with a t-statistic of 8.97 (9.23). Thus, the asymmetric correlation effect is present in both small and large firms. When we interact $I(R_i < 0, R_j < 0)$ with turnover, the average δ_3 for small firms is reduced to 0.010 and becomes insignificant; δ_4 is still significantly positive. However, trading does not capture the asymmetric correlation effect for large firms: δ_3 is still significant at 0.028 with a t-statistic of 2.93 and δ_4 is not significant at the 5% level.

Panel B in Table 4 presents the results when the two firms are both low or both high institutional-holding firms. The results in Panel B exhibit very similar pattern to those in Panel A. The similarity arises because institutional holdings are concentrated in large firms (see, for example, Gompers and Metrick, 2001). For low institutional-holding firms, turnover explains the asymmetric correlation effect. By contrast, we find that for high institutional-holding firms, the asymmetric correlation effect is not captured by turnover: δ_3 is positive and significant while δ_4 is insignificantly different from zero.

Results in Table 4, thus, imply that trading accounts for asymmetric correlation in

[5] We verify that the asymmetry in correlations is present in the shorter sample period 1980-2008 as well as the sample period 1993-2008 (sample period analyzed in the next section).

small stocks and in low institutional-holding stocks but not for the other categories of stocks. The overall asymmetric correlation effect in Table 3 is driven by small and low institutional-holding firms.

3.2 Asymmetric systematic and residual correlations

We now examine the contribution to correlations from the systematic and the unsystematic components of returns. We use the market model to compute the systematic component and the residuals. From equation (2) we can write the correlation between stock i and stock j as,

$$\begin{aligned} \mathrm{Cor}_{ij} &= \frac{\beta_i \beta_j \sigma_m^2}{\sigma_i \sigma_j} + \frac{\sigma_{\epsilon_i,\epsilon_j}}{\sigma_i \sigma_j} \\ &= \rho_{im}\rho_{jm} + \rho_{\epsilon_i,\epsilon_j} \frac{\sigma_{\epsilon_i}\sigma_{\epsilon_j}}{\sigma_i \sigma_j} \,, \end{aligned} \tag{4}$$

where σ_i (σ_j) represents the standard deviation of returns of stock i (stock j); ρ_{im} (ρ_{jm}) denotes the correlation between returns of stock i (stock j) and the market; and $\rho_{\epsilon_i,\epsilon_j}$ denotes the correlation between the residuals of stock i and stock j.

Equation (4) has two parts. The first part represents the systematic component and the second part represents the unsystematic component of correlations. In order to estimate these two components, we estimate the following market model regression for each stock each month using daily data within the month:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t} \,. \tag{5}$$

Recall that we eliminate stocks smaller than \$5 from the sample. Thus, problems associated with non-synchronous trading are not likely to be severe in our sample. We, however, also estimate equation (5) using one and two leads and lags of market return. The results from these alternative specifications are very similar to the ones reported here.

Table 5 reports the results from this analysis. When returns are negative, the systematic component of small (large) stock correlations is 0.113 (0.207) and the unsystematic component is 0.024 (0.018). Thus, the unsystematic component comprises of 17.3% of the total small stock correlation while it comprises of only 8.2% of the large stock correlation. When returns are positive the unsystematic component accounts for 15.2% (9.1%) of the small (large) stock correlation. Thus, not only does the unsystematic portion of returns of small stocks comprises of a larger fraction of the total correlation than that for large stocks, it comprises an even greater fraction of total correlation when both returns are negative.

4 Source of asymmetric correlation

Our results so far show that trading in small stocks leads to asymmetric correlation and that a large part of the increase in correlations when returns are negative is due to an increase in residual correlation. Since small stocks are predominantly held by individuals, could it be the case that trading by small retail investors causes the asymmetry in correlations? We conjecture that it is coordinated selling by retail investors in the face of declining prices that causes the asymmetry in correlations. We provide evidence for our conjecture in several steps. First, we show in Section 4.1 that it is not merely trading (as shown previously in Tables 3 and 4) but rather small trades that explain the asymmetric correlation effect when returns are negative. The idea here is that small trades are more likely to come from retail investors. Second, we show in Section 4.2 that there is coordinated selling in small trades, again presumably from retail investors, when returns are negative. This is established by examining correlations in order imbalances. Third, we show in Section 4.3 that retail investors sell following negative market returns and that this leads to the co-ordinated selling in the small stocks.

The analysis in this section uses intraday transactions data. The NYSE Trade and Quote (TAQ) dataset provides detailed transactions data for each trade as well as bid-ask quotes. Details about sample selection and filtering rules for the TAQ data can be found in Chordia, Roll, and Subrahmanyam (2000). The sample period is from 1993 to 2008. The total number of different stocks in the sample is 3,073 and the average number of stocks per month is 1,521.

4.1 Correlation and trade size

We first interact turnover with trade size to examine whether trade size has any impact on asymmetric correlations. Average daily trade size is defined as the number of shares traded each day divided by the number of daily trades. Average daily trade size is averaged over the month to get a monthly average trade size. Each month, each stock's momthly trade size is compared to its average monthly trade size in the past three months. If the trade size is greater than (less than) the average, that month is defined as the month of large trades (small trades) for that stock. We run the following time series regressions:

$$
\begin{aligned}
\text{ResCorr}_{ijt} = {} & \delta_0 + \delta_1 I(R_{it} * R_{jt} > 0) + \delta_2 I(R_{it} * R_{jt} > 0) * \text{Turnover}_{ijt} \\
& + \delta_3 I(R_{it} < 0, R_{jt} < 0) + [\, \delta_{4a} I(LT_{it}, LT_{jt}) + \delta_{4b} I(ST_{it}, ST_{jt}) \\
& + \delta_{4c} I(OT)\,] I(R_{it} < 0, R_{jt} < 0) * \text{Turnover}_{ijt} + \delta_5 \text{Turnover}_{ijt} \\
& + \delta_6 \text{Cor}_{ijt-1} + \delta_7 \text{Vol}_{ijt} + \delta_8 \text{Term}_{t-1} + \delta_9 \text{Tbl}_{t-1} + \delta_{10} \text{Def}_{t-1} \\
& + \delta_{11} I(\text{SUE}_{it} > 0, \text{SUE}_{jt} > 0) + \delta_{12} I(\text{SUE}_{it} < 0, \text{SUE}_{jt} < 0) + \eta_{ijt} \, .(6)
\end{aligned}
$$

