Firm-Level Return Comovement

Huafeng (Jason) Chen*

Shaojun (Jenny) Chen*

Feng Li**

Abstract

We study the pair-wise stock return correlations at the firm level. We find that over 85 percent of the total variation in firm-level pair-wise correlations of stock returns is not explained by an exhaustive list of variables that may affect return comovement. We also find that stocks with high correlations in the past have similar returns in the future, even after controlling for the possible determinants of the return correlations. Finally, a trading strategy based on stocks that deviate from their comovers generates abnormal returns.

August 2008

* Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC V6T 1Z2, Canada. ** Ross School of Business, University of Michigan, 701 Tappan St, Ann Arbor, MI 48109, USA. Huafeng Chen acknowledges the support from the Social Sciences and Humanities Research Council of Canada and the UBC Bureau of Asset Management. We thank seminar participants at the University of Vienna, the University of British Columbia, Connor, Clark & Lunn Investment Management, the UBC Whistler Summer Conference and the University of Oregon for their helpful comments. Please email inquiries to jason.chen@sauder.ubc.ca, jenny.chen@sauder.ubc.ca, or feng@umich.edu.

I. Introduction

"*Whirlpool corporation engages in the manufacture and marketing of home appliances worldwide. Its principal products include laundry appliances, refrigerators and freezers, cooking appliances, dishwashers, room air-conditioning equipment, and mixers and other small household appliances.*" This is an excerpt from the business summary of Whirlpool Corporation from *finance.yahoo.com*. Judging from this description of the business of Whirlpool, which stock do you think moves the most closely with Whirlpool? Probably another household appliance maker?

We estimate the stock return correlations between Whirlpool and all other stocks between 2001 and 2005. The 50 stocks with the highest correlations are reported in Table I. Maytag, one of Whirlpool's main competitors in the laundry appliances market, makes it to the list as No. 22. However, the top of the list belongs to Continental Airlines. Defying any suspicion that one airline simply by chance comoves with Whirlpool, a few other airlines also rank high in the list: Alaska Airlines is No. 4, American Airlines No. 9, and Mesa Air group No. 25. After some research, we find a potential link: Whirlpool Corporation owns a financial subsidiary that leases aircrafts.

This exercise suggests that there could be important information contained in the comovement between stock returns that we do not understand yet. In this paper, we explore systematically the determinants and persistence of firm-level pair-wise return comovements and examine their implications for future returns. More specifically, we ask the following questions. First, how much can the common variables examined in the finance literature explain the pair-wise stock return comovement? Prior research has examined the effect of specific events (such as index membership (Barberis, Shleifer, and Wurgler (2005)) on comovement. But there's virtually no systematic evidence on the determinants of the stock return comovement at the *firm* level.

To provide some first evidence on this issue, we estimate the pair-wise stock return correlations in each year using return data from the last five years. There are substantial variations in the pair-wise correlations — while the average Pearson correlation coefficient of two stocks' monthly returns in a five-year window is 0.23, the standard deviation is about 0.18. We then examine an array of variables that are likely determinants of the cash-flow correlations, discount-rate correlations, and systematic behavioral biases to explain the variations in pair-wise return correlations. Our analysis includes a long list of the following variables to explain the pair-wise stock comovement: pair-wise earnings correlations, industry membership, sales growth, size and book-to-market ratios, location, firm age, exchange membership, S&P Index membership, price level, durations of the expected future cash flows, financial leverage, upstream-downstream industry, and abnormal trading volume. We find that most of the above factors do explain pair-wise return comovement, but 85% of the variation in the total return comovement and 90% of the variation in idiosyncratic return comovement is not explained by these considerations. This suggests either that there is significant idiosyncratic variation in stock comovement or that there are systematic factors beyond those examined in the literature that explain variations in return comovement.

The second question we attempt to address is whether firms that comove in one period also comove in the next period. If two stocks' returns are highly correlated in the last five years, they are likely to be exposed to the same cash-flow and discount-rate shocks – we therefore hypothesize that pair-wise stock return comovement should be persistent over time. We find evidence in support of this hypothesis: In a univariate regression of the future return difference between two stocks on their historical return correlation, we find a negative slope coefficient, suggesting that past return comovement predicts future comovement.

More interestingly, we examine whether the unexplained portion of the return comovement is persistent. If the 85 percent of the return comovement that cannot be explained by common variables is transitory, then it's likely due to noise trading, estimation errors in our empirical implementation, or some other non-systematic factors. If, on the other hand, this unexplained variation in return comovement is also persistent, then the evidence suggests that there's information in return comovement that is not fully captured by the common variables. We regress future return divergence of a pair of stocks on its historical return correlations controlling for factors that are possible determinants of the comovement. The results indicate that the residual return comovement is negatively related to future return divergence and this effect is statistically significant. This evidence suggests that there's substantial information content in the residual return comovement.

Finally, we test whether the information contained in the return comovement is fully impounded into prices by exploring whether the divergence between the returns of two stocks which have comoved historically implies a convergence in their future returns. We identify a group of comover stocks that tend to move together with a given stock in the last five years. If this stock's return deviates from its comover stocks in a given month, we examine whether its returns converges to its comover peers in the future and provides potential trading opportunity.

We find that a trading strategy that bets on this convergence generates 5-factor alphas of up to 36% annually for an equal-weighted self-financing portfolio and 11% for a value-weighted portfolio. It represents a significant improvement over the simple short term reversal strategy documented by Jegadeesh (1990) and extends the evidence on pairs trading strategy documented in Gatev, Goetzmann, and Rouwenhorst (2006), who match stocks using minimum distance between normalized historical prices and find a trading rule that yields abnormal trading returns.

This paper fits in the literature on stock return comovement. Prior research has shown that firms with similar size and book-to-market ratios (Fama and French (1993), Boyer (2007)), index membership (Barberis, Shleifer, and Wurgler (2005), Greenwood (2007)), geographic location (Pirinsky and Wang (2006)), price level (Green and Hwang (2008)), and retail investor trading behavior (Kumar and Lee (2006)) move together in stock price. Shiller (1989) argues that the comovement between U.K. and U.S. stock prices is too large to be fully explained by comovement in dividends. Campbell and Mei (1993) examine the sources of betas, and break assets' betas with common factors into components attributable to news about future cash flows, real interest rates and excess returns. These papers mainly rely on the portfolio approach to identify the return comovement factors — as a result, they have to limit the analysis to one or two factors that may affect return comovement. In this paper, we focus on firm-level pair-wise return comovement and this allows us to examine an exhaustive list of factors that may drive the comovement. By doing so, we want to identify the variations in return comovement that can (or cannot) be explained by common variables, including those identified in the prior research.

This paper is also related to the literature on excess comovement of stock returns. This literature (e.g., Kallberg and Pasquaiello (2007)) finds that there's excess comovement between stocks at industry level, i.e., there are comovement in the stock returns of different industries that are not readily explainable by common fundamental factors. Papers in this literature generally adopt a portfolio approach and study the comovement between industry- or country-level return comovement. Our approach complements this line of research by focusing on firm-level pair-wise comovement, which enables us to examine a larger variation in the comovement and a much richer set of possible determinants. We do not directly address the issue whether there is excess comovement in returns. Rather, our goal is to provide an empirical assessment of the degree of common variables to explain return comovement — our evidence supports the existence of substantial information content in return comovement and suggests that it's not dominated by idiosyncratic noises.

Finally, our paper is also closely related to the literature on pairs trading. We extend the findings in Gatev, Goetzmann, and Rouwenhorst (2006), who show that there are abnormal returns from a return-based pair-wise relative value trading strategy. The advantage of our empirical approach is that it provides tradeable portfolio opportunities every month, so that the capital can be fully invested all the time. Our approach also produces equal weighted returns of 1.70% and 5-facotr alpha of 1.36% per month for the same universe of stocks between July 1962 and December 2002 that are used by Gatev, Goetzmann, and Rouwenhorst (2006), that produce raw returns of 0.72% and 5-factor alpha of 0.51% per month. The improvement of our strategy is partially due to the construction of the diversified comover portfolios of 50 stocks.

The rest of the paper proceeds as follows. Section II explores the determinants of pair-wise stock correlations. Section III examines the persistence of the correlations and Section IV focuses on the abnormal future returns for portfolios formed on return comovement. Section V concludes.

II. Determinants of Pair-Wise Stock Return Correlations

We first estimate the correlation between any two monthly stock return series from year t-4 to year t using CRSP data. For n stocks, there are $n*(n-1)/2$ correlations. Because the number of the observations for the correlations grows exponentially in the number of stocks, the work is computationally intensive. To reduce the computation burden, we require firms to have 60 monthly stock returns data from year t-4 to year t. We also require firms to have market equity of at least $500 million.

