Comments Welcome

An Anatomy of Pairs Trading: the Role of Idiosyncratic News, Common Information and Liquidity*

Joseph Engelberg, Pengjie Gao and Ravi Jagannathan†

First Draft: August, 2007
This Draft: February 22, 2009

Abstract

"Pairs trading" involves taking a bet that the price paths of two stocks that have historically moved together will converge again after any divergence. Consistent with the view that profits to pairs trading comes through market making, i.e., short term liquidity provision and price discovery; we find that pairs trading profits are higher when the initial divergence is due to (a) news that temporarily reduces the liquidity of one of the stocks in the pair, or (b) news that affects both stocks in the pair, and there are a priori reasons to believe that one stock reacts faster to such news. Profits are lower when the initial divergence in prices is associated with value relevant news specific to a stock in the pair. Pairs involving smaller, less liquid and more volatile stocks tend to converge faster after initial divergence. When one of the stocks in the pair is more likely to have sluggish response to common information as evidenced by less common sell side coverage and institutional holdings, the risk of a margin call due to further divergence is lower. Closing out positions that do not converge within 10 days leads to higher risk adjusted returns ignoring transactions costs when compared to holding positions for 6 months.

JEL Classification: G11, G12, G14

Key Words: Pairs Trading; News; Liquidity Shocks; Convergence Trade

*We are grateful to Robert Battalio, Jennifer Conrad, Shane Corwin, Tom Cosimano, Zhi Da, Levent Guntay, Tim Loughran, Todd Pulvino, Adam Reed, Paul Schultz, Sophie Shive, and several empolyees at Lehman Brothers for helpful discussions and comments. We also thank seminar participants at the Baruch College, University of Illinois (Chicago), University of North Carolina at Chapel Hill, University of Notre Dame, the Indiana-Purdue-Notre Dame Joint Conference on Finance, and Tykhe Capital Management for valuable feedback. The Zell Center for Risk Management and the Financial Market and Institution Center provided partial funding for this research. All errors are our own.

†Correspondence: Joseph Engelberg, Finance Area, Kenan Flagler Business School, University of North Carolina at Chapel Hill; Pengjie Gao, Finance Department, Mendoza College of Business, University of Notre Dame; and Ravi Jagannathan, Finance Department, Kellogg School of Management, Northwestern University and NBER. E-mail: joseph_engelberg@kenan-flagler.unc.edu, pgao@nd.edu, and rjaganna@kellogg.northwestern.edu.

1 Introduction

With the easy access to real time financial market data and vast amount of inexpensive computing power, quantitative trading strategies, especially those that can be implemented using machines with little human interventions have become increasingly popular. Many of these strategies, falling in the "statistical arbitrage" category, continue to earn rather large risk adjusted returns on paper even after they have been well publicized through published articles, and pose serious challenge to the "efficient market hypothesis."

One such popular statistical arbitrage strategy is "pairs trading."[1] The idea behind pairs trading is to first identify a pair of stocks with similar historical price movement. Then, whenever there is sufficient divergence between the prices in the pair, a long-short position is simultaneously established to bet that the pair's divergence is temporary and that it will converge over time. Recently, Gatev, Goetzmann and Rouwenhorst (2006, hereafter GGR) showed that a pairs trading strategy generates annual returns of 11 percent and a monthly Sharpe ratio four to six times that of market returns between 1962 and 2002. The abnormal returns associated with pairs trading are comparable in magnitude to those associated with the relative-price momentum strategy in Jegadeesh and Titman (1993). Although Pairs Trading is a contrarian strategy, it involves investing in stocks that are not the same as those associated with contrarian relative-price momentum strategies documented in Jegadeesh (1990), Lehmann (1990), and Lo and MacKinlay (1990).

Despite the large abnormal (risk-adjusted) returns associated with pairs trading, we know little about what makes pairs trading profitable. In this paper we make an attempt to fill this void in the literature by examining the following questions: Why some pairs are more profitable than others? What causes the prices of pairs to diverge and how that affects subsequent convergence?

We find that type of the news released around the divergence date plays a critical role in explaining cross-sectional variations in the profits across different pairs trading positions. We identify idiosyncratic news events from articles in the Dow Jones News Service and find that when a pair diverges because of firm-specific news, the divergence is more likely to be permanent, and hence the profitability to a pairs trading strategy is lower. We find that the profitability of a pairs trading position is related to information events that affect both firms in the pair (i.e., "common shocks"). Using a measure of information diffusion at the industry level we find some of the profitability of a pairs trading position is related to the differential responses to these common information shocks and which in turn is related to the different liquidity levels of the constituent stocks. The profitability of a position is smaller when institutional investors hold both of the constituent firms in a pair, and sell-side analysts cover both of the constituent firms. These findings suggest that the profits to pairs trading comes from two sources: providing liquidity where it is needed and taking positions when stock prices react slowly to information common to both firms.

[1] Several strategies are similar to pairs trading. Instead of relying on the statistical relationship between historical prices as in pairs trading, these strategies consider relative pricing of shares due to differences in trading locations (Froot and Dabora, 1999; Scruggs, 2006) or differences in cashflow rights and voting rights (Smith and Amoako-Adu, 1995; Zingales, 1995; Schultz and Shive, 2008).

In addition to identifying a set of economic factors related to pairs trading profits, we also investigate how these factors affect the timing of the trades, i.e., when the positions are opened and subsequently closed, and the risks associated with maintaining a position. We find that the type of news and liquidity of the underlying stocks are related to the nature of the risk and rewards associated with pairs trading positions.

Recent literature, typified by Kovajecz and Odders-White (2004) interprets the apparently high return to such strategies as being reward to investment in market making activities, i.e., trading activities that provide immediacy and price discovery in the underlying securities. Viewed that way, such strategies are also exposed to liquidity risk that affects market makers as a group more severely, in addition to standard economy wide pervasive risks that affects everyone in the economy. As pointed out by Khandani and Lo (2007), while these strategies are often uncorrelated with little market (beta) exposure, they can become highly interconnected with dramatic increase in correlations during a very short period of time thereby contributing to potential financial contagion. Consistent with that view, we find that return to pairs trading have significant exposure to the liquidity risk factor (Pastor and Stambaugh, 2003), and the U.S. Treasury / Euro Dollar (TED) spread factor which capture economy wide liquidity shocks (Asness, Moskowitz, and Pedersen, 2008). We also find that liquidity provision and price discovery components contribute about equally to the abnormal returns to pairs trading.

The rest of the paper is organized as follows. Section 2 introduces the pairs trading algorithm of GGR, describes the stock price data set we use in our study, and highlights some patterns in pairs trading profits that motivates the questions we examine in latter sections. Section 3 provides a brief review the related literature. Section 4 describes the additional data we use for examining why some pairs trading positions are more risky than others, and why some are more risky than others. Section 5 examines how profits from pairs trading are related to pair characteristics using calendar time methods. Section 6 explores variables that help explain the cross sectional variation in risk and returns on different pairs trading positions. We conclude in section 7.

2 The Pairs Trading Strategy

2.1 The Gatev, Goetzmann, and and Rouwenhurst (2006) Algorithm

At the end of each calendar month, m, we follow Gatev, Goetzmann and Rouwenhorst (2006) to identify the universe of eligible pairs using return data for all stocks during the immediately preceding 12 month "estimation period". For that purpose, we first normalize the price of every stock to equal unity at the beginning of the estimation period. Let T_m denote the number of trading days in the 12 month estimation period associated with month m. On each trading day t, $t = 1, 2, ...T_m$ from the start of the 12 month estimation priod we compute each individual stock's normalized price, P_t^i, as:

$$P_t^i = \prod_{\tau=1}^{t} 1 \times \left(1 + r_\tau^i\right) \tag{1}$$

where P_t^i is stock i's normalized price by the end of day t, τ is the index for all the trading days between the first trading day of the year till day t, and r_τ^i is the stock's total rate of return (dividends included) on day τ. To ensure the set of stocks involved in pairs trading are relatively liquid, we only include stocks with positive trading volume during the estimation period. For each calendar month, m, we then we compute the following average squared normalized price difference measure, $PD_{i,j,m}$ between stock i and stock j, for every pair of stock:

$$PD_{i,j,m} = \frac{\sum_{t=1}^{T_m} \left(P_t^i - P_t^j\right)^2}{T_m} \tag{2}$$

where T_m is the total number of trading days in the estimation period for calendar month m, P_t^i and P_t^j are the normalized prices for stock i and stock j respectively on trading day t in the estimation period. Let the standard deviation of the squared normalized price differences be given by:

$$StdPD_{i,j,m} = \sqrt{\frac{1}{T_m - 1} \sum_{t=1}^{T_m} \left[\left(P_t^i - P_t^j\right)^2 - PD_{i,j,m}\right]^2} \tag{3}$$

For each calendar month, m, we then identify 200 pairs that have the smallest average normalized price difference during its estimation period. If there are N_m stocks available for consideration at the end of month m, we need to compute $N_m \times (N_m - 1)/2$ normalized price differences, which potentially could be a very large number. We therefore choose to limit our attention to pairs from the same industry. In particular, we use the Fama-French twelve-industry industry classification scheme (Fama and French, 1997), when we compute pairwise normalized price differences.