ResCorr represents the correlations in residuals from a market model with one lag and one lead, $I(LT_{it}, LT_{jt})$ is equal to one when both firms i and j have large size trades in month t, $I(ST_{it}, ST_{jt})$ is equal to one when both firms i and j have small size trades in month t, and $I(OT)$ is equal to one when one firm has large size trades and the other firm has small size trades. In other words, the above equation decomposes the coefficient δ_4 from equation (3) into three parts. We use residual correlation as dependent variable (rather than total correlation) because as we saw in the last section, the pairwise correlations in small stocks depend on the residuals from the market model.[6]

Table 6 presents the results. Panel A shows that the coefficient δ_{4b} of the interaction between the negative returns, turnover, and small trades is 0.195, with a t-statistic of 3.17, but the interaction term δ_{4a} of negative returns, turnover, and large trades is insignificant. Further, δ_3, the coefficient on the indicator variable for negative returns, is insignificant. The evidence suggests that it is the small trades that account for the asymmetry in correlations.

Panels B and C present results for small and large firms. Consistent with Panel A, we can see that the average δ_{4b} is more than twice that the average δ_{4a}. For small firms, for example, the average δ_{4b} is 0.250, with a t-statistic of 3.80 and the average δ_{4a} is an insignificant 0.112. Moreover the average δ_{4b} is higher for small firms than that for large firms. For large firms, the average δ_{4b} is 0.093, with a t-statistic of 1.85.

Panels D and E present results for low and high institutional-holding firms. The results are similar to those when sorting on size. Once again, for low institutional-holdings firms the coefficient of the interaction between the negative returns, turnover, and small trades, δ_{4b}, is significant but the interaction term of negative returns, turnover, and large trades, δ_{4a}, is insignificant. In sum, it is the small trades, presumably by retail investors that drive the asymmetry in correlations. The asymmetry is present mainly in

[6] We estimate the regression in equation (6) using total correlation as well and the results are similar. The main difference is that the coefficient estimates are now smaller and none of the δ_3 are significant.

small stocks and stocks with low institutional holdings.

There is some concern that in recent years institutions are increasingly submitting small order sizes (see Chordia, Roll, Subrahmanyam, 2011). We address this concern in two ways: (i) While small trades may come from retail investors as well as institutions, large trades are likely to originate primarily from institutions. Since we see a clear pattern across small and large trades, this allows us to use small trades as a (noisy) proxy for retail trading, and (ii) We replicate our results in Table 6 over the period 1993-2000, which is the pre-decimalization period when institutional trade sizes were not as small as more recently, and find similar results to the ones reported here.

4.2 OIB correlations

Our next step is to examine, using order imbalances (OIB), whether retail investors herd when returns are negative. We carry out this investigation in three sub-steps. First, we show that correlations between OIB are higher in months when returns are negative and trades sizes are small. Second, we breakup OIB into that coming from small and large trades, and show that it is correlation between small trades' OIB that is the highest when returns are negative. Third, we show that there is selling (OIB is negative) in months when returns are negative and trade sizes are small. The sum of this evidence suggests that there is coordinated selling by retail investors when returns are negative.

We use the TAQ data to calculate the order imbalances. Trades are signed using the Lee and Ready (1991) algorithm. We define daily order imbalances[7] using the number of trades and shares as follows:

- OIBNUM: the daily number of buyer-initiated trades less the number of seller-initiated trades scaled by the total number of trades.
- OIBSH: the daily buyer-initiated shares purchased less the seller-initiated shares sold scaled by the total number of shares traded.

We use the daily OIBNUM and OIBSH to calculate monthly pairwise OIBNUM correlations and OIBSH correlations among stocks. We require at least 15 pairwise OIBNUM (OIBSH) points each month. This yields about 220 million monthly pairwise OIBNUM (OIBSH) correlations over the 1993-2008 period. Each month, the OIBNUM and OIBSH

[7]Note that the Lee and Ready (1991) algorithm only signs the market orders. Thus, our measures of order imbalance relate only to market orders. Since the other side of the trade is taken up by the passive limit order book, our measure of order imbalance can be considered as coming from the active trades. See Chordia, Roll, Subrahmanyam (2005) for details on implementing the Lee and Ready algorithm.

correlations are grouped into 3×3 categories (large trades, small trades or different size trades; and positive returns, negative returns or returns with different signs).

Table 7 reports the means and standard deviations of the correlations for each category. Panel A presents the results for OIBNUM and Panel B presents the results for OIBSH. For both, OIBNUM and OIBSH, the correlations during months of small trades and negative returns have the highest means. The mean monthly pairwise correlation for the order imbalances measured as number of trades (shares traded) is 0.071 (0.026). Moreover, for each row the average correlation for negative returns is highest among three columns; for each column the average correlations for small trades is highest among the three rows. Given the large number of observations and small standard deviations, the mean difference between each category is statistically significant. Thus, there is correlated trading in small size trades when returns are negative suggesting that the retail investors[8] seem to be trading simultaneously when prices decline. Since negative returns are usually accompanied by selling, the evidence also suggests that retail investors simultaneously sell when prices decline. We will provide more direct evidence on selling shortly.