For the determinants of the comovement, we consider an array of variables that are potentially related to the comovement of firms' cash flows and discount rates. We also consider variables that are related to investors' trading behavior that could also cause stock comovement (e.g., Barberis, Shleifer, and Wurgler (2005)). More specifically, the variables we consider include:

Earnings correlation. If the earnings of two stocks exhibit a strong correlation, their stock returns should also exhibit a strong correlation. We compute the earnings correlations as follows. We use data from Compustat monthly Price, Dividends, and

Earnings file. For each quarter, we construct the return on equity *ROE* as ratio of the 12 month earnings per share to the book value of equity per share (ERN over BKV). To mitigate the influence of outliers, we set *ROEs* that are greater than 10, or less than -10, to be 10 and -10, respectively. We then compute the correlation between 20 quarterly *ROE* series of any two stocks, between year *t-4* and year *t*. We call this variable earnings correlation, *Earncorr*. We require firms to have all 20 quarterly *ROE* data between year *t-4* and year *t* to be included in our sample. Alternatively, we construct correlations for changes in *ROEs*, and we call that variable *Earncorr_ch*.

Growth. While earnings correlation captures how historical earnings comove between two stocks, firm growth captures the expected future earnings. We compute five-year sales growth for each firm-year and calculate the absolute value of difference in log growth rates to measure the growth difference of two firms.

Industry. If firms are in the same industry, they may face similar business conditions and therefore are likely to share cash-flow and discount-rate shocks. We construct dummy variables for a stock pair if they are in the same industry. For any two stocks, if they have the same 4-digit SIC industry code, we construct a dummy variable *Ssic4* to be 1, and 0 otherwise. Similarly we set dummy variables *Ssic3*, *Ssic2*, *Ssic1* to be 1 for stocks that have the same 3-digit SIC codes, 2-digit SIC codes, or 1-digit SIC codes.

Size. If firms are of similar size, they may have similar exposure to risk factors (for example, Fama and French (1993)), and therefore have similar expected returns. Alternatively, investors may categorize assets into different styles and move funds among these styles depending on their relative performance, as suggested by Barberis and Shleifer

(2003). Because size is one common style, investors' trading behavior can induce stocks of similar size to move together. For market value of equity, we use Compustat total shares outstanding multiplied by the fiscal year end price (25*199). Size is the logarithm of the market value of equity. We construct the absolute value difference in size *Dsize*.

Book-to-market ratios. If two firms have similar book-to-market ratios, they may have similar exposure to some fundamental risks; alternatively, investors may view stocks with similar book-to-market ratios as having the same style (Fama and French (1993) and Boyer (2007)). We construct book value of equity as total assets minus total liabilities (6-181). Book-to-market equity is then the ratio of the book equity to the market value of equity. We construct the absolute value difference in log book-to-market ratios *Dlogbtm*.

Geographic location. If firms are located in geographically proximate areas, they may be subject to some common shocks and therefore similar returns (e.g., Pirinsky and Wang (2005)). We construct a same State dummy variable *Sstate* to be 1, if two firms are headquartered in the same state, and 0 otherwise. We use the historical SIC code supplied by CRSP for this task.

Firm age. Firms with similar ages are likely to be in similar stages of life cycle, and therefore have comovement. We construct firm age as the difference between the current year and the first year that the firm appears in CRSP plus 1. We then construct the absolute value difference between logarithm of ages of two firms *i*, and *j*, as $Dage = |\log(Age_i) - \log(Age_j)|$.

Exchange listing. If two stocks are on the same exchange, similar market microstructure issues may create stock return comovement. We set the variable *Slisting* to be 1, if two stocks are listed in the same stock exchange, and 0 otherwise.

S&P Index. Two stocks belonging to the same S&P major Index (Utilities, Transportation, Financial Index) may move together (e.g., Barberis, Shleifer and Wurgler (2005)). We create a dummy variable *Sindex* that equals 1 if two firms belong to the same S&P major index, S&P mid cap index or S&P small cap index in a given year, and 0 otherwise.

Price difference. Prior literature has shown that firms with similar share prices comove together (see Green and Hwang (2008)). We therefore construct a variable *Dprice* as the absolute difference in the log price at the end of a given year.

Duration. Firms with similar equity duration are likely to move in similar fashions. We follow Dechow, Sloan, and Soliman (2004) and construct a variable that measures the difference of duration of the cash flows of two stocks, *Dduration*.

Financial leverage. Firms with similar financial leverage ratios (Long Term Debt/Total Assets) may respond to changes in economic conditions in similar ways. We construct the absolute value difference in two firms' financial leverage ratios, *DLeverage*.

Upstream-downstream industry. Industries that have customer-supplier relations often benefit from one another's survival. Successes in the customers should lead to more order flows to the suppliers. If the market fully incorporates this information, then good news for the customer should be interpreted as good news for the supplier as well, thus creating comovement between the customer and the supplier. We follow Menzly and Ozbas (2006), and use the information from Bureau of Economic Analysis' Input-Output Benchmark Survey. We construct a dummy variable *Indlink* to be 1 if two firm's industries have customer-supplier relation, and 0 otherwise.

Trading volume. We calculate the abnormal trading volume of a given stock using the residual from a regression of monthly trading volume on year trend and monthly dummies with data from the last 36 months. The rationale for examining trading volume is that it's likely to be related to liquidity factors and investor trading behavior that may affect return comovement. For instance, Kumar and Lee (2006) find that firms with similar retail investor trading patterns tend to comove. We therefore construct the correlation between two firm's abnormal trading volumes *VolumeCorr*.

Table II provides the descriptive statistics of the pair-wise return correlations. Our sample consists of almost 3 million pair-years of stocks over the sample period. On average, the Pearson correlation between any two stocks' monthly returns is 0.23 and the standard deviation of this correlation is 0.18. The most negative correlation between any two stocks is -0.583 and the most positive is 0.964.

The mean value of *Ssic1* is 0.18, indicating that 18 percent of our sample pairs are from the same one-digit SIC code industry. About 5 percent of the firm-pairs are from the same state, as suggested by the mean *Sstate* of 0.053. The mean value of *Slisting* (0.76) suggests that about 76 percent of our sample pairs are from the same exchange. About 48% of the stock pairs belong to the same S&P 500, S&P mid-cap, or S&P small cap indices.

We estimate the regression of return correlations on the variables that are hypothesized to predict comovement. In Table III, we present the results for the Fama-

MacBeth regression of stock return correlations on the determinants using data from 1976 to 2005. Because return correlations are auto-correlated by construction, we use Newey-West procedure to compute the standard errors.[1]

Consistent with the univariate correlations, the variables do explain the pair-wise return correlations in ways as expected and almost all the variables have statistically significant coefficients. In column (1), both earnings correlation and correlation of changes in earnings load up statistically significant with t-statistics of 5.06 and 3.52, respectively. However, the magnitudes seems small 0.031 and 0.014 — given both the dependent variable and independent variable are correlation coefficients, the perfect model would produce coefficients much closer to 1. The average R-squared of the regression is only 2.85%. This indicates that pair-wise earnings comovement and growth difference can explain less than 3% of the total variation in the pair-wise returns correlations.

In column (2), the four same-industry dummies (*ssic1* to ssic4) are also included in the regression. All of the industry dummies show up positive and statistically significant. For instance, the coefficient on *ssic4* is 0.054 (t=4.11), which suggests that, if two firms are within the same four-digit SIC industry, then compared with a pair of firms that are in the same three-digit SIC industry (but not in the same four-digit SIC industry), their return correlation is higher by 0.054. Compared with two firms that do not share any industry membership, pair of firms that are in the same four-digit SIC industry will have a return correlation coefficient higher by 0.223 (0.054+0.064+0.058+0.047). The R-squared also

[1] To mitigate the concerns of within-pairs correlations, we also adjusted the standard errors by clustering at the pair level, and the results remain qualitatively similar with mostly higher statistical significance.

increases dramatically when the industry dummies are included—the new R-squared is 8.74%.

In column (4), we further include the distances in firm size and book-to-market ratio (*Dsize* and *Dlogbtm*), the same-state dummy (*Sstate*), the distance in firm age (*Dage*), the same exchange listing dummy (*Slisting*), the same S&P index membership (*Sindex*), the distance in log prices (*Dprice*), the distance in durations (*Dduration*), the distance in financial leverage ratios (*Dleverage*), Industrial customer-supplier link (*Indlink*), and Abnormal volume correlations (*Volumecorr*) to explain return comovement. Not surprisingly, the further the distance in size, book-to-market, and price, duration, financial leverage, the lower the correlation in stock returns of any two firms. Everything else equal, two firms headquartered in the same state, listed in the same exchange list, in the same stock index, and with similar trading volume patterns tend to comove in stock returns. The only two factors that show up statistically insignificant are *Dage*, the absolute value difference in log age, and *Indlink*, the industrial customer-supplier link.