For each calendar month m, those 200 pairs that have the smallest normalized price difference during month $m's$ estimation period will be "eligible" for investment during the 12 month "eligibility period" immediately following calendar month m. If the prices of the stocks in an eligible pair diverge by more than $2 \times StdPD_{i,j,m}$then we buy the "cheap" stock in the pair and sell the "expensive" one. As in GGR, we wait one day after divergence before investing in order to mitigate the effects of bid-ask bounce, trading halt, and other market microstructure induced irregularities. If the pair later converges, i.e., when their normalized prices cross for the first time after divergence, we unwind our position and wait for the pair to diverge again. If the pair diverges but does not converge within 6 months, we close the position and call this "no convergence." We also consider a "cream-skimming" strategy which closes the position on pairs that have not converged within 10 days.

We calculate buy-and-hold portfolio returns to the pairs trading strategy as in GGR to avoid the transaction cost associated with daily rebalancing. Let $p(l^i, s^i)$ – which we will write as p^i for brevity – indicate the pair of stock l^i and stock s^i for the pair. We let D^i indicate the most recent day of divergence for pair p^i. When we invest in the pair one day after divergence, we let the first coordinate (l) indicates the stock in which we go long and the second coordinate (s) indicates the stock in which we go short. We indicate the return to stock l^i on day t as $R_t(l^i)$ and the return to stock s^i on day t as $R_t(s^i)$ so that the return for p^i on day t is defined as,

$$R_t(p^i) = R_t(l^i) - R_t(s^i) \tag{4}$$

then the return to a portfolio of N pairs on day t is

$$R_t^{Portfolio} = \sum_{i=1}^{N} W_t^i R_t(p^i) \tag{5}$$

where the weight W_t^i is defined as

$$W_t^i = \frac{\varpi_t^i}{\sum_{j=1}^{N} \varpi_t^j} \ ,$$

and

$$\varpi_t^j = (1 + R_{t-1}(p^j)) \times (1 + R_{t-2}(p^j)) \times ... \times (1 + R_{D^i+1}(p^j)) \ .$$

In words, we use the N pairs that are held in the portfolio on day t , and calculate the daily return to the portfolio as the weighted average of the returns to the N pairs on day t where the weight (ϖ_t^i) given to the return of pair i on day t is determined by its cumulative return in the portfolio ending on day $t-1$ relative to the sum of the cumulative returns of all pairs in the portfolio.

2.2 Stock Price Data

Stock prices, returns, trading volume and shares outstanding are obtained from the Center for Research in Security Prices (CRSP) database. We only retain common shares (share code = 10 or 11) traded on NYSE, AMEX or NASDAQ (exchange code = 1, 2 or 3). We use daily stock price and return data for the period January 1992 to June 2006. We begin forming pairs in January 1993 based on the estimation period, January 1992 to December 1992. The last estimation period in our sample is January 2004 to December 2004. Eligible pairs corresponding to the last estimation period remain eligible till December 2005. Pairs that opened in December 2005 and did not coverge within the stipulated maximum of 6 months would have been closed in June 2006.

2.3 Patterns in Pairs Trading Profits

Figure 1 graphs the mean pair-return in *event-time* where event day T is $(T + 1)$ days after the pair diverges (at day 0). Consistent with GGR, skipping a day after the divergence of the pair prior to taking position mitigates microstructure effects, such as the first-order negative serial correlation induced by the bid-ask bounce. The figure clearly illustrates the profitability from pairs trading declines substantially in event time. For example, event day 1 and 2 generate a mean return of 23 and 13 basis points respectively but after event day 4 the mean pair-returns from pairs trading never reaches 10 basis points and after event day 20 the average daily return hovers and falls below 5 basis points (see the solid line). A five-day moving average plot (the dashed line) - which smoothes out the daily return variations - paints essentially the same picture.

Panels A, B, and C of Figure 2 present the empirical distribution of the probability of pair convergence within the next 5, 10 and 20 days in event time, conditional on the pair having not converged k days after their first divergence, for $k = 1, 2, 3,...,49,50$. For example, the probability of a pair converging within the next 20 days after event day 1 is 28% but the probability of a pair converging within the next 20 days after event day 30 (i.e. given a pair has not converged during the first 26 event days) is 20%. The figures demonstrate that after event day 7 the probability of convergence declines monotonically across all three plots. In other words, if a pair diverges and has not converged within the first 7 days it becomes increasing unlikely to converge. Finally, figure 3 plots the empirical distribution (along with a kernel density estimate) of the time to convergence conditional on convergence (i.e., given a pair converges, figure 3 shows the empirical frequency of time to that convergence). Given the results from figures 1 and 2 it is not surprising that the mode of this distribution is 8 days. Taken together, the evidence suggests that the profitability generated from pair trading position is short-lived and declines over time. That is consistent with the view that pairs trading is one of the quant tools available to those who make a market in the underlying stocks for locating which stocks will reward liquidity provision and price discovery at given points in time.

The event-time evidence presented in this section motivates much of our empirical work. First, our finding that the profitability from pairs trading is much larger on days close to divergence suggest that the divergence date is not some random day when a pair's spread reaches an arbitrary threshold. These divergence dates are critical. To better understand pairs trading, we need to better understand what happened on the divergence date and what pair characteristics contributed to the divergence.

Second, while the profits to a pairs trading strategy are large near the divergence date and then decline monotonically, the profitability remains economically and statistically significant for months after the first 10 days. Concerning statistical significance, after the first 10 days we find that in 83 of the following 100 days the average return is greater than zero. A binomial test easily rejects the null hypothesis that average returns from pairs trading during this later period is random around zero (p-value less than 0.001%). Concerning economic significance, we will show in Section 3.2 that a trading strategy which commits to holding a pair for as long as 6 months after divergence earns a higher return *per pair* than a strategy which commits to holding a pair for only 10 days (208 basis points versus 83 basis points). This observation suggests the profits from pairs trading could come from different sources. That is, while some factors may contribute to profits from pairs trading at the shorter horizon, some others may contribute to profits from pairs trading at the longer horizon. For example, liquidity provision may involve a shorter term than price discovery.

Third, if convergence of some pairs does not happen until several months later, we need to understand the risks an arbitrageur faces when he holds his long-short pair position over a non-trivial horizon. In particular, what are the factors related to the speed of convergence and what are the factors related to the divergence of the arbitrage spread before convergence? In what follows we make an attempt to answer these questions, though not necessarily in that order.

3 Related Literature

In this section we provide a brief discussion of several mechanisms we deem relevant for explaining cross-sectional variations in the profits from a pairs trading strategy, and the related literature. As we have discussed in previous sections, some factors may be more relevant to explain the short-term profits from pairs trading, such as a liquidity shock, while others may be useful to explain longer-term profits, such as the level of liquidity and the speed of information diffusion.

3.1 Liquidity and Asset Prices

The large difference of returns from short and long holding horizons and the exponential decline of profits after initial divergence suggest that liquidity may play a role in explaining the source of profits from pairs trading. Conrad, Hameed, and Niden (1994) find that the short-term reversal strategy's profits increase with trading volume. On the other hand, Cooper (1999) finds reversal strategy's profits decrease with trading volume. Avramov, Chordia and Goyal (2006) show that the largest return reversals from the contrarian trading strategy occur in high turnover and illiquid stocks. Gervais, Kaniel, and Mingelgrin (2001), document that extreme short-run trading volume (measured as turnover) changes precede large return changes in the same direction without any return reversal effects. Therefore, both the theoretical literature and prior empirical literature provide several possibilities. On the one hand, to the extent trading volume captures the non-information driven liquidity demand, and the change of volume captures the sudden change of liquidity demand, trading volume induced reversal effects may contribute positively to the profits from the pairs trading. On the other hand, if the sudden change of trading volume also captures informational effects due to increased visibility of the stocks (Gervais, Kaniel, and Mingelgrin, 2001), then the change of trading volume may contribute negatively to profits from pairs trading. The two effects will of course offset each other.