We classify each 'month' into small trades-month or large trades-month based on average trade size in Panels A and B of Table 7 and then look at correlations between OIB. Another way of looking at correlations between small trades is to directly classify each 'trade' as a small trade or a large trade. We examine each trade from 1993 through 2008 and classify it as small or big depending upon whether the dollars transacted per trade were less than or greater than $10,000. We then compute the daily small and big order imbalances per stock defined as follows:

- SOIBNUM: the daily number of buyer-initiated small trades less the number of seller-initiated small trades scaled by the total number of trades.
- BOIBNUM: the daily number of buyer-initiated big trades less the number of seller-initiated big trades scaled by the total number of trades.
- SOIBSH: the daily buyer-initiated shares purchased in small trades less the seller-initiated shares sold in small trades scaled by the total number of shares traded.
- BOIBSH: the daily buyer-initiated shares purchased in big trades less the seller-initiated shares sold in big trades scaled by the total number of shares traded.

[8] Since our measure of order imbalances signs only the market orders or the active as opposed to passive orders, when we use the term retail (institutional) investors in relation to the order imbalances, we mean the active retail (institutional) investors.

We then use these daily small and big order imbalances to compute the monthly pairwise correlations based on at least 15 pairwise order imbalances each month. Panel C of Table 7 presents the results for order imbalances measured using number of trades and Panel D presents correlations between order imbalances in terms of shares traded.

The average correlations in order imbalances measured using small trades are always higher than those measured using large trades. For instance, with negative monthly returns, the pairwise average correlation for SOIBNUM at 0.046 is higher than that of BOIBNUM at 0.039. While the SOIBNUM (SOIBSH) correlations are higher than BOIBNUM (BOIBSH) for all other return combinations as well, the SOIB correlations are substantially higher than BOIB correlations when returns are negative. For instance, with the order imbalance measured using trades, the average pairwise correlation for SOIBNUM is 18% higher than BOIBNUM when returns are negative, 11% higher when returns are positive and only 9% higher when the monthly pairwise returns have opposing signs. The differences in the average correlations measured using small and large trades are even larger for negative returns when the order imbalances are measured in terms of shares traded.

We now ask whether small investors are selling or buying stocks when the returns are negative. Table 8 presents the average order imbalance for small and large trades. It is clear that the average small trade imbalance is negative when the returns are negative. For instance, when returns are negative, the average SOIBNUM (SOIBSH) is -0.054% (-0.269%) while the average BOIBNUM (BOIBSH) is 0.376% (0.353%). This suggests that the source of the negative returns is the selling by retail investors.

The results in this subsection strongly suggest that there is herding in small trades, presumably because of trading by retail investors. This is consistent with the results of Barber, Odean, and Zhu (2009) who use their retail brokerage data to show that individual investors herd. The negative returns seem to be caused by the co-ordinated selling activity of small investors and this is what leads to the asymmetry in correlations. The question then arises: What causes this co-ordinated selling by retail investors?

4.3 Search for the source of co-ordinated selling

We conjecture that retail investors sell in the face of declining market prices possibly triggered by margin calls or stop loss orders and this leads to the co-ordinated selling in small stocks. To test this hypothesis we first sort the sample months on the basis of market returns in the first five trading days[9] of the month and examine the behavior of the small trades, specially in the small stocks, during the remaining days of the month. The idea is that negative market returns in the initial days of the month trigger co-ordinated selling by retail investors in the remaining days of the month.

Table 9 presents the results. Panel A presents the results based on first five days market returns being negative and positive while Panel B presents the results conditional on the average daily market returns being two standard deviations higher or lower than the mean of the daily returns. We sort firms into large and small on the basis of market capitalization at the beginning of the month and examine returns, order imbalances, correlations between returns, and correlations between order imbalances over the remaining seventeen (typically) trading days of the month. The table reports the overall means of all these variables.

Let us focus first on returns. When the market return is negative during the first five days of the month, then over the remaining days of the month, the mean small stock returns are 0.025% whereas the large stock returns are 0.037%. On the other hand, when the initial market return is positive, the small stock returns are 0.072% while the large stock returns are 0.034% over the remaining days of the month. Thus, small stock returns are lower when the initial market returns are negative and higher when the initial market returns are positive. This suggests that large stocks experience a slight reversal as compared to the small stocks which experience a continuation pattern. The pairwise return correlation and the residual correlation patterns are also higher with the initial negative market returns than those with initial positive market return. A similar pattern obtains in Panel B.

We now turn to the order imbalances. All the order imbalances are positive except for those of the small stocks when the initial market returns are negative. This does suggest that there is more selling in small stocks following the initial negative market returns. Focusing on the number of trades in Panel A, we see that the mean SOIBNUM is -1.33% whereas BOIBNUM is -0.44%. Thus, the small sell trades exceed the small buy trades by more than what the large sell trades exceed the large buy trades. However, SOIBSH is -1.28% whereas the BOIBSH is -1.95% suggesting that while there are fewer large trades, these trades involve the sale of a large number of shares. When considering the large stocks with negative initial market returns, BOIBNUM and BOIBSH are both positive and larger than SOIBNUM and SOIBSH respectively, suggesting that amongst large stocks the increase in returns following negative market returns is caused mainly

[9] We find similar results when conditioning on market returns over the first eleven trading days of the month.

by the large traders.

In order to show that there is co-ordinated selling in small stocks by retail traders following the initial negative market returns, we finally examine the pairwise correlations in the small and large order flows across large and small stocks. Consider Panel A with negative initial market returns. For small stocks SOIB pairwise correlation is 0.039 (0.032) when measured in terms of number of trades (number of shares) while the BOIB pairwise correlation is 0.016 (0.007) when measured in terms of number of trades (shares). Thus, for small stocks the pairwise correlation in order flow is higher when the trade size is small. A similar pattern obtains for small stocks when the initial market returns are positive. However, in every case for the small stocks, the ratio of correlations between SOIB and BOIB is larger when the initial market return is negative than when it is positive. On the other hand, for large stocks, the pairwise correlation in order flow is actually higher for BOIB when the order flow in measured in terms of number of trades. The results in Panel B, when conditioning on market returns being two standard deviations from the mean, are qualitatively very similar.