Overall, even when all the determinants are included in column (4), there is a lot of variation in the return comovement that cannot be explained—when we include all the explanatory variables, the average R^2 is still less than 15%.[2] The interesting question, then, is whether the 85 percent of unexplained variations is largely noise or something systematic that is not easily incorporated in the common economics variables.

We also estimate idiosyncratic stock returns by calculating the residual returns from a five-factor asset-pricing model (the market, size, and book-to-market factors documented

[2] One concern is that our pair-wise correlations are redundant when the number of stocks n is much larger than the time period T, in that $n*T$ return observations generate $n*(n-1)/2$ pair-wise correlations. To address this concern, we carefully choose $n*T$ non-redundant correlations, and the results remain similar.

in Fama and French (1993), the momentum factor based on Jegadeesh and Titman (1993), and the liquidity factor in Pastor and Stambaugh (2003)) and calculate the pair-wise correlations of the residual returns. The idiosyncratic returns filter out many of the factors that are known in the literature to drive return comovement and focusing on the idiosyncratic returns comovement allows us to examine other variables.

Table IV shows the regression of the idiosyncratic return comovement on the determinants. A puzzling result is that in Column (3), size and book-to-market still significantly affect the idiosyncratic return comovement, even though the systematic components of stock returns related to size and book-to-market are supposedly filtered out with the five-factor model. The results indicate that most factors that affect the pair-wise total return comovement also impact the residual return comovement in a similar fashion.

III. Persistence of Return Comovement

If two stocks' returns are highly correlated, they are likely to be exposed to cash-flow and discount-rate shocks in similar ways. We hypothesize that stocks that are highly correlated historically have similar realized returns in the future.

Specifically we do the following test:

$$\left| \frac{1}{\sigma_{i,t}} r_{i,t+1} - \frac{1}{\sigma_{j,t}} r_{j,t+1} \right| = a + b_1 * corr_{i,j,t} + e_{t+1},$$

where $corr_{i,j,t}$ is the correlation coefficient in returns between firm i and firm j estimated using data between t-4 and t, $r_{i,t+1}$ is the realized return of firm i in period $t+1$,

$\sigma_{i,t}$ is the standard deviation of the stock returns of firm *i* estimated using data between *t-4* and *t*, and |.| is the operator for absolute value.

We include the ratio of the return volatility when calculating the absolute value of the differences in realized future returns to adjust for the scale differences in returns. The logic for including the volatility ratio adjustment is the following: two stocks may be perfectly correlated but may have different returns. For example, stock A's return is always twice of stock B's return. Adjusting for their volatility ratio helps mitigate such issues.[3] Our prediction is that the coefficient b_1 should be negative since stocks that are highly correlated have similar discount-rate exposure and cash-flow shocks, and thus similar realized return.

The results are presented in Table V. In the first column, the coefficient on *Corr* is -0.324 with a t-value of -7.70. This is consistent with the argument that return comovement is persistent and historical return correlations are systematically related to future return divergence.

We then examine whether the residual return comovement (i.e., the 85 percent of the variations in comovement that is not explained by the common variables in Table III) is associated with future return differences by estimating the following equation:

$$\left| \frac{1}{\sigma_{i,t}} r_{i,t+1} - \frac{1}{\sigma_{j,t}} r_{j,t+1} \right| = a + b_1 * corr_{i,j,t} + b * Control_{i,i,t} + e_{i,j,t+1}$$

The control variables include differences in the common determinants of stock returns, such as difference in size, difference in book-to-market ratios for each pair of

[3] However, the empirical results are similar, if we do not include the volatility ratio adjustment, i.e., if we use $| r_{i,t+1} - r_{j,t+1} |$ as the dependent variable.

stocks. The timing for constructing the variables is the following: for monthly observations from July of year *t+1* to June of year *t+2*, we match with the correlation computed with data from January of year *t-4* to December of year *t*. The other explanatory variables are for the fiscal years that end in year *t*.

The Fama-MacBeth regression results are reported in Table V columns (2) to (4). Cross sectional regression is run for each month between July 1977 and June 2007. The average coefficients and Newey-West adjusted t-statistics are reported.

With all the control variables included, *Corr* is still negatively and significantly associated with the return divergence in the future. For instance, in column (4), the coefficient on *Corr* is -0.264 (t-statistic=-7.13). This indicates that the unexplained portion of the pair-wise return comovement is unlikely to be simply noise or estimation error, but rather has substantial information content about future returns.

The interpretation of the results on the control variables is difficult, because the return correlation *Corr* is already included in the regression and the coefficients on the control variables are the residual effect. Of the control variables, many are not statistically significant, with industry memberships being the most notable exceptions. Conditional on current returns comovement, earnings correlation and abnormal volume correlation do not have any implications for future returns: the coefficient on *Earncorr* is -0.004 with a t-value of -1.03, and the coefficient on *Volumecorr* is 0.001 with a t-value of 0.05.

IV. Profitability of a Comovement-based Trading Strategy

In this section, we explore a trading strategy based on the historical return comovement. Essentially, this test examines whether the information contained in stock comovement is fully impounded into the prices.

Our comovement based trading hypothesis is that if in the previous month, a stock's return, *Lret,* deviates from its comover portfolio returns, *Cret*, then in the following month, this deviation should be reversed, i.e., the stock should experience abnormally higher returns the next month. We identify the comover portfolio in the following way. For each stock *i* in year *t+1*, we compute the Pearson correlation coefficients between the stock returns of stock *i* and all other stocks using monthly return data from January of year *t-4* to December of year *t*. We then find 50 stocks that have the highest correlations with this stock. In each month of year *t+1*, we compute the comover portfolio return as the equal weighted average return of the 50 comover stocks, *Cret*.

In Table VI, we report the value weighted and equal weighted portfolio returns based on a double sort of previous month's stock return and comover portfolio return. The holding period is also one month. Consistent with the findings in Fama (1965) and Jegadeesh (1990), stocks exhibit a short term reversal: stocks with low returns in the previous month have high returns in the current month. In all the columns, the portfolio returns in the next month decrease monotonically from the first row (i.e., firms with low lag returns) to the fifth row (i.e., firms with high lag returns). For instance, conditional on having low lag comover return, the average return of firms with low lag return is 0.93 percent per month, whereas the return of firms with high lag returns is only 0.42 percent.

We also find that given the stock's lagged return, if the comover portfolio has higher return in the previous month, then this stock is likely to have high return this month. In other words, stocks that deviated from their comovers tend to reverse that deviation. This effect is also monotonic — in every row of Table VI, the portfolio returns increase monotonically from column 1 to column 5 when the lag comover returns increase. A portfolio that longs stocks with high lag comover returns and shorts those with low lag comover returns generates positive returns for each row (i.e., within each quintile portfolios sorted on lag firm returns).

Table VII reports the Fama-MacBeth regressions of monthly returns on previous month's comover return *Cret*, previous month's own return *Lret*, and other control variables such as size and log book-to-market ratios. For returns between July of year *t+1* and June of year *t+2*, we match with Size and book-to-market-equity at the fiscal year end in year *t*. *Price* is the logarithm of price in the previous month. Variables are demeaned before forming interaction terms. All regressions are for the sample period July 1951 to December 2007.

We find that Comover portfolio return *Cret* predicts returns positively and lagged return *Lret* predicts return negatively. In column (1), the coefficient on *Cret* is 0.228 (t=12.93) and that on *Lret* is -0.069 (t=-17.46). The fact that *Cret* is statistically significant even when *Lret* is included in explaining future returns suggests that it's not driven by the simple short-term reversal story. In column (2), we further include size and book-to-market in the regression. The coefficients on *Cret* and *Lret* both remain statistically significant.

We then interact *size* and *Price* with *Cret* and *Lret* to gauge the cross-sectional differences of the effects for firms with different size and stock price levels. The evidence in columns (3) to (5) shows that the effects are more pronounced for small stocks and low price stocks, as indicated by the negative (positive) coefficient on the interaction term of *Size* and *Price* and *Cret* (*Lret*). This evidence indicates that both the short-term reversal and comover trading profits are more pronounced for smaller firms and firms with lower price, which can be due to greater bid-ask bounces or the slower information diffusion of these less liquid firms.

In columns (6) to (9), we construct a new variable *RetDiff* as the scaled return difference to combine the effect of comover portfolio return and the lag return. It is $beta^{C}*(Cret\text{-}Rf) - (Lret - Rf)$, where $beta^{C}$ is the regression coefficient of a firm's monthly return on its comover portfolio return in the most recent 5 years.[4] The use of $beta^{C}$ addresses the issue of different return volatilities between any two stocks. The regression coefficient of future returns on *RetDiff* captures the comovement-based trading strategy's predicting power.