It is known that the *level of liquidity* may affect asset prices (Amihud and Mendelson, 1986). Moreover, the theoretical model of Campbell, Grossman and Wang (1993), suggests that non-information driven liquidity demand - the sudden *change of liquidity level*, i.e., *liquidity shocks* - causes temporary price pressure, conditional on the level of liquidity. Prices reverse back when such liquidity demand is accommodated. Consistent with such a theoretical argument, Llorente, Michaely, Saar and Wang (2002) find non-information driven hedging trades are related to the short-run return reversal effect. In the context of pairs trading, less liquid stocks are more likely to diverge for non-information reasons. Meanwhile, a lower level of liquidity may keep arbitrageurs at bay, which could contribute to a prolonged period of price divergence. Which of these two forces are more likely to prevail is ultimately an empirical question.

3.2 Information and Asset Prices

News is ubiquitous and plays a crucial role in financial markets, but it is far from clear how and when news gets impounded into asset prices. There have been many empirical studies that have found that future returns can be predicted from firm-level news such as earnings announcements (Ball and Brown, 1968, Bernard and Thomas, 1989), equity issuance (Loughran and Ritter, 1995; Loughran and Ritter, 1997), open market share repurchase (Ikenberry, Lakonishok, and Theo, 1995), dividend initiations and omissions (Michaely, Thaler, and Womack, 1995), among others. Even the well-known momentum anomaly is related to firm-level news. Chan (2003) finds evidence that the momentum effects only exist among firms that have had news in the previous month. Several papers have proposed non-risk based models to better understand the information processing mechanism of investors that would generate these return patterns.

More recent work has begun to examine news and future returns using a more complete collection of news events like those reported in the Dow Jones News Service or the Wall Street Journal without specifically attributing the nature of the news. Mitchell and Mulherin (1994) study the relationship between the number of news announcements from Dow Jones & Company and aggregate market trading volumes and returns, and find strong relationship between the amount of news and market activity. Tetlock et al. (2008) finds that the market underreacts to the linguistic content of news articles in the Dow Jones News Service and the Wall Street Journal, while Tetlock (2008) finds that the market overreacts to repeated news stories which suggests a differential response of the market to news and media coverage. Vega (2006) attempts to disentangle news and coverage by using the contemporaneous firm return and finds a difference in the way news and coverage relates to the Post Earnings Announcement Drift. Using a large sample of Wall Street Journal articles, Frank and Antweiler (2006) look at a large cross-section of firm news events and find that the market underreacts to some events and overreacts to others but they do not attempt to distinguish news from coverage. In this paper, we make a distinction between "news" and "coverage", and examine how market may respond differently to "news" and "coverage".

Hong and Stein (1999) build a heterogeneous belief model in which the economy is populated by two groups of bounded rational investors. The key assumption is that information is impounded into asset prices slowly as a group of "newswatchers" slowly acquire information. Consistent with the prediction of their model, Hong, Lim and Stein (2000) show momentum effects are weaker among firms with low analyst coverage. Cohen and Frazzini (2008) find evidence that information in the equity price of a customer firm incorporates slowly into the price of a supplier firm. Menzly and Ozbas (2006) find similar evidence across industries linked through a supply chain.

Our paper also explores how information diffusion affects asset prices. In the context of a relative valuation strategy like pairs trading, *two* kinds of information are important: idiosyncratic (firm-level) news and common (industry-level) news. If investors overreact to the idiosyncratic news of one stock in the pair which pushes its price away from its fundamental value as proxied by the price of the second stock in the pair, then there would be profits in the form of pairs trading as price converges to fundamental value. However, if information diffuses slowly into prices, then the presence of idiosyncratic news should create permanent differences in prices and have a negative affect on pairs trading profits. Therefore, idiosyncratic news may be related to returns from pairs trading, but different characterization of the reactions to news has different predictions.

With respect to common information, simple underreaction or overreaction are not enough to explain profits from the pairs trading. Two stocks may underreact or overreact, but if the extent and timing of underreaction or overreaction is the same, then convergence trading will not be profitable. It is the *relative* underreaction or overreaction that matters for pairs trading. If market frictions allow some information to be impounded into one stock in the pair more quickly, this will create a lead-lag relationship between the two stocks in a pair (see Conrad and Kaul, 1989; Lo and MacKinlay, 1990; Hong, Torous, and Valkanov, 2007 for investigations of this lead-lag relationship among individual stock returns and portfolio returns).

What market frictions will create such a differential response to common information? We consider three aspects of the constituents of the pairs: the underlying institutional shareholder ownership structure, the sell-side analyst's coverage and the liquidity of the shares. If slow information diffusion is at least partially due to costly information acquisition or illiquidity, then we anticipate such slow information diffusion to be more pronounced among stocks less commonly held by the institutions, less commonly covered by analysts, or less liquid.

3.3 Limits to Arbitrage

The "limits to arbitrage" literature dates back to De Long, Shleifer, Summers and Waldman (1990) and Shleifer and Vishny (1997) and suggests that various market frictions may impede arbitragers from eliminating mispricing. These market frictions include transaction costs, short-sale constraints and idiosyncratic risk, among others. First, after taking into account transaction costs, net returns from apparently profitable asset pricing anomalies attenuate or completely disappear.[2] GGR provide some estimates of after-transaction cost net returns. Their results show pairs trading profits decrease but not enough to explain pairs trading profits. Second, arbitrage trades usually involve both long- and short- positions to hedge away systematic risks, but short-sale constraints may impede the implementation of such strategies.[3] Idiosyncratic risk also limits the ability to execute an arbitrage and has been called "the single largest cost faced by arbitrageurs" (Pontiff, 2006).[4]

[2] See, for example, Lesmond, Schill and Zhou (2004), Korajczyk and Sadka (2004) on momentum profits; Batalio and Mendenhall (2006) on the post earnings announcement drfit (PEAD); Hanna and Ready (2005) on Haugen and Baker (1996) accounting and return based stock screening model; Mitchell and Pulvino (2001) on merger arbitrage; Scherbina and Sadka (2007) on the analyst disagreement anomaly.

[3] See Mitchell, Pulvino and Stafford (2002), and Lamont and Thaler (2003) for the discussion of short-sale constraints - in particular, the extremely high short-rebate rates - on negative stub value trades.

[4] Idiosyncratic risk is shown to be related to the close-end fund discount (Pontiff, 1995), merger arbitrage (Baker and Savasogul, 2002), index addition and deletion (Wurgler and Zhuravskaya, 2002), the book-to-market effect (Ali, Hwang, and Trombley, 2003), post earnings announcement drift (Mendenhall, 2004) and distressed security investment (Da and Gao, 2008), among others.

3.4 Risk in Statistical Arbitrage Strategies

Several aspects of the risks associated with statistical arbitrage strategies have received attention in the theoretical literature. One thread explores divergence risk and horizon risk, as well as their implications for asset prices. Divergence risk refers to the arbitrageurs face when their arbitrage positions may be wiped out before eventual convergence due to exacerbated mispricing. Horizon risk refers to the risk that convergence may not be realized during a fixed time horizon. Xiong (2001) considers wealth-constrained arbitrageurs. He shows that arbitrageurs may in general stabilize prices, nevertheless he finds that there are situations when arbitrgeurs can further exacerbate mispricing (the "amplification effect"). Liu and Longstaff (2004) directly model the divergence risk in arbitrage trading. One important implication from their model is that such divergence risk may preclude rational arbitrageurs from taking large positions to completely eliminate the temporary mispricing. Jurek and Yang (2005) consider divergence risk with uncertainties about both the magnitude of the mispricing and the convergence horizon. They derive an optimal investment policy and show the arbitrageur's position in convergence trades is subject to a threshold level - beyond which the arbitrage position decreases. In this paper we examine the divergence risk and horizon risk associated with pairs trading, and how those risks are systematically related to news, liquidity and the information environment, using statistical reduced form models.