To summarize briefly the results in this section, we provide evidence that negative market returns leads to co-ordinated selling by retail investors (proxied by small trades) in small-cap stocks. It is this co-ordinated selling that leads to higher correlations in small stock returns when small stock returns are negative.

5 Conclusions

We search for the source of the asymmetric correlations between individual stocks. We find that the trading activity governs the asymmetric correlation phenomenon in individual stocks with high retail concentrations (i.e. small stocks/low institutional-holding stocks). More specifically, the positive coefficient in the regression of correlation on negative returns is entirely captured by the interaction of return and the trading activity. An increase in trading volume when returns are negative leads to an increase in correlations. The results suggest that co-ordinated selling activity across stocks drives the asymmetry in correlations in small stocks.

We explore who is behind this co-ordinated selling. We conjecture that the trading activity of retail investors causes the asymmetric correlations. This conjecture is supported by our empirical results: it is the small trades, presumably by retail investors that drives the asymmetry in correlations; there is correlated selling in small size trades when returns are negative suggesting that retail investors seem to be selling simultaneously when prices decline. We provide support for the argument that it is an initial decline in market returns that leads to this co-ordinated selling in the small stocks/low institutional-holding stocks.

References

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001, "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics* 61, 43-76.

Ang, A., and G. Bekaert, 2000, "International Asset Allocation with Time-varying Correlations," *Review of Financial Studies* 15, 1137-1187.

Ang, A., and J. Chen, 2002, "Asymmetric Correlations of Equity Portfolios," *Journal of Financial Economics* 63, 443-494.

Ang, A., and J. Chen, and Y. Xing, 2006, "Downside Risk," *Review of Financial Studies* 19, 1191-1239.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The Cross-Section of Volatility and Expected Returns," *Journal of Finance* 61, 259-299.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2009, "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence," *Journal of Financial Economics* 91, 1-23.

Bae, K. H., G. A. Karolyi, and R. M. Stulz, 2003, "A New Approach to Measuring Financial Market Contagion," *Review of Financial Studies* 16, 717-64.

Ball, R., and S. P. Kothari, 1989, "Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns," *Journal of Financial Economics* 25, 51-74.

Barber, B. M., T. Odean, and N. Zhu, 2009, "Do Noise Traders Move Markets?," *Review of Financial Studies*, Forthcoming.

Barberis, N., A. Shleifer, and J. Wurgler, 2005, "Comovement," *Journal of Financial Economics* 75, 283-317.

Bekaert, G., and G. Wu, 2000, "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies* 13, 1-42.

Campbell, J. Y., and L. Hentschel, 1992, "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics* 31, 281-318.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance* 56, 1-44.

Cho, Y. H., and R. F. Engle, 2000, "Time-Varying Betas and Asymmetric Effects of News: Empirical Analysis of Blue Chip Stocks," Working Paper, National Bureau of Economic Research, Cambridge, MA.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, "Commonality in Liquidity," *Journal of Financial Economics* 56, 3-28.

Chordia, T., R. Roll, and A. Subrahmanyam, 2005, "Evidence on the Speed of Convergence to Market Efficiency," *Journal of Financial Economics* 76, 271-292.

Chordia, T., R. Roll, and A. Subrahmanyam, 2011, "Recent Trends in Trading Activity and Market Quality," forthcoming *Journal of Financial Economics.*

Conrad, J., M. Gultekin, and G. Kaul. 1991, "Asymmetric Predictability of Conditional Variances," *Review of Financial Studies* 4, 597-622.

French, K. R., G. W. Schwert, and R. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics* 19, 3-29.

Gompers, P.A. and A. Metrick, 2001, "Institutional Investors And Equity Prices," *Quarterly Journal of Economics* 116, 229-259.

Green, T.C., and B. Hwang, 2009, "Price-Based Return Comovement," *Journal of Financial Economics* 93, 37-50.

Harvey, C. R., and A. Siddique, 2000, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance* 55, 1263-95.

Hong, Y., J. Tu, and G. Zhou, 2007, "Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation," *Review of Financial Studies* 20, 1547-1581.

Hvidkjaer, S., 2006, "A Trade-based Analysis of Momentum," *Review of Financial Studies* 19, 457-491.

Hvidkjaer, S., 2008, "Small Trades and the Cross-Section of Stock Returns," *Review of Financial Studies* 21, 1123-1151.

Jones, C., G. Kaul, and M. Lipson, 1994, "Transactions, Volume, and Volatility," *Review of Financial Studies* 7, 631-651.

Kaniel, R., G. Saar, and S. Titman, 2008, "Individual Investor Trading and Stock Returns," *Journal of Finance* 63, 273-310.

Karolyi A., and R. Stulz, 1996, "Why Do Markets Move Together? An Investigation of US-Japan Stock Return Comovements," *Journal of Finance* 51, 951-986.

Kumar A., and C. Lee, 2006, "Retail Investor Sentiment and Return Comovements," *Journal of Finance* 61, 2451-2486.

Lee, C., and M. Ready, 1991, "Inferring Trade Direction from Intraday Data," *Journal of Finance* 46, 733-747.

Long, F., and B. Solnik, 2001, "Extreme Correlation of International Equity Markets," *Journal of Finance* 56, 649-676.

Newey, W., and K. West, 1987, “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica* 55, 703-708.

Newey, W., and K. West, 1994, “Automatic Lag Selection in Covariance Matrix Estimation,” *Review of Economic Studies* 61, 631-653.

Schwert, G. W., 1989, “Why Does Stock Market Volatility Change Over Time?,” *Journal of Finance* 44, 1115-1153.

Table 1: Summary Statistics

The table reports means, medians, and standard deviations of the correlations. Panel A, Panel B, and Panel C provide the summary statistics for all time periods, expansionary periods, and recessionary periods (defined by the National Bureau of Economic Research), respectively. Panel D and Panel E report the summary statistics for both positive returns and negative returns. We divide the stocks into large size group and small size group based on market capitalization of the firm at the beginning of month. All column is the summary results for all stock pairwise correlations without breaking stocks down into small and large group. Small/Small column is the summary results for pairwise correlations of both stocks in the small size group. Large/Large column is the summary results for pairwise correlations of both stocks in the large size group. Small/Large column is the summary results for stocks in different groups. The sample period is from 1963 to 2008. Stocks with price less than $5 are excluded from the sample.