Not surprisingly, we find *RetDiff* positively predict next month's return and the effect is highly statistically significant, even after controlling for size and book-to-market. In column (7), the coefficient on *RetDiff* is 0.082 with a t-statistic of 17.67. The effect of *RetDiff* is more pronounced in low price stocks and small stocks.

Table VIII reports the portfolio returns based on time-series tests if we use the scaled return difference *RetDiff* as the single sorting variable. In each month, we form 10

[4] Alternatively, we can construct the simple return difference as *Cret-Lret*. The empirical results based on this specification are similar, with slightly lower magnitude.

portfolios, Decile 1 through Decile 10 based on previous month's *RetDiff*. Our holding period is one month. We now use all stocks that had 60 months of returns in the most recent 5 years. Our sample period is from January 1931 to December 2007. In Panel A, we report raw returns, Fama-French 3 factor alphas, and 5 factor alphas for the value weighted portfolios, where the 4th and 5th factors are momentum factor and short term reversal factor. [5] An examination of the raw returns and alphas of the decile portfolios shows that stocks with high *RetDiff* (high previous month's comover return and low own stock return), have higher subsequent returns. For the value weighted portfolios, the zero-cost portfolio Decile 10 – Decile 1 (i.e., longing Decile 10 and shorting Decile 1) generates a return of 1.40% per month (t=9.28). The hedge portfolio has a three-factor adjusted alpha of 1.23% with a t-value of 8.32 and five-factor adjusted alpha of 0.91% (t=6.61). In addition to the significant hedge portfolio alphas, the alphas increase almost monotonically from Decile 1 to Decile 10, indicating that sorting on *RetDiff* systematically drives the hedge portfolio returns.

The equal weighted portfolios generate even higher dispersion in returns. Panel B of Table VIII reports the raw returns, 3-factor alphas and 5-factor alphas for equal weighted portfolios sorted by *RetDiff*. For the equal weighted portfolios, the zero-cost portfolio Decile 10 – Decile 1 (i.e., longing Decile 10 and shorting Decile 1) generates a return of

[5] The short term reversal factor is provided by Kenneth French. Six value-weight portfolios are formed on size and prior (t-1) returns to construct ST_Rev. The portfolios, which are formed monthly, are the intersections of 2 portfolios formed on size (market equity, ME) and 3 portfolios formed on prior (t-1) return. The monthly size breakpoint is the median NYSE market equity. The monthly prior (t-1) return breakpoints are the 30th and 70th NYSE percentiles. ST_Rev is the average return on the two low prior return portfolios minus the average return on the two high prior return portfolios, *ST_Rev=1/2 (Small Low + Big Low) - 1/2(Small High + Big High).*

3.59% per month (t=18.69). The three-factor alpha for the self-financing portfolio is 3.17% per month (t=18.30). The five-factor alpha is 3.00% (t=17.76).

In Panel C of Table VIII, we provide another robustness check of our results. Our findings could be driven by market microstructure issue: thinly traded stocks are simply switching back and forth between the bid and ask prices and the return data for these stocks reflect these bounces. To further examine this issue, we look at the performance of the strategy on the high price stocks, among which the bid ask bounce is less of a concern. Panel C reports the profitability of this trading strategy for stocks with a price over $5. The alphas of the value weighted hedge portfolio become smaller but remain economically and statistically significant — the three factor alpha remains highly significant at 1.08% per month (t=7.60) and the five-factor alpha is now 0.75% per month (t=5.74).

For an example of this trading strategy, we return to Whirlpool. As tabulated in Table I, we identify the 50 stocks that move the most closely with Whirlpool between 2001 and 2005 and Figure 1 plots the returns of Whirlpool and its comover portfolio between 2001 and February 2006. From the plot, it's easy to see that Whirlpool and its comovers are almost perfectly synchronized in returns between 2001 and 2005. In January 2006, however, the comover portfolio returned 3.6% while Whirlpool lost 3.7%. In February 2006, Whirlpool more than made it up by returning a whopping 11.3% while its comover portfolio stayed flat at 0.7%.

To examine the common factor-loadings characteristics of the comovement-based decile portfolios, Table IX reports the loadings of these portfolios with respect to the 5 factors. For the value weighted portfolios, the self-financing portfolio (Decile 10 – Decile 1) loads positively on the market, and very strongly on the short term reversal factor, but loads negatively on momentum. Its loading on SMB, HML are both economically and statistically insignificant. For the equal weighted portfolios, the self-financing portfolio loads positively on the market, SMB, HML and especially short term reversal. Again it loads negatively on the momentum factor.

Our trading strategy is closely related to the pairs trading strategy in Gatev, Goetzmann, and Rouwenhorst (2006, in the rest of this section referred to as GGR). GGR identify pairs by minimizing the sum of squared deviations between two normalized price series in the previous 12 months. A position is opened within the next 6 months the day after when prices diverge by more than two historical standard deviations. The position is unwound at the next crossing of the prices, or at the end of the 6 month period. They report 0.72% per month for the equal weighted hedge portfolio, with 5 factor alpha of 0.51% per month.

Our trading strategy is an improvement of the pairs trading strategy in that we provide tradeable opportunities at any given point of time. In GGR, at any given point of time, the number of pairs traded are unknown in advance and it's not clear how much capital should be invested in each pair. On top of this ease of implementability, our strategy also provides much higher trading profits.

In Panel A of Table X, we try to implement our trading strategy on GGR's sample to make these two strategies comparable. Following them, we only trade stocks that have positive trading volumes on each day of the previous year. We use the sample period of July 1962 to December 2002. We also skip a day between the one-month formation period and the one-month holding period. We find that our equal weighted hedge portfolio generate raw average returns of 1.70% per month, and 5 factor alpha of 1.36%, more than twice than those in pairs trading.

We believe one of the reasons that our trading strategy performs better is that we use a diversified comover portfolio of 50 stocks, while GGR rely on one matching stock to carry out their pairs trading strategy. In Panel B of Table X, we test whether using only one stock to form the comover portfolio is sufficient for the comovement-based strategy. We go back to our sample of January 1931 to December 2007 and value weighted methodology to be comparable to our earlier results. The results indicate that trading strategy is still profitable with a five-factor alpha of 0.43% (t=3.70). However, it is less than half of our baseline strategy's alpha of 0.91%. We suspect that the idiosyncratic component of one comover stock return is too high for it to provide a better benchmark for what a stock's return should have been.

We also suspect that another reason for our improved performance may lie on how we identify comovers or pairs. In our strategy, we use the stocks with high correlations. In GGR, pairs are determined by minimizing the sum of squared deviations between two normalized price series in the previous 12 months, and thus all pairs are co-integrated. Our approach allows us to get useful information from pairs with high correlation, but perhaps with very different volatilities. For example, if two stocks are perfectly correlated, but one stock return is always twice of the other stock (perhaps driven by difference in financial leverage ratios), these two stocks would not be considered pairs in GGR. They would be comovers in our strategy.

V. Conclusions

In this paper, we embark on the first systematic study of firm-level pair-wise stock return correlations. We find that while common variables such as cash-flow comovement, industry affiliation, size, and book-to-market are important determinants of return comovement, more than 85 percent of the variation in return correlations is not explained by an exhaustive list of existing explanatory variables. We also find that comovement and the unexplained part of comovement are persistent. Finally, we extend the results in Gatev, Goetzmann, and Rouwenhorst (2006) by showing that a comovement-based strategy can generate significant abnormal returns and this abnormal return is incremental to the simple short-term reversal strategy.

Reference

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

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

Boyer, H. Brian, 2007, Comovement among stocks with similar book-to-market ratios, working paper, Brigham Young University.

Campbell, John Y., and Jianping Mei, 1993, Where do betas come from? Asset price dynamics and the sources of systematic risk, *Review of Financial Studies*, Vol. 6, No. 3, pp. 567-592.

Dechow, Patricia, Richard Sloan, and Mark Soliman, 2004, Implied Equity Duration: A New Measure of Equity Risk, *Review of Accounting Studies*.

Fama, Eugene F., 1965, The behavior of stock market prices, *Journal of Business* 38, 34-105.

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.

Gatev, Evan, William N. Goetzmann, and K. Geert Rouwenhorst, 2006, Pairs Trading: Performance of a Relative Value Arbitrage Rule, *Review of Financial Studies* 19, 797-827.

Green, T. Clifton, and Byoung-Hyoun Hwang, 2008, Price-Based Return Comovement, Working Paper, Emory University.

Greenwood, Robin, 2007, Excess Comovement: Evidence from cross-sectional variation in Nikkei 225 weights, *Review of Financial Studies*, forthcoming.