4 Data Description and Summary Statistics

4.1 Data Sources

As mentioned in Section 2, Stock prices, returns, trading volume and shares outstanding are obtained from the Center for Research in Security Prices (CRSP) database. Accounting information is extracted from the Standard & Poor's Compustat annual files. To ensure accurate matching between CRSP and Compustat databases, we use CRSP-LINK database produced by the Center for Research in Security Prices (CRSP). To compute the proportional quoted spreads, we use the TAQ database disseminated by NYSE, and filter out all irregular trades following the procedure outlined in Bessembinder (2003). Quarterly institutional holdings are extracted from the CDA/Spectrum 13f database produced by Thomson/Reuters. Sell-side analyst coverage information is obtained from the "detailed files" of the Institutional Broker's Estimate System (I/B/E/S) database maintained by Thomson/Reuters.

Our database of news events are all Dow Jones News Service (DJNS) articles downloaded from Factiva between 1993 and 2005. Factiva is a database that provides access to archived articles from thousands of newspapers, magazines, and other sources, including more than 400 continuously updated newswires such as the Dow Jones newswires. The Dow Jones newswire is the newswire which covers North American markets (including NYSE, AMEX and NASDAQ) and companies. According to Chan (2003), "by far the services with the most complete coverage across time and stocks are the Dow Jones newswires. This service does not suffer from gaps in coverage, and it is the best approximation of public news for traders." We match the unique company codes assigned by Factiva to the CRSP PERMNO as in Engelberg (2008). The matching is done using a combination of ticker extraction from the DJNS articles as well as textual matching of the company names in Factiva and CRSP.

4.2 Variable Definitions

We outline the main variables used in this paper in this section. Several more complex variables are defined shortly in the section where we discuss the motivation behind their construction and the associated empirical results.

Avg_PESPR - the pair's average proportional effective spreads, measured in the previous ten days prior to the event day.

Avg_PESPR_Change - the change of the average of the pair's proportional effective spreads, measured in the previous five days leading to the event day minus the pair's average proportional effective spreads, measured in the previous tenth to the sixth days prior to the event day.

Avg_dTurn - the pair's average daily turnover ratio, measured in the previous ten days prior to the event day.

Avg_dTurn_Change - the change of the average of the pair's daily turnover ratio, measured in the previous five days leading to the event day; minus the pair's average daily turnover ratio, measured in the previous tenth to the sixth days prior to the event day.

Avg_Ret_pst1mth - the pair's average cumulative returns over the one month prior to the event month (event month is the month when the event date occurs).

Avg_Ret_pst12mth - the pair's average cumulative return over the eleven months prior to the second month to the event month.

Avg_Ret_pst36mth - the pair's average cumulative return over the 24 months prior to the 12 month to the event month.

Avg_BM - the pair's average book to market equity ratios measured using the most recently available book equity value, and the market equity values during the month ending at the beginning of the previous month.

Log_Avg_MktCap - the natural logarithm of market capitalization of firms in billion dollars using last available market capitalization t during the pair estimation period.

Avg_mRetVola - the average of the pair's monthly return residual volatilities estimated using daily returns during the pair estimation period.

Common_Holding - for the continuous version of this variable, it is computed as the number of institutions holding both stocks in the pair during the quarter prior to the event quarter (the quarter the event date occurs), divided by the maximum number of institutions holding stock one or stock two of the pair during the same quarter. For the binary version of this variable, if the number of institutions holding two stocks of the pair is less than fifty, the Common_Holding indicator variable takes the value of one; and zero otherwise.

Common_Coverage - for the continuous version of this variable, it is computed as the number of brokerage houses (as identified by the brokerage code in I/B/E/S), divided by the maximum number of brokerage houses covering stock one or stock two of the pair during the same quarter. For the binary version of this variable, if the number of brokerage houses covering two stocks of the pair is less than or equal to two, the Common_Coverage indicator variable takes the value of one; and zero otherwise.

Abnormal Return - is a binary variable which takes the value of one if one stock in the pair has an absolute return greater than two standard deviations of the daily return calculated over the previous 21 trading days (a month).

News - a binary variable which takes the value of one if at least one stock in the pair has both a news article in the Dow Jones News Service on the day of divergence and an abnormal return.

No News (Coverage) - a binary variable which takes the value of one if at least one stock in the pair has a news article in the Dow Jones News Service on the day of divergence but neither stock has an abnormal return.

Size_Rank - a binary variable which takes the value of one if the average size percentile of the pair is below 50-th of NYSE decile breakpoints, and zero otherwise.

4.3 Summary Statistics

Table 1 provides sample mean, median, first quartile, third quartile and standard deviations of the pair's characteristics. There are a few points of interest from the table. First, the stocks in our sample are, on average, larger firms. The average NYSE size rank of our paired stocks is 65th percentile so we should be less concerned about the implementability of a pairs trading strategy from these stocks. When we look at the kinds of industries that make up our pairs in Panel B, we find that almost half of our pairs (44.38%) come from the financial industry and there is also significant representation from utilities (22.52%) and manufacturing (13.96%). This may be due to the fact that the prices of stocks within these industries might comove with macro information about interest rates, energy prices and commodity prices. When we sort on pairs based on whether they are listed on the same exchange or different exchanges we find that pairs on "mixed" exchanges lead to more pairs trading profits and that this result is statistically significant.

Figures 4 and 5 give the number of pairs and the number of unique pairs that are invested in calendar time during the sample period. As can be seen, except for the first six months and the last six months of the sample period, there are at least 200 invested unique pairs under the 10 day "cream skimming" strategy and 600 invested unique pairs under the 6 month strategy on each trading day. To see why the number of unique pairs is different from the total number of pairs invested in, consider the pair, say IBM and DELL. Suppose IBM-DELL was in the eligible list of pairs constructed using the 12 month estimation period for month t. Suppose further that IBM-DELL is in the eligibliity list constructed for month t+1 as well. The eligible list of pairs constructed in month t will remain eligible till month $t + 12$, and that constructed in month $t + 1$ will remain valid till month $t+13$. Suppose IBM-DELL opened (i.e.,the squared deviation of IBM's normalized price from DELL's price exceeded twice the estimation period standard deviations of the squared normalized price differences) for the first time in month t+2. In that case IBM-DELL will be invested in twice – i.e., will get twice the weight in the portfolio of invested pairs. Therefore the total number of invested pairs will in general be larger than the number of unique invested pairs, since we count IBM-DELL in month t eligibility list as being different from IBM-DELL in month t+1 eligibility list.

Figure 6 gives the distribution of the number of invested pairs in event time, with day 0 being the first day following the computation of the list of eligible pairs using 12 preceding monthly returns on stocks. The number of open pairs for the 10 day cream skimming strategy peaks at 25 by the end of the first month, declines to 10 by the end of 2 months and thereafter declines slowly to about 5 by the end of the year; suggesting that the probability of a pair opening before subsequent closing is much higher initially, and stabilizes after the first two months at a lower level. The number of open pairs in the 6 month strategy occurs much later at 6 months, indicating that substantial fraction of pairs do not converge after they first open within 10 days, and a signficant number continue to remain open even after 6 months.

Table 2 investigates the distribution of a selected set of corporate events - quarterly earnings announcements, seasoned equity offerings, mergers and acquisitions, and debt issuance - within a two-day window leading to the date of divergence, $[t-1, t]$, where t is the day of divergence. Panel A examines all pairs that diverge, and Panel B examines all pairs that diverge and there is at least one piece of news coverage on at least one stock of the pair on the divergence date. Quarterly earning announcements is the most frequently identified event. They occur among six percent of the opened pairs, and eight percent of the opened pairs with news coverage on the divergence date. This table shows no single type of corporate event news dominates around the date of divergence. Thus it is unlikely that all divergence can be reliably attributed to one single event. This table also shows that using news coverage constructed from Dow Jones News Service is necessary because it significantly enlarges the collection of news associated with the stock.

One concern the reader may have is about the disproportionately large number of index addition and deletion events, which could induce potentially permanent divergence of pairs. In an untabulated analysis, we find that among the $27,703$ pairs retained, only 69 pairs experienced index addition or deletion during the event window of $[t-30, t]$, 23 pairs experienced index addition or deletion during the event window of $[t-1, t]$, and 9 pairs experienced index addition or deletion on date t, where t is the date of divergence. In summary, index addition and deletion events are unlikely to be the major events behind the divergence of pairs.