	All	Small/Small	Large/Large	Small/Large
Panel A: All sample				
Mean	0.129	0.095	0.172	0.117
Median	0.128	0.092	0.175	0.116
Standard deviation	0.234	0.231	0.235	0.231
Panel B: Expansion				
Mean	0.117	0.089	0.161	0.109
Median	0.117	0.087	0.164	0.109
Standard deviation	0.232	0.229	0.231	0.228
Panel C: Recession				
Mean	0.175	0.131	0.240	0.166
Median	0.176	0.126	0.245	0.166
Standard deviation	0.246	0.242	0.241	0.242
Panel D: Both returns are positive				
Mean	0.149	0.121	0.189	0.139
Median	0.150	0.118	0.194	0.140
Standard deviation	0.222	0.222	0.222	0.221
Panel E: Both returns are negative				
Mean	0.169	0.138	0.223	0.164
Median	0.169	0.135	0.227	0.164
Standard deviation	0.242	0.239	0.241	0.237

Table 2: Cross-Sectional Determinants of Correlation

We estimate the following Fama-MacBeth cross-sectional regression:

$$C_{ijt} = \gamma_{0t} + \gamma_{1t}\left(\frac{X_{it} + X_{jt}}{2}\right) + u_{ijt}\,.$$

The dependent variable, C_{ijt}, is the correlation of stock i and stock j in month t, calculated using the intra-month data. Unless otherwise noted, all explanatory variables are the average values for stock i and stock j. Beta is estimated following the methodology described in Fama and French (1992). Size is market capitalization in billions of dollars, BM is book-to-market, Turnover is monthly share trading volume divided by shares outstanding, and Vol is monthly return volatility. $I(\cdot)$ refers to the indicator function. R_i and R_j denote stock i and stock j returns. $R_{i,1-12}$ and $R_{j,1-12}$ denote past year cumulative returns for stocks i and j, respectively. I(Similar Prices) is equal to one if stock i and stock j have similar prices, defined as $|\text{Price}_i - \text{Price}_j| / \max(\text{Price}_i, \text{Price}_j) < 25\%$. I(Same Size/BM group) is the indicator function to denote stocks that are within same size and BM group. I(S&P500) is equal to one if both stocks are in the S&P 500 index. I(Same Industry) is equal to one if both stocks are in the same industry. The t-statistics, adjusted by Newey and West (1987) standard errors, are shown in parenthesis. The last row shows the average $\overline{R}^2$ in percent. The sample period is from 1963 to 2008. Stocks with price less than \$5 are excluded from the sample.

	(1)	(2)	(3)
Constant	0.103	-0.399	-0.471
	(20.12)	(-16.31)	(-26.30)
Beta	0.037	0.052	0.044
	(7.92)	(11.20)	(9.32)
Ln(Size)		0.059	0.055
		(29.12)	(26.20)
Ln(BM)		0.005	0.004
		(4.43)	(2.98)
I($R_i * R_j > 0$)		0.005	0.004
		(5.32)	(3.20)
I($R_i < 0, R_j < 0$)		0.011	0.008
		(6.30)	(4.89)
I($R_{i,1-12} * R_{j,1-12} > 0$)		0.006	0.005
		(6.30)	(4.37)
I($R_{i,1-12} < 0, R_{j,1-12} < 0$)		0.003	0.003
		(1.93)	(2.77)
Ln(Turnover)			0.017
			(4.00)
Ln(Vol)			0.039
			(3.50)
I(Same Size/BM group)			0.016
			(17.01)
I(Similar Prices)			0.022
			(16.89)
I(S&P500)			0.023
			(10.99)
I(Same Industry)			0.036
			(8.86)
$\bar{R}^2$	1.01	7.32	9.30

Table 3: Asymmetric Correlation and Turnover

We estimate the following time series regressions:

$$\begin{aligned}
\mathrm{Cor}_{ijt} &= \delta_0 + \delta_1 I(R_{it} * R_{jt} > 0) + \delta_2 I(R_{it} * R_{jt} > 0) * \mathrm{Turnover}_{ijt} \\
&\quad + \delta_3 I(R_{it} < 0, R_{jt} < 0) + \delta_4 I(R_{it} < 0, R_{jt} < 0) * \mathrm{Turnover}_{ijt} \\
&\quad + \delta_5 \mathrm{Turnover}_{ijt} + \delta_6 \mathrm{Cor}_{ijt-1} + \delta_7 \mathrm{Vol}_{ijt} + \delta_8 \mathrm{Term}_{t-1} + \delta_9 \mathrm{Tbl}_{t-1} \\
&\quad + \delta_{10} \mathrm{Def}_{t-1} + \delta_{11} I(\mathrm{SUE}_{it} > 0, \mathrm{SUE}_{jt} > 0) + \delta_{12} I(\mathrm{SUE}_{it} < 0, \mathrm{SUE}_{jt} < 0) + \eta_{ijt}\ ,
\end{aligned}$$

where Cor_{ijt} is the correlation between stocks i and j in month t, R is the return, $\mathrm{Turnover}_{ijt}$ is the average of turnover of the two stocks, and Vol_{ijt} is the average of the two stock volatilities. I($\cdot$) is the indicator function. Term, Tbl and Def denote the term spread, the T-bill yield and the default spread, respectively. SUE is the standardized unexpected earnings, computed as the most recently announced quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. The table reports the (cross-sectional) means of the parameter estimates and their t-statistics, which take into account cross-sectional correlation in estimators of each regression. We also report the percentage of parameter estimates with t-statistics less than -1.96 and greater than 1.96. The sample period is from 1963 to 2008. Stocks with price less than \$5 are excluded from the sample.