Jegadeesh, Narasimhan, 1990, Evidence of Predictable Behavior of Security Returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance*, 65- 91.

Kallberg, Jarl and Paolo Pasquariello, 2007, Time-series and cross-sectional excess comovement in stock indexes, *Journal of Empirical Finance*.

Kumar, Alok, and Charles M.C. Lee, 2006, Retail Investor Sentiment and Return Comovements, *Journal of Finance* 61: 2451-2486.

Menzly, Lior, and Oguzhan Ozbas, 2006, Cross-Industry Momentum, working paper, University of Southern California.

Pastor, Lubos, and Robert Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy* 111, 642-685.

Pirinsky, Christo A. and Qinghai Wang, 2005, Does Corporate Headquarters Location Matter for Stock Returns?, *Journal of Finance*, forthcoming.

Shiller, Robert, 1989, Comovements in Stock Prices and Comovements in Dividends, *Journal of Finance* 44: 719-729.

Table I: Comovers of Whirlpool

This table reports the 50 stocks that have the highest return correlations with Whirlpool Corporation (ticker: WHR, CRSP permno: 25419, SIC: 3633) between January 2001 and December 2005. Return correlations are estimated using monthly return observations. A stock is required to have 60 valid return observations to be included in the sample.

	Ticker	Company Name	SIC	Correlation
1	CAL	CONTINENTAL AIRLINES INC	4512	0.736
2	NAV	NAVISTAR INTERNATIONAL CORP	3713	0.722
3	BWA	BORGWARNER INC	5013	0.706
4	ALK	ALASKA AIRGROUP INC	4511	0.696
5	JNS	JANUS CAP GROUP INC	6282	0.692
6	AVX	A V X CORP NEW	3675	0.682
7	AVT	AVNET INC	3679	0.680
8	WDR	WADDELL & REED FINANCIAL INC	6211	0.673
9	AMR	A M R CORP DEL	4511	0.663
10	DOV	DOVER CORP	3534	0.657
11	ITW	ILLINOIS TOOL WORKS INC	3565	0.656
12	ESST	E S S TECHNOLOGY INC	3670	0.655
13	VVI	VIAD CORP	7389	0.653
14	TSG	SABRE GROUP HOLDINGS INC	7375	0.653
15	HSC	HARSCO CORP	3443	0.652
16	SCUR	SECURE COMPUTING CORP	7372	0.651
17	TGT	TARGET CORP	5331	0.649
18	CDNS	CADENCE DESIGN SYSTEMS INC	7372	0.648
19	BDC	BELDEN C D T INC	3357	0.647
20	GPI	GROUP 1 AUTOMOTIVE INC	5511	0.647
21	HIG	HARTFORD FINANCIAL SVCS GROUP	6411	0.647
22	MYG	MAYTAG CORP	3639	0.642
23	RSYS	RADISYS CORP	3674	0.642
24	FD	FEDERATED DEPT STORES INC DEL	5311	0.639
25	MESA	MESA AIR GROUP INC NEV	4510	0.637
26	DGII	DIGI INTERNATIONAL INC	3577	0.635
27	ETN	EATON CORP	3714	0.635
28	LEA	LEAR CORP	3714	0.635
29	HON	HONEYWELL INTERNATIONAL INC	3724	0.633
30	TROW	T ROWE PRICE GROUP INC	6211	0.633
31	PHTN	PHOTON DYNAMICS INC	3550	0.632
32	URI	UNITED RENTALS INC	7353	0.632
33	AMG	AFFILIATED MANAGERS GROUP INC	6282	0.630
34	CTB	COOPER TIRE & RUBBER CO	3011	0.630
35	RHI	ROBERT HALF INTERNATIONAL INC	7363	0.630
36	SWK	STANLEY WORKS	3429	0.629
37	JDAS	J D A SOFTWARE GROUP INC	7370	0.629
38	CD	CENDANT CORP	6531	0.628
39	GBL	GAMCO INVESTORS INC	6282	0.627
40	TFX	TELEFLEX INC	3841	0.627
41	MAR	MARRIOTT INTERNATIONAL INC NEW	7011	0.627

42	AN	AUTONATION INC DEL	5511	0.627
43	PCP	PRECISION CASTPARTS CORP	3324	0.626
44	WY	WEYERHAEUSER CO	2421	0.626
45	NTRS	NORTHERN TRUST CORP	6020	0.622
46	VSH	VISHAY INTERTECHNOLOGY INC	3676	0.622
47	NFS	NATIONWIDE FINANCIAL SERVICES IN	6311	0.621
48	CAT	CATERPILLAR INC	3531	0.620
49	GT	GOODYEAR TIRE & RUBR CO	3011	0.619
50	SUNW	SUN MICROSYSTEMS INC	3570	0.619

Table II: Summary Statistics

This table reports the summary statistics for the determinants of stock return correlations between 1976 and 2005. Each observation represents a stock pair year. *Corr* is the stock return correlation between two stocks using monthly returns between year *t-4* and *t*. Earnings correlation (*Earncorr*) is calculated from Compustat Price, Dividends, and Earnings dataset as follows. In each quarter, we construct the return on equity *ROE* as the ratio of the 12 month earnings per share to the book value of equity per share (ERN over BKV). To mitigate outliers, we set *ROEs* that are greater than 10, or less than -10, to be 10 and -10, respectively. We then compute the correlation between 20 quarterly *ROE* of any two stocks A and B, between year *t-4* and year *t*. *Earncorr_ch* is the correlation between two firms' changes in *ROE*s. *Dgrowth* is the absolute value difference in 5-year log sales growth rates. For any two stocks A and B, if they have the same 4-digit SIC industry code, we construct a dummy variable *Ssic4* to be 1, and 0 otherwise. Similarly we set dummy variables *Ssic3*, *Ssic2*, *Ssic1* to be 1 for stocks that have the same 3-digit SIC codes, 2-digit SIC codes, 1-digit SIC codes, respectively. Market equity is the product of Compustat total shares outstanding and the fiscal year end price (25*199). Size is the logarithm of the market equity at the fiscal year end in year *t*. *Dsize* is the absolute value difference in size. The book value of equity is the total assets minus total liabilities. Book-to-market equity is then the ratio of the book equity to the market value of equity at the fiscal year end in year *t*. *Dlogbtm* is the absolute value difference in log book-to-market ratios. *Sstate* is a dummy variable that is 1, if two firms are located in the same state, and 0 otherwise. Firm age is the difference between the current year and the first year that a firm appears in CRSP plus 1. *Dage* is the absolute value difference between logarithm of ages of two firms *i*, and *j*, $|\log(Age_i)-\log(Age_j)|$. *Slisting* is 1 if two stocks are listed on the same exchange, and 0 otherwise. *Sindex* is 1 if two stocks belong to the same S&P major, mid cap, or small cap indices, and 0 otherwise. *Dprice* is the absolute value difference in log prices per share. *Dduration* is the absolute value difference in cash flow durations. *Dleverage* is the absolute value difference in financial leverage ratios (long term debt/total assets). *Indlink* is 1 if two firms belong to industries that have customer-supplier links, and 0 otherwise. *Volumecorr* is the correlations between two firms' abnormal trading volumes. A stock is included in the sample only if it has 60 valid monthly returns, 20 quarterly earnings, and market equity of at least $500 million.

Variable	Observations	Mean	Std	Min	Max
year	2,824,654	1998	5.654	1976	2005
Corr	2,824,654	0.226	0.184	-0.583	0.964
Earncorr	2,824,654	0.066	0.447	-0.991	0.998
Earncorr_ch	2,824,654	0.050	0.385	-0.996	0.996
Dgrowth	2,824,654	0.413	0.379	0	5.607
Ssic4	2,824,654	0.012	0.111	0	1
Ssic3	2,824,654	0.020	0.141	0	1
Ssic2	2,824,654	0.049	0.216	0	1
Ssic1	2,824,654	0.177	0.382	0	1
Dsize	2,824,654	1.280	1.028	0	6.906
Dlogbtm	2,824,654	0.693	0.573	0	6.394
Sstate	2,824,654	0.053	0.224	0	1
Dage	2,824,654	0.655	0.489	0	2.315
Slisting	2,824,654	0.757	0.429	0	1
Sindex	2,824,654	0.484	0.500	0	1
Dprice	2,824,654	0.626	0.623	0	10.533
Dduration	2,824,654	1.693	1.801	0	41.699
Dleverage	2,824,654	0.157	0.119	0	0.903
Indlink	2,824,654	0.791	0.407	0	1.000
Volumecorr	2,824,654	0.096	0.198	-0.778	0.852