5 Calendar Time Pairs Trading Portfolio Returns

This section motivates and performs a series of asset pricing tests on calendar-time pairs trading portfolios. Calendar time portfolios with returns constructed as in Section 2 are useful because they approximate the returns to an arbitrageur who executes a pairs trading strategy. Our calendar time portfolios are overlapping at the monthly level. For example, we begin forming the portfolios in January of 1993 based on the estimation period of January 1992 - December 1992. The top 200 pairs in this estimation period are eligible to open from January 1993 to December 1993 and, given that the pair opens, may be held for as long as 6 months under the standard strategy (i.e. a pair may be held into June of 1994 if it diverges in December of 1993 and never converges). Next, the top 200 pairs from the estimation period February 1992 - January 1993 are eligible to open for one year beginning February of 1993. And so on. The last month in which a new 200 pairs becomes eligible is for the period January 2005 to December 2005 and pairs which open in December 2005 may be held as long as June 2006 under the standard strategy.

Construction of the overlapping portfolios in this way will make it so that months in the beginning and ending of the portfolio holding period may have few stocks (depending on the opening and closing events of the pairs) - especially when we perform double-sorts. In some cases of double sorting, we may have no stocks in the portfolio in January of 1993 or June of 2006. For this reason, our number of observations (months) for the standard strategy may be 161 instead of 162.

For the majority of this section, the portfolios are sorted in different ways in order to demonstrate the effect of timing, firm-level news, industry-level news and liquidity on the profitability (alpha) from these calendar-time portfolios. The evidence presented here suggests strong heterogeneity in the performance of portfolios sorted on these variables.

5.1 Profitability and Timing

To formally investigate the timing and profits from pairs trading, we examine pairs trading strategies which hold the long-short position for various lengths of time. As we have discussed earlier, a strategy which holds the position for a short window after divergence seems to earn higher returns. This conjecture is confirmed in Table 3. The profits to a "cream-skimming" strategy that holds the position no longer than 10 days after divergence earns superior returns to a standard strategy which holds the pair no longer than 6 months after divergence. Just like the standard strategy, the cream-skimming strategy requires pairs to fully converge before they are eligible to be invested in again after divergence. This means that the 6-month strategy and the 10-day strategy will require the exact same number of round-trip transactions.[5] Monthly returns are regressed against the three Fama-French factors to adjust for exposure to economy wide pervasive risk; and two style factors, a momentum factor and a short-term reversal factor, to adjust for the possibility that pairs trading profits may be driven by the same forces that drive the well documented price momentum phenomenon in stocks.

The standard strategy with the maximum holding horizon of six months generates factor-model

[5] To see this in an example, suppose a pair diverges on day 1, converges on day 15, diverges again on day 40 and never converges. Under the standard strategy, we would open our position in the pair on day 3 (recall that we wait one day after divergence) and close our position by convergence on day 15. Then, we would open a position again in the pair on day 42 and close the position 126 days later on day 167. This entails a total of 2 roundtrip transactions in the pair. Under the cream-skimming strategy, we would open our position in the pair on day 3 and close our position on day 12 (since we hold the positions for a maximum of 10 days). Then, we would open a position again in the pair on day 42 and close the position 10 days later on day 51. This also entails a total of 2 roundtrip transactions in the pair.

adjusted return of 70 basis points (bpts) per month in our sample between 1993 and 2006. The return is comparable to the monthly factor-model adjusted return of 51 to 65 basis points per month reported by GGR (see table 4 of GGR). The main difference between our results and those reported by GGR lies in the choice of the stock universe to construct pairs. GGR mainly focus on pairs constructed from all available stocks in CRSP (subject to some exclusion criteria), while we construct pairs from industry sectors (but subject to the same exclusion criteria).[6] Also, consistent with factor regressions in GGR, we find that the pairs trading returns load negatively on the momentum factor and but positively on the short-run reversal factor. However, factor models do not explain much the time-series variation in pairs trading return. Usually, the *R-squared* from the regression is not high (about 30%).

What is most interesting to us is that the cream-skimming strategy earns a monthly alpha of 175 basis points (bpts) compared to a monthly alpha of 70 basis points (bpts) for the standard pairs trading strategy, while the factor loadings and the statistical significance of these two strategies barely change. However, the standard strategy earns more *per pair* than the cream-skimming strategy. We hold a pair after it opens for an average of 66 trading days for a total return of 208 bps per pair under the standard strategy and hold a pair for an average of 10 trading days for a total return of 83 basis points (bpts) under the cream-skimming strategy.

5.2 Profitability and Liquidity

To capture the level and change of liquidity, we introduce the pairwise average proportional effective spreads ($PESPR$), and the change of pairwise average proportional effective spreads ($\Delta PESPR$). In Table 4 Panel A, we consider the returns from pairs trading with a ten day maximum holding horizon. In this panel, we first split the sample into two portfolios based on the average market capitalization of stocks in the pair; then we further sort the pairs based on the average proportional effective spreads ($PESPR$), where the monthly factor-model adjusted returns are reported in the left columns; or sort the pairs into tercile portfolios based on the change of the average proportional quoted spreads ($\Delta PESPR$), where the monthly factor-model adjusted returns are reported in the right columns. Panel B is similar to Panel A except in B the maximum holding period is six months.

The numbers in Table 4 demonstrates that the *level of liquidity* has a persistent effect on the profits from pairs trading but that the *change in liquidity* ("liquidity shock") has a temporary effect. When we define the level of liquidity as the average proportional effective spread during the estimation period, we find a strong and positive relationship between it and the profits from a standard pairs trading strategy, but a statistically weaker and positive relationship between it and the cream-skimming strategy. Pairs from the most illiquid tercile outperform those from the most liquid tercile by 70 to 80 basis points per month when the holding horizon is ten days and 20 to 50 basis points per month when the holding horizon is six months. The effect is stronger among pairs with smaller average market capitalization.

[6] Indeed, if we compare the pairs trading return in this paper with the pairs trading return reported in table 3 of GGR, we see our returns are comparable.

However, when we define a change in illiquidity as the difference between the average proportional effective spread computed during the five days before divergence and the average proportional effective spread computed during the estimation period, we find a positive relationship between it and the profits from the subset of pairs with smaller average market capitalization with the cream-skimming strategy but no statistically detectable relationship between it and the standard strategy. In summary, this table provides evidence that some of the short-term profits from pairs trading are rewards for providing immediate liquidity and that the long term profits are larger among illiquid stocks.

5.3 Profitability and Idiosyncratic News

Pairs trading has two key events: divergence and convergence. Here we examine whether characteristics of a pair's divergence are related to its convergence. We have argued early that the profits from pairs trading are related to the information event which creates divergence. In particular, the profits from pairs trading should be small if the divergence event is caused by idiosyncratic news to a constituent of the pair and should be large if the divergence is caused by common news in the presence of market frictions.

To examine the effect of idiosyncratic news events, we use articles from the Dow Jones News Service retrieved from Factiva to identify corporate news stories about stocks in the pair and form portfolios based on whether there was news on the day of divergence. There are two major empirical issues related to the application of Factiva news database. First, as noted by Tetlock (2008) and Vega (2006), there a distinction between "news" and "coverage". News refers to the once non-public information which becomes publicly known upon reporting; but coverage refers reprinting or repackaging previously publicly available information. To decipher real news events from simple coverage, for each stock in the pair we calculate the standard deviation of market model adjusted excess returns over the past 21 days before divergence. If either stock in the pair has an abnormal return, we look to see if it also has a news story. Only when there is both a news story and an abnormal return, we designate that there is a piece of news, rather than press coverage.[7] Second, as many authors have found (D'Avolio, 2003; Fang and Peress, 2008; and Engelberg, 2008), media coverage of firms is strongly related to firm size. Therefore, before constructing portfolios we first sort by the size of the firm to disentangle the size effect.

Our results are reported in Table 5. "No Abnormal Return" means neither stock in the pair had an absolute excess return on the day of divergence that was greater than two historical standard deviations. "News" means that at least one stock in the pair had an abnormal return on the day of divergence and had a story in the Dow Jones News Service. "No News" means that at least one stock in the pair had an abnormal return on the day of divergence but that (or those) stock(s) did not have a news story. Table 5 illustrates that the profits from a standard pairs trading strategy are smaller when a member of the pair has news on the day of divergence and that this differential profitability is both economically and statistically significant. For large (small) stocks the difference

[7] This time-series identification approach resembles the approach in Vega (2006).

in monthly alpha is 34 (30) basis points (bpts).