	δ_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}	δ_{12}
Mean	0.273	-0.001		0.035		-0.212	0.010	0.000	-2.102	-3.039	0.653	0.002	0.004
t(Mean)	54.34	-1.20		9.32		-3.78	2.75	16.12	-12.00	-5.39	6.42	1.43	1.79
$\%(t < -1.96)$	1.01	3.15		0.85		7.12	1.99	1.78	14.32	20.21	2.29	3.21	2.38
$\%(t > 1.96)$	27.32	2.76		9.21		4.19	5.54	16.89	2.12	3.01	7.39	4.38	4.31
Mean	0.263	-0.003	-0.002	0.015	0.475	-0.112	0.026	0.000	-2.218	-2.831	0.682	0.002	0.003
t(Mean)	47.45	-0.59	-0.12	1.50	3.01	-1.63	3.09	10.11	-10.03	-5.16	6.43	1.59	1.72
$\%(t < -1.96)$	1.12	2.98	3.32	1.99	1.73	5.32	1.58	1.15	14.12	20.32	2.23	3.11	2.43
$\%(t > 1.96)$	23.21	2.30	3.10	4.32	5.90	5.01	6.60	19.37	2.92	3.90	6.90	4.54	4.25

Table 4: Asymmetric Correlation and Turnover For SubSamples

We estimate the same time-series regression as in Table 3 but report only coefficients δ_1–δ_4. We further divide the sample into various groups based on market capitalization in Panel A and institutional holding in Panel B. The left part (right part) in Panel A reports results that both firms i and j are small firms (large firms). Each month stocks are ranked from 1 (lowest) to 10 (highest) based on market capitalization. A firm is defined as a small firm (large firm) if the time-series average ranking is less than 5 (greater than 5). The left part (right part) in Panel B reports results that both firms i and j are low (high) institutional-holding firms. Each quarter stocks are ranked from 1 (lowest) to 10 (highest) based on institutional holdings. A firm is defined as a low (high) institutional-holding firm if the time-series average ranking is less than 5 (greater than 5). The sample period is from 1963 to 2008 in Panel A, and from 1980 to 2008 in Panel B. Stocks with price less than \$5 are excluded from the sample.

	δ_1	δ_2	δ_3	δ_4	δ_1	δ_2	δ_3	δ_4
Panel A: Subsamples based on market capitalization								
	Small				Large			
Mean	-0.002		0.034		-0.001		0.035	
t(Mean)	-1.27		8.97		-1.42		9.23	
$\%(t < -1.96)$	2.86		1.01		3.73		0.68	
$\%(t > 1.96)$	2.20		8.60		2.43		9.04	
Mean	-0.002	-0.126	0.010	0.592	-0.000	-0.053	0.028	0.324
t(Mean)	-0.15	-1.01	1.02	4.21	0.52	-0.76	2.93	1.75
$\%(t < -1.96)$	2.57	2.86	2.76	2.01	2.83	3.25	1.90	2.20
$\%(t > 1.96)$	2.37	2.39	3.98	6.54	2.80	2.76	6.08	4.09
Panel B: Subsamples based on institutional-holding								
	Low				High			
Mean	-0.002		0.035		-0.002		0.039	
t(Mean)	-2.74		9.59		-2.23		11.36	
$\%(t < -1.96)$	3.21		0.67		3.21		0.39	
$\%(t > 1.96)$	2.39		9.19		2.03		9.98	
Mean	0.002	-0.108	0.014	0.520	-0.003	0.023	0.032	0.184
t(Mean)	0.38	-1.00	1.45	3.54	-0.83	0.01	3.37	1.39
$\%(t < -1.96)$	2.61	3.47	2.11	2.08	2.84	2.86	1.83	2.23
$\%(t > 1.96)$	2.79	2.52	4.03	6.82	2.51	2.90	6.65	4.12

Table 5: Asymmetric Systematic and Residual Correlation

We estimate the following regression each month:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t} \ ,$$

where $R_{i,t}$ is stock i daily return and $R_{m,t}$ is the market daily return on day t. Correlations of returns with the market are denoted by ρ_{im} and ρ_{jm}. Residual correlation is $\rho_{\epsilon_i \epsilon_j}$. Return standard deviations are σ_i and σ_j. Residual standard deviations are σ_{ϵ_i} and σ_{ϵ_j}. The table reports the various components of the following decomposition of correlation:

$$\rho_{ij} = \rho_{im}\rho_{jm} + \rho_{\epsilon_i,\epsilon_j} \frac{\sigma_{\epsilon_i}\sigma_{\epsilon_j}}{\sigma_i \sigma_j} \ .$$

Left panel ($R_i < 0, R_j < 0$) reports the statistics when both returns are negative, while the right panel ($R_i > 0, R_j > 0$) reports the statistics when both returns are positive. Sample is grouped into small firms and large firms based on the market capitalization of the firms at the beginning of each month. The sample period is from 1963 to 2008. Stocks with price less than $5 are excluded from the sample.

	$R_i < 0, R_j < 0$			$R_i > 0, R_j > 0$		
	Small/Small	Large/large	Small/Large	Small/Small	Large/large	Small/Large
$\rho_{im}\rho_{jm}$	0.113	0.207	0.150	0.0995	0.169	0.125
$\rho_{\epsilon_i \epsilon_j}$	0.0288	0.0253	0.0262	0.0192	0.0217	0.0204
$\sigma_{\epsilon_i}\sigma_{\epsilon_j}$	0.0014	0.0008	0.0011	0.0013	0.0007	0.0010
$\sigma_i \sigma_j$	0.0017	0.0011	0.0015	0.0014	0.0009	0.0012
$\rho_{\epsilon_i \epsilon_j}\sigma_{\epsilon_i}\sigma_{\epsilon_j}/\sigma_i\sigma_j$	0.024	0.018	0.019	0.018	0.017	0.017
ρ_{ij}	0.137	0.225	0.169	0.117	0.186	0.142