Table III: Determinants of Stock Return Correlations

This table reports Fama-Macbeth regression results of the pair-wise stock return Pearson correlation coefficient (*Corr*) on the determinants between 1976 and 2005. Each observation represents a stock pair year. *Corr* is the stock return correlation between two stocks using monthly returns between year *t-4* and *t*. Earnings correlation (*Earncorr*) is calculated from Compustat Price, Dividends, and Earnings dataset as follows. In each quarter, we construct the return on equity *ROE* as the ratio of the 12 month earnings per share to the book value of equity per share (ERN over BKV). To mitigate outliers, we set *ROEs* that are greater than 10, or less than -10, to be 10 and -10, respectively. We then compute the correlation between 20 quarterly *ROE* of any two stocks A and B, between year *t-4* and year *t*. *Earncorr_ch* is the correlation between two firms' changes in *ROE*s. *Dgrowth* is the absolute value difference in 5-year log sales growth rates. For any two stocks A and B, if they have the same 4-digit SIC industry code, we construct a dummy variable *Ssic4* to be 1, and 0 otherwise. Similarly we set dummy variables *Ssic3*, *Ssic2*, *Ssic1* to be 1 for stocks that have the same 3-digit SIC codes, 2-digit SIC codes, 1-digit SIC codes, respectively. Market equity is the product of Compustat total shares outstanding and the fiscal year end price (25*199). Size is the logarithm of the market equity at the fiscal year end in year *t*. *Dsize* is the absolute value difference in size. The book value of equity is the total assets minus total liabilities. Book-to-market equity is then the ratio of the book equity to the market value of equity at the fiscal year end in year *t*. *Dlogbtm* is the absolute value difference in log book-to-market ratios. *Sstate* is a dummy variable that is 1, if two firms are located in the same state, and 0 otherwise. Firm age is the difference between the current year and the first year that a firm appears in CRSP plus 1. *Dage* is the absolute value difference between logarithm of ages of two firms *i*, and *j*, $|\log(Age_i)-\log(Age_j)|$. *Slisting* is 1 if two stocks are listed on the same exchange, and 0 otherwise. *Sindex* is 1 if two stocks belong to the same S&P major, mid cap, or small cap indices, and 0 otherwise. *Dprice* is the absolute value difference in log prices per share. *Dduration* is the absolute value difference in cash flow durations. *Dleverage* is the absolute value difference in financial leverage ratios (long term debt/total assets). *Indlink* is 1 if two firms belong to industries that have customer-supplier links, and 0 otherwise. *Volumecorr* is the correlations between two firms' abnormal trading volumes. A stock is included in the sample only if it has 60 valid monthly returns, 20 quarterly earnings, and market equity of at least $500 million. Newey-West adjusted t-statistics are reported below the coefficients.

Dependent Variable: Return Correlation				
	1	2	3	4
Earncorr	0.031	0.027	0.026	0.023
	(5.06)	(5.42)	(5.02)	(6.15)
Earncorr_ch	0.014	0.012	0.012	0.010
	(3.52)	(2.86)	(2.9)	(2.46)
Dgrowth	-0.040	-0.031	-0.029	-0.029
	(-3.86)	(-2.88)	(-2.90)	(-2.69)
Ssic4		0.054	0.051	0.047
		(4.11)	(3.94)	(4.19)
Ssic3		0.064	0.063	0.061
		(5.96)	(5.97)	(6.49)
Ssic2		0.058	0.054	0.054
		(6.02)	(5.69)	(6.62)
Ssic1		0.047	0.048	0.040
		(7.94)	(8.18)	(6.54)
Dsize			-0.008	-0.008
			(-2.69)	(-2.50)
Dlogbtm			-0.024	-0.010
			(-5.31)	(-2.21)
Sstate				0.009

				(5.14)
Dage				0.001
				(0.25)
Slisting				0.022
				(2.65)
Sindex				0.023
				(4.70)
Dprice				-0.013
				(-3.97)
Dduration				-0.004
				(-4.49)
Dleverage				-0.079
				(-4.73)
Indlink				-0.015
				(-1.56)
Volumecorr				0.071
				(6.10)
Average Obs	94,155	94,155	94,155	94,155
Average R^2	2.85%	8.74%	10.02%	14.31%

Table IV: Determinants of Idiosyncratic Stock Return Correlations

This table reports Fama-Macbeth regression results of the pair-wise Pearson correlation coefficient of idiosyncratic stock return on the determinants between 1976 and 2004. Each observation represents a stock pair year. The idiosyncratic return is the residual from a regression of return on the Fama-French 3 factors, momentum factor and Pastor-Stambaugh liquidity factor using monthly returns between year *t-4* and *t*. Earnings correlation (*Earncorr*) is calculated from Compustat Price, Dividends, and Earnings dataset as follows. In each quarter, we construct the return on equity *ROE* as the ratio of the 12 month earnings per share to the book value of equity per share (ERN over BKV). To mitigate outliers, we set *ROEs* that are greater than 10, or less than -10, to be 10 and -10, respectively. We then compute the correlation between 20 quarterly *ROE* of any two stocks A and B, between year *t-4* and year *t*. *Earncorr_ch* is the correlation between two firms' changes in *ROE*s. *Dgrowth* is the absolute value difference in 5-year log sales growth rates. For any two stocks A and B, if they have the same 4-digit SIC industry code, we construct a dummy variable *Ssic4* to be 1, and 0 otherwise. Similarly we set dummy variables *Ssic3*, *Ssic2*, *Ssic1* to be 1 for stocks that have the same 3-digit SIC codes, 2-digit SIC codes, 1-digit SIC codes, respectively. Market equity is the product of Compustat total shares outstanding and the fiscal year end price (25*199). Size is the logarithm of the market equity at the fiscal year end in year *t*. *Dsize* is the absolute value difference in size. The book value of equity is the total assets minus total liabilities. Book-to-market equity is then the ratio of the book equity to the market value of equity at the fiscal year end in year *t*. *Dlogbtm* is the absolute value difference in log book-to-market ratios. *Sstate* is a dummy variable that is 1, if two firms are located in the same state, and 0 otherwise. Firm age is the difference between the current year and the first year that a firm appears in CRSP plus 1. *Dage* is the absolute value difference between logarithm of ages of two firms *i*, and *j*, $|\log(Age_i)-\log(Age_j)|$. *Slisting* is 1 if two stocks are listed on the same exchange, and 0 otherwise. *Sindex* is 1 if two stocks belong to the same S&P major, mid cap, or small cap indices, and 0 otherwise. *Dprice* is the absolute value difference in log prices per share. *Dduration* is the absolute value difference in cash flow durations. *Dleverage* is the absolute value difference in financial leverage ratios (long term debt/total assets). *Indlink* is 1 if two firms belong to industries that have customer-supplier links, and 0 otherwise. *Volumecorr* is the correlations between two firms' abnormal trading volumes. A stock is included in the sample only if it has 60 valid monthly returns, 20 quarterly earnings, and market equity of at least $500 million. Newey-West adjusted t-statistics are reported below the coefficients.

Dependent Variable: Idiosyncratic Return Correlation				
	1	2	3	4
Earncorr	0.020	0.015	0.015	0.014
	(3.92)	(4.32)	(4.01)	(4.68)
Earncorr_ch	0.021	0.019	0.019	0.019
	(2.98)	(3.23)	(3.32)	(3.26)
Dgrowth	-0.038	-0.026	-0.025	-0.024
	(-3.58)	(-3.39)	(-3.35)	(-3.01)
Ssic4		0.117	0.115	0.111
		(15.20)	(15.49)	(15.41)
Ssic3		0.038	0.039	0.037
		(3.06)	(3.04)	(2.92)
Ssic2		0.105	0.103	0.101
		(10.81)	(10.84)	(10.53)
Ssic1		0.026	0.026	0.024
		(7.55)	(8.09)	(6.33)
Dsize			-0.012	-0.012
			(-4.52)	(-4.38)

Dlogbtm			-0.005	0.003
			(-6.39)	(0.84)
Sstate				0.010
				(4.29)
Dage				-0.002
				(-1.16)
Slisting				0.020
				(1.73)
Sindex				0.001
				(0.79)
Dprice				-0.009
				(-2.66)
Dduration				-0.002
				(-3.29)
Dleverage				-0.015
				(-2.00)
Indlink				-0.006
				(-1.01)
Volumecorr				0.042
				(4.02)
Average obs	96,701	96,701	96,701	96,701
Average R^2	1.82%	9.09%	9.52%	10.46%