Because the news variable may be correlated with other variables, we perform a cross-sectional regression which allows us to determine if our result is robust to including several control variables. Every pair opening is an observation and the left hand side variable is the total return to the long/short position in the pair. Foreshadowing some of the results in Table 9, the univariate results hold up well even after we control for other firm characteristics like market capitalization, book-to-market, turnover, and past returns accumulated over various horizons.

5.4 Profitability and Common Information

So far we have focused on firm-specific information. Of course, not all information is of this form. Two steel firms may have news about their respective firms (like labor disputes or equity/debt issues) but there also may be news about the industry in which they operate (like traded steel prices or proposed regulation) that affect both firms. Here we consider how this kind of "common information" is related to the returns from pairs trading.

We extend Mech (1993), Chordia and Swaminathan (2000), Hou and Moskowitz (2005), and Hou (2006) by computing the average delay of a firm's stock price to industry shocks which we call our "industry information diffusion measure". At the end of December of each year, we regress each individual stock's weekly returns on a contemporaneous return and prior four weeks' returns of the market and industry portfolios over the previous three years,

$$r_{i,t} = \alpha_j + \beta_0 R_{M,t} + \delta_0 R_{I,t} + \epsilon_{i,t}, \tag{6}$$

$$r_{i,t} = \alpha_j + \beta_0 R_{M,t} + \sum_{n=1}^{4} \beta_n R_{M,t-n} + \delta_0 R_{I,t} + \epsilon_{i,t}, \tag{7}$$

$$r_{i,t} = \alpha_j + \beta_0 R_{M,t} + \sum_{n=1}^{4} \beta_n R_{M,t-n} + \delta_0 R_{I,t} + \sum_{n=1}^{4} \delta_n R_{I,t-n} + \epsilon_{i,t}. \tag{8}$$

where the industry portfolio's construction follows the Fama and French (1997) twelve-industry industry classification, and the industry portfolio returns are taken from Ken French's website.

After obtaining the regression estimates of (6), (7) and (8), we compute three versions of the industry information diffusion measure. To control for any possible lagged response to the market return, we include four lags of the market return in the regression. The first measure is the fraction of variation of the contemporaneous individual stock returns explained by lagged industry portfolio returns. That is, it is one minus the ratio of the R^2 from the regression (7) restricting $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$ divided by the R^2 from the regression (8) with no restrictions.

$$IND\_D_1 = 1 - \frac{R^2_{\delta_n=0,\forall n\in[1,4]}}{R^2} \tag{9}$$

Intuitively, the larger the value of this number, the more return variation is captured by lagged industry returns and the slower the rate of industry information diffusion. Since the IND_D_1 measure does not distinguish between shorter and longer lags or the precision of the estimates, we consider two alternative measures:

$$IND\_D_2 = \frac{\sum_{n=1}^{4} n\delta_n}{\delta_0 + \sum_{n=1}^{4} \delta_n} \tag{10}$$

$$IND\_D_3 = \frac{\sum_{n=1}^{4} n \left[\frac{\delta_n}{se\left(\delta_n\right)}\right]}{\frac{\delta_0}{se\left(\delta_0\right)} + \sum_{n=1}^{4} \left[\frac{\delta_n}{se\left(\delta_n\right)}\right]} \tag{11}$$

where $se(\cdot)$ is the standard error of the coefficient estimates. Following Hou and Moskowitz (2005), we ignore the sign of the lagged coefficients because most of the lagged coefficients are either zero or positive.

For any individual pair, we compute the pairwise industry information diffusion measure by considering the difference of each pair's industry information diffusion measure:

$$DIF\_IND\_D_k = \left|IND\_D_k^1 - IND\_D_k^2\right| \tag{12}$$

where $k = 1, 2, 3$ denotes the version of individual industry information diffusion measure outlined in (9), (10) and (11). We consider the absolute value of the difference of each stock's industry information diffusion measure within a pair because such difference captures the difference in the lead and lag relationship with respect to the common industry level information. Since the results from these three versions of information diffusion measures are qualitatively similar, we choose to focus our attention on $DIF\_IND\_D_{k=3}$ defined by (12), which is derived from IND_D_3 in (11).

Table 6 reports the results for portfolios sorted on the industry diffusion measure given in (11) and (12). For the overall sample, as shown in Panel A, when the difference of the industry information diffusion rates of the stocks in a pair is large, the monthly portfolio return is about 90 basis points (bpts); and when the difference is small the monthly portfolio return is about 50 basis points. The return spreads between the large and small diffusion rate portfolios is about 30 basis points and statistically significant at the one percent level.

Such a difference is unlikely to be entirely driven by the difference of pair's average size. As reported in Panel B and Panel C, when we first split the sample of pairs into two portfolios based on pairwise average market capitalization, most of the return spreads come from large market capitalization pairs rather than small market capitalization pairs. For example, among the large market capitalization pairs, when the difference of the industry information diffusion rates of the stocks in a pair is large, the monthly portfolio return is about 80 basis points (bpts); and when the difference is small, the monthly portfolio return is about 30 basis points. The return spreads between these two portfolios are 40 basis points and statistically significant at one percent level. For the small market capitalization pairs, though as the difference of information diffusion rates for the underlying stocks increase, monthly returns from the pairs portfolios increase as well and there is not much spread among these portfolios. This is likely due to the fact that among small market capitalization pairs, there is not much difference in the industry information diffusion measure. In summary, Table 6 demonstrates that when the two stocks in the pair have large (small) differences in diffusion rates, the profits to a pairs trading strategy are also large (small). There is evidence that when common information diffuses into stocks at differential rates, it can create the prices of related stocks to temporarily move apart.

By taking an approach similar to Hong, Lim and Stein (2000), we consider two alternative and indirect measures to capture the relative information diffusion rates. Hong, Lim and Stein (2000) test whether the slow information diffusion model of Hong and Stein (1999) can explain the momentum anomaly by forming portfolios based on analyst coverage. They find that - controlling for firm size - if a firm has fewer analysts then it is more likely to experience momentum. Momentum is a univariate strategy so that it is natural for Hong, Lim and Stein (2000) to compute the number of analysts that cover a particular firm; pairs trading is a bivariate strategy so that it is natural for us to compute the number of analysts that cover *both* firms. This first measure is called "common analyst coverage". For those pairs where both stocks are covered by analysts from the same brokerage house, there should be relatively small difference in information diffusion rates. In this case, the profits from pairs trading should be smaller. We also construct a measure based on common institutional holdings. For those pairs where both stocks are held by the same institutional investors, there should be a relatively small difference in information diffusion rates. In this case, the profits from pairs trading should also be smaller.[8]

Table 7 and Table 8 provide evidence consistent with these hypotheses. Pairs with few common analysts outperform pairs with many common analysts by 40 basis points per month and the spreads are statistically significant at one percent level. Similarly, pairs with few common institutional holdings outperform pairs with more common institutional holdings by 50 basis points (bpts) per month and such spreads are statistically significant at one percent level. Splitting the sample based on the average market capitalization of the pairs reveals that the most of the spreads between high versus low common analyst coverage or common institutional holding portfolios come from the large average market capitalization pairs. Large market capitalization pairs with few common analysts outperform pairs with many common analysts by 30 basis points per month and the spreads are statistically significant at the five percent level. In addition, large market capitalization pairs with few common institutional holdings outperform pairs with more common institutional holding by 20 basis points (bpts) per month and such spreads are statistically significant at the ten percent level.

[8] An interesting question is why some pairs sometimes covered by the same brokerage (or held by the same institutional investors), but some pairs are not covered by the same brokerage house (or held by the same institutional investors) at some other times. This may be due to categorical thinking in the investment process as suggested by Mullainathan (2000), and Barberis and Shleifer (2003).

5.5 Exposure to Liquidity Risk Factors

5.5.1 Equity Market Liquidity Risk

Several recent theoretical papers consider how returns to convergence trading is related to market liquidity. For instance, Kondo (2008) suggests that liquidity risk (Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005) is related to the arbitrageur's profits, which have a left-skewed distribution. Kondo's model is also related to funding liquidity and the market liquidity channel in Brunnermeier and Pedersen (2008), which underscores the importance of funding liquidity risk.