Table 6: Residual Correlation and Trade Size

We estimate the following time series regressions:

$$
\begin{aligned}
\text{Res}_{ijt} \quad = \quad & \delta_0 + \delta_1 I(R_{it} * R_{jt} > 0) + \delta_2 I(R_{it} * R_{jt} > 0) * \text{Turnover}_{ijt} + \delta_3 I(R_{it} < 0, R_{jt} < 0) \\
& + [\delta_{4a} I(LT_{it}, LT_{jt}) + \delta_{4b} I(ST_{it}, ST_{jt}) + \delta_{4c} I(OT)] I(R_{it} < 0, R_{jt} < 0) * \text{Turnover}_{ijt} \\
& + \delta_5 \text{Turnover}_{ijt} + \delta_6 \text{Res}_{ijt-1} + \delta_7 \text{Vol}_{ijt} + \delta_8 \text{Term}_{t-1} + \delta_9 \text{Tbl}_{t-1} + \delta_{10} \text{Def}_{t-1} \\
& + \delta_{11} I(\text{SUE}_{it} > 0, \text{SUE}_{jt} > 0) + \delta_{12} I(\text{SUE}_{it} < 0, \text{SUE}_{jt} < 0) + \eta_{ijt} \,.
\end{aligned}
$$

Most of the variables are described in Table 3. Res_{ijt}, is the residual correlation of stock i and stock j in month t, calculated using the intra-month data. $I(LT_{it}, LT_{jt})$ is equal to one when both firms i and j have large size trades in month t, $I(ST_{it}, ST_{jt})$ is equal to one when both firms i and j have small size trades in month t, and $I(OT)$ is equal to one when one firm has large size trades and the other firm has small size trades. Sample is divided further based on market capitalization and institutional holding, as described in Table 4. The sample period is from 1993 to 2008. Stocks with price less than \$5 are excluded from the sample.

	δ_1	δ_2	δ_3	δ_{4a}	δ_{4b}	δ_{4c}
Panel A: All firms						
Mean	0.000	-0.010	0.002	0.080	0.195	0.139
t(Mean)	0.34	-0.28	1.43	1.03	3.17	2.10
$\%(t < -1.96)$	1.54	2.12	2.02	2.23	1.02	1.68
$\%(t > 1.96)$	1.95	1.92	2.60	2.80	4.30	3.60
Panel B: Small capitalization firms						
Mean	0.000	-0.015	0.001	0.112	0.250	0.208
t(Mean)	0.15	-1.32	0.32	1.80	3.80	2.97
$\%(t < -1.96)$	1.30	2.75	2.73	1.61	1.18	1.77
$\%(t > 1.96)$	1.92	1.92	2.35	2.59	6.28	4.06
Panel C: Large capitalization firms						
Mean	0.000	-0.006	0.003	0.038	0.093	0.072
t(Mean)	0.78	-0.11	1.32	0.35	1.85	1.63
$\%(t < -1.96)$	1.67	2.56	2.17	1.92	1.21	1.32
$\%(t > 1.96)$	2.19	1.67	2.83	2.85	3.27	2.86
Panel D: Low institutional-holding firms						
Mean	0.000	-0.011	0.000	0.100	0.218	0.232
t(Mean)	0.02	-1.23	0.00	1.72	3.39	2.43
$\%(t < -1.96)$	1.99	2.57	2.20	1.50	2.08	1.93
$\%(t > 1.96)$	1.63	2.01	2.22	2.60	5.86	4.10
Panel E: High institutional-holding firms						
Mean	0.000	-0.005	0.001	0.045	0.089	0.062
t(Mean)	1.11	-0.55	1.10	0.59	1.65	1.52
$\%(t < -1.96)$	1.66	2.65	2.32	1.73	1.85	1.62
$\%(t > 1.96)$	2.32	1.95	3.02	2.27	3.05	2.90

Table 7: Correlations in Order Imbalance

We calculate order imbalance (OIB) as the difference between all buy trades and all sell trades, small order imbalance (SOIB) as the difference between small buy trades and small sell trades, and large order imbalance (BOIB) as the difference between large buy trades and large sell trades, where a trade is classified as a small or a large trade depending on whether dollars transacted per trade were less than or greater than $10,000. All order imbalances are calculated in number of trades (OIBNUM, SOIBNUM, BOIBNUM) and in number of shares (OIBSH, SOIBSH, BOIBSH). Pairwise monthly correlations are calculated using these daily measures. The table reports means and standard deviations of these correlations for various subsamples. Each month is classified into '$R_i > 0, R_j > 0$' (returns of both stocks are greater than 0), '$R_i < 0, R_j < 0$' (returns of both stocks are less than 0) and 'Other' (returns of the two stocks have different signs). Panels A and B further classify the months into 'ST$_i$, ST$_j$' (small trade month for both stocks), 'LT$_i$, LT$_j$' (large trade month for both stocks), and 'ST$_i$, LT$_j$' (large trade month for one stock and small trade month for the other stock). A month is classified as LT or ST for a stock if the stock's average trade size in that month greater or smaller than its average trade size in the past three months. Thus each month is placed into one of 3×3 categories in Panels A and B and into one of 3×1 categories in Panels C and D. Panel A reports correlations between pairwise OIBNUM while Panel B reports correlations between pairwise OIBSH. Panel C reports correlations between pairwise SOIBNUM (first row), between pairwise BOIBNUM (second row), and between pairwise SOIBNUM and BOIBNUM (third row). Panel D reports correlations between pairwise SOIBSH (first row), between pairwise BOIBSH (second row), and between pairwise SOIBSH and BOIBSH (third row). The sample period is from 1993 to 2008. Stocks with price less than $5 are excluded from the sample.