Table V: Implications of Comovement for Future Returns

This table reports Fama-Macbeth regression of monthly stock return difference (scaled by volatility), $|\frac{1}{\sigma_{i,t}} r_{i,t+1} - \frac{1}{\sigma_{j,t}} r_{j,t+1}|$, on pair-wise stock correlation and a group of control variables between July 1977 and June 2007. Each observation represents a stock pair. The dependent variable is the absolute value difference in monthly stock return for a stock pair between July of year *t+1* and June of year *t+2*. *Corr* is the stock return correlation between two stocks using monthly returns between year *t-4* and *t*. Market equity is the product of Compustat total shares outstanding and the fiscal year end price (25*199). Size is the logarithm of the market equity at the fiscal year end in year *t*. *Dsize* is the absolute value difference in size. The book value of equity is the total assets minus total liabilities. Book-to-market equity is then the ratio of the book equity to the market value of equity at the fiscal year end in year *t*. *Dlogbtm* is the absolute value difference in log book-to-market ratios. Earnings correlation (*Earncorr*) is calculated from Compustat Price, Dividends, and Earnings dataset as follows. In each quarter, we construct the return on equity *ROE* as the ratio of the 12 month earnings per share to the book value of equity per share (ERN over BKV). To mitigate outliers, we set *ROEs* that are greater than 10, or less than -10, to be 10 and -10, respectively. We then compute the correlation between 20 quarterly *ROE* of any two stocks A and B, between year *t-4* and year *t*. *Earncorr_ch* is the correlation between two firms' changes in *ROE*s. *Dgrowth* is the absolute value difference in 5-year log sales growth rates. For any two stocks A and B, if they have the same 4-digit SIC industry code, we construct a dummy variable *Ssic4* to be 1, and 0 otherwise. Similarly we set dummy variables *Ssic3*, *Ssic2*, *Ssic1* to be 1 for stocks that have the same 3-digit SIC codes, 2-digit SIC codes, 1-digit SIC codes, respectively. *Sstate* is a dummy variable that is 1, if two firms are located in the same state, and 0 otherwise. Firm age is the difference between the current year and the first year that a firm appears in CRSP plus 1. *Dage* is the absolute value difference between logarithm of ages of two firms *i*, and *j*, $|\log(Age_i)-\log(Age_j)|$. *Slisting* is 1 if two stocks are listed on the same exchange, and 0 otherwise. *Sindex* is 1 if two stocks belong to the same S&P major, mid cap, or small cap indices, and 0 otherwise. *Dprice* is the absolute value difference in log prices per share. *Dduration* is the absolute value difference in cash flow durations. *Dleverage* is the absolute value difference in financial leverage ratios (long term debt/total assets). *Indlink* is 1 if two firms belong to industries that have customer-supplier links, and 0 otherwise. *Volumecorr* is the correlations between two firms' abnormal trading volumes. A stock is included in the sample only if it has 60 valid monthly returns, 20 quarterly earnings, and market equity of at least $500 million. Newey-West adjusted t-statistics are reported below the coefficients.

Dependent Variable: $\|\frac{1}{\sigma_{i,t}} r_{i,t+1} - \frac{1}{\sigma_{j,t}} r_{i,t+1}\|$				
	1	2	3	4
Corr	-0.324	-0.311	-0.278	-0.264
	(-7.70)	(-7.65)	(-7.05)	(-7.13)
Dsize		0.010	0.011	0.010
		(1.46)	(1.57)	(1.41)
Dlogbtm		0.027	0.024	0.014
		(3.04)	(2.54)	(1.77)
Earncorr			-0.006	-0.004
			(-1.13)	(-1.03)
Earncorr_ch			0.000	0.000
			(0.03)	(-0.08)
Dgrowth			0.035	0.036
			(1.39)	(1.30)
Ssic4			0.007	-0.001

			(0.36)	(-0.06)
Ssic3			-0.045	-0.042
			(-4.10)	(-4.52)
Ssic2			-0.025	-0.028
			(-2.66)	(-3.03)
Ssic1			-0.024	-0.020
			(-3.95)	(-4.48)
Sstate				-0.004
				(-0.85)
Dage				0.004
				(1.02)
Slisting				0.016
				(1.92)
Sindex				-0.018
				(-2.55)
Dprice				0.012
				(1.95)
Dduration				0.004
				(1.32)
Dleverage				0.017
				(0.57)
Indlink				0.014
				(1.57)
Volumecorr				0.001
				(0.05)
Average Obs.	89,025	89,025	89,025	89,025
Average R2	0.72%	1.29%	2.05%	3.72%

Table VI: Portfolio Returns sorted on Previous Month's Return and Previous Month's Comover Portfolio Return

This table reports the average returns for 25 portfolios formed on lagged return and lagged comover portfolio return between 1931 and 2007. For each month in year *t+1*, portfolio is formed based on previous month's stock return and previous month's comover portfolio return, and held for one month. The comover portfolio is the equal weighted portfolio of 50 stocks that have the highest return correlations with a given stock between year *t-4* and year *t*. *T*-statistics are reported below the average returns.

Panel A: Value weighted

	Low Lag Comover Return	2	3	4	High Lag Comover Return	L/S
Low Lag Return	0.93%	1.54%	1.90%	2.35%	2.47%	1.71%
	3.61	6.02	7.12	8.12	8.27	7.64
2	0.84%	1.15%	1.39%	1.70%	2.10%	1.26%
	3.83	5.52	6.87	7.42	8.21	6.19
3	0.57%	1.00%	0.91%	1.37%	1.69%	1.12%
	2.61	5.00	4.50	7.14	8.00	6.45
4	0.35%	0.76%	0.98%	1.18%	1.34%	0.99%
	1.61	3.71	4.87	5.76	5.91	5.54
High Lag Return	0.42%	0.62%	0.70%	0.86%	0.95%	0.56%
	1.38	2.37	2.96	3.76	4.07	2.47

Panel B: Equal weighted

	Low Lag Comover Return	2	3	4	High Lag Comover Return	L/S
Low Lag Return	1.96%	2.43%	2.70%	3.09%	3.20%	1.73%
	6.23	7.96	8.36	8.95	9.16	8.00
2	0.79%	1.43%	1.77%	1.90%	2.51%	1.76%
	3.25	5.82	7.13	7.52	8.55	9.25
3	0.66%	1.08%	1.43%	1.60%	2.12%	1.46%
	2.86	4.74	6.32	7.32	8.54	9.74
4	0.33%	0.96%	1.13%	1.33%	1.80%	1.47%
	1.36	4.01	4.87	5.68	7.00	9.25
High Lag Return	0.10%	0.43%	0.54%	0.59%	0.86%	0.73%
	0.30	1.56	2.11	2.38	3.27	2.97

Table VII: Fama-MacBeth Regressions of Monthly Returns

This table reports the Fama-MacBeth regressions of monthly returns on lagged variables. *Cret* is the previous month's comover portfolio return. For each month in year *t+1*, the comover portfolio is the equal weighted portfolio of 50 stocks that have the highest return correlations with a given stock between year *t-4* and year *t*. *Lret* is the previous month's stock return. *RetDiff* is $beta^C$*(*Cret-Rf*) – (*Lret* – *Rf*), where $beta^C$ is the regression coefficient of a firm's monthly return on its comover portfolio return in the most recent 5 years. For returns between July of year *t+1* and June of year *t+2*, we match with *Size* and book-to-market equity at the fiscal year end in year *t*. The market value of equity is Compustat total shares outstanding multiplied by the fiscal year end price (25*199). *Size* is the logarithm of the market value of equity. The book value of equity is the total assets minus total liabilities (6-181). *Logbtm* is logarithm of the ratio of the book equity to the market value of equity. *Price* is the logarithm of price in the previous month. Variables are demeaned before forming interaction terms. All the regressions are for the sample period July 1951 to December 2007. *T*-statistics are reported below the coefficients.

	1	2	3	4	5	6	7	8	9
Cret	0.228	0.191	0.191	0.186	0.170				
	12.93	12.45	12.47	12.05	11.16				
Lret	-0.069	-0.073	-0.070	-0.073	-0.059				
	-17.46	-18.04	-17.83	-17.69	-15.87				
RetDiff						0.082	0.082	0.077	0.065
						18.14	18.72	18.42	17.56
Size		-0.001	-0.001	-0.001	-0.001		-0.001	-0.001	-0.001
		-2.81	-2.93	-2.63	-1.9		-2.66	-2.67	-2.11
Logbtm		0.002	0.002	0.002	0.003		0.002	0.002	0.003
		4.19	4.28	4.29	5.13		3.98	3.97	4.85
Size*Cret			-0.013						
			-1.99						
Size*Lret			0.002						
			1.61						
Size*RetDiff								-0.005	
								-3.19	
Cret*Lret				0.030					
				0.20					
Price					0.000				0.000
					-0.39				0.01
Price*Cret					0.033				
					9.56				
Price*Lret					-0.052				
					-3.67				
Price*RetDiff									-0.036
									-10.10
Avg. Obs.	1994	1994	1994	1994	1994	1994	1994	1994	1994
Avg. R^2	0.025	0.047	0.053	0.051	0.067	0.011	0.038	0.041	0.056

Table VIII: Portfolio Formed on Return Difference

This table reports the value weighted and equal weighted returns for portfolios that are formed on return difference (*RetDiff*). *Cret* is the previous month's comover portfolio return. For each month in year *t+1*, the comover portfolio is the equal weighted portfolio of the 50 stocks that have the highest return correlations with a given stock between year *t-4* and year *t*. *Lret* is the previous month's stock return. *RetDiff* is $beta^C$*(*Cret-Rf*) – (*Lret* – *Rf*), where $beta^C$ is the regression coefficient of a firm's monthly return on its comover portfolio return in the most recent 5 years. The three factors are excess market return, SMB, and HML. The five factors are the three factors, plus momentum factor, and short term reversal factor. Panel A reports value weighted portfolios formed using all stocks with 60 monthly returns in the last 5 years. Panel B reports equal weighted portfolios formed using all stocks with 60 monthly returns in the last 5 years. Panel C reports value weighted portfolios using stocks with prices at least $5 per share at end of previous month. The sample period is January 1931 to December 2007.