Table 9 considers several alternative factor models with different equity market liquidity factors. Panel A considers the pairs trading with a holding horizon of ten days; and Panel B considers the pairs trading with a holding horizon of six months. In the columns (1) and (2) of each panel, the liquidity factors are respectively the value-weighted version and equally-weighted version of Pastor-Stambaugh liquidity factor (Pastor and Stambaugh, 2003). In columns (3) and (4) of each panel, the liquidity factors are the fixed-cost and variable-cost components of the spreads liquidity constructed by Sadka (2006). Due to availability of liquidity risk factors, the sample period for regressions (1) and (2) is from January, 1993 to December, 2004; and the sample period for regressions (3) and (4) is from January, 1993 to December, 2005. The equally weighted version of the liquidity factors in Pastor and Stambaugh (2003), and the liquidity risk factor from the variable-costs component of total spreads in Sadka (2006) are negatively correlated with the returns from pairs trading, especially when the holding horizon is relatively short such as ten days. The ability of these factors to explain the returns in the time-series regressions is limited though, especially for short holding horizon. The *R-squared* from these regression are usually quite low, especially for ten day holding horizon - about 6 to 9%; but increases to about 30% for holding horizon of six months. The alphas of pairs trading (the intercept terms of these regressions) hardly change with these additional liquidity risk factors.

5.5.2 Funding Liquidity Risk and Other Macro Risks

Brunnermeier, Nagel and Pedersen (2008), and Asness, Moskowitz and Pedersen (2008) show foreign exchange carry trades and value/momentum strategy returns are related to funding liquidity risk. To explore the pairs trading exposure to this macro liquidity risk, we adopt a funding liquidity risk proxy, the U.S. Treasury-Eurodollar (TED) spreads proposed by these authors. Krishnamurthy and Vissing-Jorgensen (2008) suggest the AAA/T-bill spreads capture the convenience yields of the U.S. treasury securities to the investors. We adopt the AAA/T-bill spreads to proxy for the demand-side driven liquidity premium in the economy. To link the profits from pairs trading to the long-run consumption risk (Bansal and Yaron, 2004; Parker and Julliard, 2005; Malloy, Moskowitz and Vissing-Jorgensen, 2007; Jagannathan and Wang, 2007). To capture the business cycle risk, we use default spreads, which is computed as the Moody's BAA minus AAA bond yield spreads (Fama and French, 1999; Jagannathan and Wang, 1996; among others).

To construct US Treasury-Eurodollar (TED) spreads (Asness, Moskowitz and Pedersen. 2008), we obtain the 3-month LIBOR rate (in US$) from the ECONSTATS database, and 3-month Treasury Bill rates from Federal Reserve Board H15 release. We also obtain the Moody's BAA, and Moody's AAA corporate bond rates from Federal Reserve Board H15 release to construct BAA/AAA spreads and AAA/T-Bill spreads. In the construction of the long-run consumption growth rates, per capital real nondurable goods quarterly consumption are derived from Table 2.1 (line 6, real nondurable goods quarterly consumption) and Table 2.3.6 (line 38, population) National Income and Product Account (NIPA) database. Long-run consumption growth is the future three-year growth in consumption, measured as the sum of log quarterly consumption growth from quarter q to $q + 12$ (both inclusive).[9]

The results are reported in Table 10. In general, the exposures of pairs trading to long-run consumption growth, AAA/T-bill spreads and default spreads are low and statistically insignificant. However, the exposure to the U.S. Treasury-Eurodollar (TED) spreads is high and statistically significant. This is the case for the strategy with a holding period of up to ten days or up to six months. These results suggest that although pairs trading may have little exposure to macroeconomic risk factors, its exposure to the funding liquidity risk is large. When the TED spreads are wide, borrowing is difficult. At the same time, the returns from the pairs trading are high. One interpretation of the relationship is that arbitragers who are enforcing the relative pricing of stocks within the pair are constrained to participate in the market, which leaves the relatively wide spreads.

Taken together, the macro risk factors, particular the macro liquidity risk, could potentially explain some of the profits from the pairs trading.

6 Event-Time Cross-Sectional Evidence

Thus far we have shown differential profitability from pairs trading when we sort on liquidity, news and information diffusion variables in calendar-time portfolios. Here we examine these results in event-time using cross-sectional regressions, where the unit of observation is the opening (divergence) of a pair. The event-time approach has several advantages over the calendar-time approach. First, we can run cross-sectional regressions that allows us to include a battery of control variables so that we are more confident about the economic and statistical significance of our main variables of interest. In the calendar-time approach, sorting along several dimensions creates portfolios with very few pairs, thereby incresing the noise to signal ratio. Second, the event time cross-sectional approach also allows us get a more complete picture of the cycle of pairs trading: (1) the opening of the pairs (2) the evolution of the pairs along the path of convergence, and (3) the termination of the pairs via natural or forced convergence. The analysis required to understand this life cycle is beyond a simple linear factor regression. Therefore, when necessary, we introduce several econometric techniques to meet our needs.

[9] LIBOR rates can be accessed from the following website: http://www.econstats.com/r/rlib_ _d1.htm; Federal Reserve H15 release can be access from the following website: http://www.federalreserve.gov/releases/h15/data.htm; and NIPA data can be accessed from the following website: http://www.bea.gov/bea/dn/nipaweb

6.1 Regression Analysis of Profits from the Pairs Trading

In Table 11 we analyze how short- and long-term pairs trading profits are related to a set of pair characteristics. In the calculation of standard errors, we cluster by industry, year and month, following Petersen (2008).[10] As in section 5, we are particularly interested in how firm-specific idiosyncratic news, common information are related to the pairs trading profits, and how they interact with the underlying institutional share holding structure, information intermediary information production, and liquidity levels.

We use a set of standard control variables, including pairwise average book-to-market equity, the logarithm of pairwise average market capitalization, and pairwise past cumulative returns at the horizons of one month, one year and three years. At the individual stock level, these are shown to be related to future returns (Brennan, Chordia and Subramayahm, 1998). On average, pairs of small stocks and growth stocks earn higher pairs trading profits. Although calendar time pairs trading profits are negatively correlated to the momentum factor as shown in Table 2, we find little evidence in the time-series cross-sectional regressions. The pairwise average cumulative 12-month returns are not statistically significant. Comparing Panel A and Pane B, we see that pairs with low past one-month returns earn higher profits, especially when we hold the position for up to 6 months. Notice that more volatile stocks earn higher returns in both long- and short- horizons. That could be due to Avg_mRetVola proxying for illiquidity of the stock that is not adequately captured by the Avg_PESPR and Avg_PESPR_Change variables, or due to arbitrage requiring a higher reward because of the higher risk in pairs involving more volatile stocks.

Table 11 also provides evidence about how liquidity and trading volume influence pairs trading profits. To capture the level and change of liquidity, we introduce the pairwise average proportional effective spreads (*PESPR*) estimated during the portfolio formation period, and the change of pairwise average proportional effective spreads (ΔPESPR) five trading days leading to the divergence of the pairs. We also consider average turnover rates estimated during the portfolio formation period, and change of average turnover rates five trading days leading to the divergence of the pairs. Our results are largely consistent with the prior literature. We find that, depending on the return horizons, the level and the change of turnover and proportional effective spreads are related to pairs trading profits in an interesting way. With the 10-day holding restriction, the only variable that is reliably related to the profits from pairs trading is the pairwise average change of the proportional effective spreads. At the longer horizon of six months, both the pairwise average proportional effective spreads and pairwise average turnover are related to profits from pairs trading. Stocks with higher pairwise average proportional effective spreads and low pairwise average turnover earn higher profits. The level of liquidity, captured by turnover and the level of spreads are related to profits from pairs trading in a longer horizon, which suggests non-information driven liquidity demand plays an important role in explaining returns accrued to pairs trading. What is interesting

[10] We also compute the standard errors using Fama-MacBeth approach by first estimating a pooled regression monthly then average the monthly regression coefficients to compute the Fama-MacBeth regression coefficients. The results are qualitatively similar so we present the regression results clustered by year, month and industry throughout.

is that, at short-term, the change of liquidity level, or the liquidity shock, subsumes the level of liquidity in explaining the returns of pairs trading. This is consistent with the model of Campbell, Grossman and Wang (1993), which emphasize the temporary nature of liquidity demand shock and its relation to asset prices.