	$R_i > 0, R_j > 0$		$R_i < 0, R_j < 0$		Other	
	Mean	Std	Mean	Std	Mean	Std
Panel A: Corr(OIBNUM$_i$, OIBNUM$_j$)						
ST$_i$, ST$_j$	0.063	0.239	0.071	0.241	0.060	0.239
LT$_i$, LT$_j$	0.052	0.237	0.058	0.239	0.050	0.238
ST$_i$, LT$_j$	0.055	0.237	0.062	0.240	0.054	0.238
Panel B: Corr(OIBSH$_i$, OIBSH$_j$)						
ST$_i$, ST$_j$	0.026	0.233	0.026	0.240	0.024	0.239
LT$_i$, LT$_j$	0.017	0.232	0.021	0.239	0.018	0.238
ST$_i$, LT$_j$	0.021	0.232	0.024	0.239	0.021	0.238
Panel C: Correlations between SOIBNUM and BOIBNUM						
Corr(SOIBNUM$_i$, SOIBNUM$_j$)	0.040	0.241	0.046	0.241	0.038	0.241
Corr(BOIBNUM$_i$, BOIBNUM$_j$)	0.036	0.236	0.039	0.236	0.035	0.236
Corr(SOIBNUM$_i$, BOIBNUM$_j$)	0.037	0.237	0.043	0.239	0.036	0.238
Panel D: Correlations between SOIBSH and BOIBSH						
Corr(SOIBSH$_i$, SOIBSH$_j$)	0.030	0.241	0.039	0.241	0.024	0.239
Corr(BOIBSH$_i$, BOIBSH$_j$)	0.017	0.233	0.019	0.234	0.015	0.234
Corr(SOIBSH$_i$, BOIBSH$_j$)	0.023	0.237	0.026	0.240	0.018	0.237

Table 8: Average Order Imbalance

OIBNUM and OIBSH denote the order imbalance in number of transactions and shares. SOIBNUM and SOIBSH (BOIBNUM and BOIBSH) denote the order imbalance in number of transactions and shares for small trades (large trades). All variables are the average values for stock i and stock j. The table reports means and standard deviations (in percentage) of the order imbalance variables for various subsamples. Each month is classified into '$R_i > 0, R_j > 0$' (returns of both stocks are greater than 0), '$R_i < 0, R_j < 0$' (returns of both stocks are less than 0) and 'Other' (returns of the two stocks have different signs). The sample period is from 1993 to 2008. Stocks with price less than \$5 are excluded from the sample.

	$R_i > 0, R_j > 0$		$R_i < 0, R_j < 0$		Other	
	Mean	Std	Mean	Std	Mean	Std
Panel A: OIBNUM						
OIBNUM	3.452	9.142	0.322	9.375	1.663	9.513
SOIBNUM	1.459	7.288	-0.054	7.124	0.540	7.312
BOIBNUM	1.993	3.612	0.376	3.612	1.123	3.693
Panel B: OIBSH						
OIBSH	6.319	8.764	0.084	9.934	2.524	9.423
SOIBSH	1.176	3.421	-0.269	3.846	0.281	3.829
BOIBSH	5.143	7.613	0.353	8.331	2.243	7.941

Table 9: Market Impact

Sample months are divided into groups based on the market return, R_m, for the first five days for each month. Panel A makes this division based on the sign of the market return while Panel B makes this division based on whether the market return is higher or lower than the two times the standard deviation of market returns. Stocks are grouped into small firms and large firms based on the market capitalization of the firms at the beginning of each month. The table reports average daily percentage return, return correlations, residual correlations, as well as order imbalance (SOIBNUM, BOIBNUM, SOIBSH and BOIBSH) means and their correlations in each group for the remaining days of the month. The sample period is from 1993 to 2008. Stocks with price less than $5 are excluded from the sample.

Panel A: Based on market return in the first five days

	$R_m < 0$			$R_m > 0$		
	Small/Small	Large/Large	Small/Large	Small/Small	Large/Large	Small/Large
Return Mean	0.025	0.037	0.028	0.072	0.034	0.055
Return Corr	0.141	0.226	0.169	0.118	0.186	0.132
Residual Corr	0.020	0.019	0.017	0.014	0.016	0.011
SOIBNUM Mean	-1.329	1.319	0.196	0.164	0.951	0.641
SOIBNUM Corr	0.039	0.066	0.047	0.033	0.061	0.041
BOIBNUM Mean	-0.442	2.421	1.098	0.041	2.312	1.481
BOIBNUM Corr	0.016	0.069	0.034	0.011	0.067	0.031
SOIBSH Mean	-1.282	0.128	-0.062	0.180	0.287	0.272
SOIBSH Corr	0.032	0.050	0.033	0.023	0.044	0.028
BOIBSH Mean	-1.951	5.632	2.021	0.149	5.331	3.193
BOIBSH Corr	0.007	0.029	0.020	0.009	0.026	0.016

Panel B: Based on two-std of market return in the first five days

	R_m two-std below mean			R_m two-std above mean		
	Small/Small	Large/Large	Small/Large	Small/Small	Large/Large	Small/Large
Return Mean	0.020	0.049	0.032	0.097	0.028	0.040
Return Corr	0.156	0.203	0.180	0.102	0.178	0.153
Residual Corr	0.024	0.016	0.020	0.012	0.018	0.015
SOIBNUM Mean	-1.924	1.831	0.341	0.237	0.802	0.643
SOIBNUM Corr	0.048	0.061	0.053	0.030	0.062	0.048
BOIBNUM Mean	-0.732	2.932	1.230	0.108	2.031	1.430
BOIBNUM Corr	0.020	0.063	0.042	0.020	0.068	0.039
SOIBSH Mean	-1.731	0.032	-0.342	0.202	0.252	0.231
SOIBSH Corr	0.036	0.045	0.039	0.020	0.047	0.032
BOIBSH Mean	-1.976	6.002	2.382	0.290	5.129	3.012
BOIBSH Corr	0.012	0.023	0.017	0.006	0.028	0.019

Figure 1: Average Correlations

The figure plots the equally weighted correlations (solid line), the market capitalization value-weighted correlations (heavier solid line) and value-weighted market returns (dashed line). The bottom panel shows the 12-month moving average of the correlations and market returns from the top panel. Columns highlight NBER-dated recessions.