Panel A: Value weighted portfolios

	Raw Return	3-Factor Alpha	5-Factor Alpha
Decile 1	0.45%	-0.70%	-0.45%
2	0.65%	-0.43%	-0.17%
3	0.74%	-0.26%	-0.13%
4	0.93%	-0.03%	0.04%
5	0.97%	0.04%	0.04%
6	1.17%	0.24%	0.18%
7	1.16%	0.18%	0.10%
8	1.35%	0.33%	0.26%
9	1.53%	0.39%	0.39%
Decile 10	1.86%	0.52%	0.46%
Decile 10-Decile 1	1.40%	1.23%	0.91%
t-statistics	9.28	8.32	6.61

Panel B: Equal weighted portfolios

	Raw Return	3-Factor Alpha	5-Factor Alpha
Decile 1	0.00%	-1.45%	-1.15%
2	0.61%	-0.74%	-0.53%
3	0.94%	-0.36%	-0.22%
4	1.07%	-0.20%	-0.14%
5	1.22%	-0.03%	-0.01%
6	1.37%	0.11%	0.05%
7	1.56%	0.25%	0.16%
8	1.78%	0.41%	0.32%
9	2.22%	0.69%	0.68%
Decile 10	3.59%	1.72%	1.85%
Decile 10-Decile 1	3.59%	3.17%	3.00%
t-statistics	18.69	18.30	17.76

Panel C: Value weighted portfolios, price>=5

	Raw Return	3-Factor Alpha	5-Factor Alpha
Decile 1	0.45%	-0.63%	-0.45%
2	0.63%	-0.42%	-0.19%
3	0.74%	-0.24%	-0.16%

4	0.89%	-0.07%	0.06%
5	0.97%	0.04%	0.02%
6	1.13%	0.19%	0.15%
7	1.15%	0.19%	0.11%
8	1.32%	0.35%	0.21%
9	1.43%	0.39%	0.30%
Decile 10	1.63%	0.45%	0.30%
Decile 10-Decile 1	1.17%	1.08%	0.75%
t-statistics	8.23	7.60	5.74

Table IX: Time-Series Test

This table reports the factor loadings for portfolios that are formed on return difference (*RetDiff*). *Cret* is the previous month's comover portfolio return. For each month in year *t+1*, the comover portfolio is the equal weighted portfolio of the 50 stocks that have the highest return correlations with a given stock between year *t-4* and year *t*. *Lret* is the previous month's stock return. *RetDiff* is $beta^{C}*(Cret-Rf) - (Lret - Rf)$, where $beta^{C}$ is the regression coefficient of a firm's monthly return on its comover portfolio return in the most recent 5 years. The five factors are the excess market return, SMB, HML, momentum factor, and short term reversal factor. The sample period is January 1931 to December 2007.

Panel A: Value weighted portfolios						
	Alpha	Rm-Rf	SMB	HML	MOM	ST Rev
Decile 1	-0.45%	1.13	0.28	0.08	0.00	-0.31
2	-0.17%	1.10	-0.01	0.09	-0.04	-0.27
3	-0.13%	1.01	-0.09	0.10	-0.03	-0.12
4	0.04%	0.98	-0.12	0.09	-0.02	-0.08
5	0.04%	0.96	-0.15	0.05	-0.02	0.02
6	0.18%	0.94	-0.08	0.05	-0.01	0.08
7	0.10%	0.98	-0.10	0.07	-0.01	0.12
8	0.26%	1.01	0.01	0.01	-0.04	0.15
9	0.39%	1.10	0.09	0.04	-0.12	0.17
Decile 10	0.46%	1.25	0.32	0.07	-0.14	0.27
Decile 10 - Decile 1	0.91%	0.12	0.04	-0.02	-0.14	0.58
t-statistics	6.61	4.36	0.96	-0.46	-4.39	15.10

Panel B: Equal weighted portfolios						
	Alpha	Rm-Rf	SMB	HML	MOM	ST Rev
Decile 1	-1.15%	1.02	1.01	0.31	-0.14	-0.20
2	-0.53%	1.04	0.62	0.33	-0.10	-0.13
3	-0.22%	1.01	0.52	0.32	-0.09	-0.06
4	-0.14%	0.99	0.43	0.35	-0.06	0.01
5	-0.01%	0.96	0.41	0.33	-0.09	0.09
6	0.05%	0.93	0.52	0.33	-0.06	0.16
7	0.16%	0.99	0.54	0.35	-0.06	0.19
8	0.32%	1.01	0.64	0.36	-0.09	0.23
9	0.68%	1.06	0.90	0.41	-0.16	0.22
Decile 10	1.85%	1.13	1.40	0.57	-0.37	0.32
Decile 10 - Decile 1	3.00%	0.11	0.39	0.26	-0.23	0.52
t-statistics	17.76	3.34	7.86	5.49	-5.97	11.09

Table X: Difference between our Trading Strategy and Pairs Trading

This table reports the returns for portfolios that are formed on return difference (*RetDiff*). *Cret* is the previous month's comover portfolio return. For each month in year *t+1*, the comover portfolio is the equal weighted portfolio of the 50 stocks (for Panel A) or 1 stock (for Panel B) that have the highest return correlations with a given stock between year *t-4* and year *t*. *Lret* is the previous month's stock return. *RetDiff* is $beta^{C}$*(*Cret-Rf*) – (*Lret* – *Rf*), where $beta^{C}$ is the regression coefficient of a firm's monthly return on its comover portfolio return in the most recent 5 years. In Panel A, equal weighted portfolios are formed on stocks that have positive trading volumes every day in the previous 12 months and we skip one day between the one-month formation period and one-month holding period. In Panel B, value weighted portfolios are formed on all stocks that have 60 monthly returns in the most recent 5 years. The three factors are excess market return, SMB, and HML. The five factors are the three factors, momentum factor, and short term reversal factor. The sample period is July 1962 to December 2002 for Panel A, and January 1931 to December 2007 for Panel B.

Panel A: Portfolios are for stocks traded every day, skip one day, equal weighted

	Raw Return	3-Factor Alpha	5-Factor Alpha
Decile 1	0.22%	-0.93%	-0.60%
2	0.48%	-0.70%	-0.51%
3	0.64%	-0.53%	-0.39%
4	0.83%	-0.33%	-0.26%
5	1.03%	-0.16%	-0.06%
6	1.21%	0.00%	0.06%
7	1.33%	0.10%	0.18%
8	1.46%	0.21%	0.30%
9	1.69%	0.41%	0.53%
Decile 10	1.91%	0.54%	0.76%
Decile 10-Decile 1	1.70%	1.47%	1.36%
t-statistics	8.32	7.2	8.52

Panel B: Only one stock in the comover portfolio, value weighted

	Raw Return	3-Factor Alpha	5-Factor Alpha
Decile 1	0.59%	-0.58%	-0.24%
2	0.59%	-0.43%	-0.27%
3	0.80%	-0.18%	-0.04%
4	0.82%	-0.16%	-0.04%
5	0.89%	0.00%	0.00%
6	1.04%	0.10%	0.09%
7	1.19%	0.24%	0.13%
8	1.27%	0.28%	0.10%
9	1.40%	0.35%	0.18%
Decile 10	1.54%	0.34%	0.19%
Decile 10-Decile 1	0.95%	0.92%	0.43%
t-statistics	7.22	6.95	3.70

Figure 1: Whirlpool and Comover Portfolio Return

This figure plots the returns for Whirlpool and its equal weighted comover portfolio return. The comover portfolio is the equal weighted portfolio of the 50 stocks that have the highest return correlations with Whirlpool between 2001 and 2005. The sample period is from January 2001 to February 2006.