We also find that the idiosyncratic news variable is significant at both the short-term and long-term horizons. It is statistically significant at the five-percent level when the pairs are forced to close in ten trading days and it is significant at one-percent level when pairs are forced to close in six months. In both cases, it is also economically significant. On average, for the ten-day holding horizon, pairs with news earns 40 basis points *less* than otherwise similar pairs; and for the six-month holding horizon, pairs with news on one of the constituent stock earns 120 basis points *less* than otherwise similar pairs. In sharp contrast, pairs with just media coverage - but not news - do not seem to earn returns any different from stocks without any media coverage. These results provide confirmatory evidence that idiosyncratic news creates permanent differences in the prices of the stocks in the pair and therefore less profitability from a pairs trading strategy.

Fourth, the common institutional holding (*Common_ Holding*) and the common analyst coverage (*Common_ Analyst*) measures are related to the profits from pairs trading. In the second columns of Panel A and Panel D, we consider a continuous version of these two variables as defined in Section 4.2, which essentially count how many institutions hold both stocks in the pair, and how many brokerage houses cover both stocks in the pair. In the third and fourth columns of Panel A to Panel D, we consider binary version of these variables, which take the value of one if the number of institutions holding both stocks in the pair is less than the sample median (about 63 institutions), or if the number of brokerage house covering both stocks in the pair is less than the sample median (about 2 brokerage houses). At both short and long horizons, the institutional ownership structure of the pair matters for the profits from pairs trading. Columns three in Panel A and Panel B indicate that, compared to otherwise similar pairs, if there are few institutional investors holding both stocks within the pair during the quarter prior to the divergence of the pair, the pairs trading profits increase about 70 to 80 basis points on average per pair. The impact of the information intermediary structure on pairs profits are weaker. If there are fewer than two brokerage houses covering both stocks within the pair, the profits from pairs trading are indeed stronger: the magnitude is about 60 basis points per pair more for the longer holding horizon. However, the number of brokerage houses covering both stocks of the pair has no impact on the profits for the shorter horizon. These results are consistent with the idea that institutions can impound information into prices more quickly (which is why institutional ownership of the paired firms are important for the short-horizon) and the information produced by intermediaries like analysts takes more time to be impounded into prices (which is why analyst coverage is important for the long-horizon).

Regressions reported in Panel C and Panel D are similar to those in Panel A and Panel B of Table 11, except we include the industry information diffusion measure (*DIF_ FF12_ D3*), and its interactions with liquidity (*Liquidity*), institutional ownership structure (*Common_ Holding*), information intermediary structure (*Common_ Analyst*), and size (Size) binary variables. In all cases, the industry information diffusion measure are statistically significant. The larger the value of the industry information diffusion measure, the larger the difference of individual stock's speed of response to industry common information within the pair, and the larger the profits from pairs trading. Furthermore, the interactions between industry information diffusion measure and liquidity (*Liquidity*), institutional ownership structure (*Common_Holding*), information intermediary structure (*Common_Analyst*) are all statistically significant at least five-percent significance level for one of the holding horizons. That is, the impact of the difference of individual stock's speed of response to industry common information is particularly strong among less liquid stocks, stocks with fewer common institutional holding or analyst coverage.

Finally, we point out that the interaction between pairwise average size and the industry information diffusion measure is insignificant. This is consistent with our interpretation that even though liquidity (*Liquidity*), institutional ownership structure (*Common_Holding*), information intermediary structure (*Common_Analyst*) may be related to the average size of the pair, they seem to capture something more than the size effect. Moreover, they represent market frictions in the form of transactions costs and information costs that exacerbate the differential response of paired stocks to common information which we have argue is a channel by which profits are made in pairs trading.

6.2 Probability of a Pair Opening for Investment

We begin our analysis of the life cycle of pairs trading with the binary divergence event (the "opening" event). This is event-day in which the pair's squared price difference is larger than twice the standard deviation of the squared price differences during the estimation period for that pair. The logistic regression analysis of the pair's daily opening probability is reported in Table 12. On each day and for each pair, we consider whether the pair remains "closed" or becomes "open", and relate this divergence event to a set of pair-specific characteristics using a logistic regression. In the calculation of standard errors, we cluster by industry, year and month, following Petersen (2008).

As shown by the first regression in the Panel A of Table 12, if eligible for trading, the pair consisting of stocks associated with higher average proportional effective spreads, sudden increase in the proportional effective spreads, lower turnover rates, sudden increase in turnover rates, higher past two-to-three year cumulative returns, lower market capitalization, lower book to market equity, and higher idiosyncratic volatilities is more likely to open on a particularly day.

Regressions 2 to 4 in Panel A show that the common institutional holding (*Common_Holding*) and the common analyst coverage (*Common_Analyst*) measures are related to the probability of pair opening either individually or together. In these regressions, the common institutional holding (*Common_Holding*) and the common analyst coverage (*Common_Analyst*) are continuous variables. Serving as a robustness check, regressions 5 is similar to regression 4, but the institutional ownership structure (*Common_Holding*) and the information intermediary structure (*Common_Analyst*) are categorical variables. These three regressions show that the probability of a pair opening is significantly lower for those pairs with both stocks held by a larger number of *the same institutions*, or covered by a large number of *the same analysts*.

Regressions 6 in Panel A adds another binary variable (*Size_ Rank*) to the independent variables in regression 5, which takes the value of one if the pairwise market capitalization is lower than the sample median. After inclusion of this variable, the magnitude and statistical significance of the institutional ownership structure (*Common_ Holding*) and information intermediary structure (*Common_ Analyst*) categorical variables do not change significantly. Regression 7 excludes the institutional ownership structure (*Common_ Holding*) and information intermediary structure (*Common_ Analyst*) categorical variables from regression 6. The magnitude and statistical significance of size (*Size_ Rank*) categorical variable remain similar to those in regression 6. Therefore, it is clear that the institutional ownership structure (*Common_ Holding*) and information intermediary structure (*Common_ Analyst*) categorical variables provide additional information beyond the size.

In Table 12, the specification of regressions 1 to 5 in Panel B is similar to regression 4 in Panel A. The difference lies in the additional industry information diffusion measure (*DIF_ F12_ D3*), and its interaction with liquidity (*Liquidity*), institutional ownership structure (*Common_ Holding*), information intermediary structure (*Common_ Analyst*), and average pairwise market capitalization (*Size_ Rank*). With the exception of the interaction term between the industry information diffusion measure (*DIF_ F12_ D3*) and the institutional ownership structure (*Common_ Holding*), the industry information diffusion measure and its interaction with liquidity, information intermediary, and average pairwise market capitalization are statistically significant at one percent level. Regression 1 show the daily opening probability of the pair increases, when the difference in the relative speed of prices adjustment to industry common information of the stocks in a pair decreases (*i.e.*, the larger the value of *DIF_ F12_ D3*). Regressions 2, 4 and 5 show the relationship between the daily opening probability of the pair and the information diffusion measure is stronger among the pairs which are less liquid, covered by smaller number of the same analysts, or smaller average pairwise market capitalization.

6.3 Determinants of the Time-to-Convergence

After a pair opens, we analyze its time-to convergence using survival analysis. For that purpose we use the survival analysis model that Lo, MacKinlay and Zhang (2002) use for examining the time to execution of limit orders. Survival analysis is a statistical technique commonly used for analyzing life-times and failure-times. Here we use it to analyze the time-to-convergence. For convenience, we provide a brief description of survival analysis in Appendix A, based on Lo, MacKinlay and Zhang (2002).

Table 13 reports the survival analysis of time-to-convergence of the pairs in our sample. Several noteworthy observations emerge. First, in all cases, the scale and shape parameter estimates are all statistically significant at one percent level, which hints that the choice of generalized gamma distribution as the baseline distribution is preferred to some other more restrictive distribution assumptions.

Second, pairs consisting of stocks associated with higher average proportional effective spreads, sudden increase in the proportional effective spreads, lower turnover rates, sudden decrease in turnover rates, higher past twelve-month returns, lower market capitalization, lower book to market equity, and higher idiosyncratic volatilities have shorter times-to-convergence and, thus, less horizon risk. On the one hand, it is more difficult for the rational arbitragers to arbitrage away the anomalous returns from these stocks. On the other hand, they are exactly those stocks which are less liquid, smaller, and more volatile stocks. In another word, these stocks have higher "holding risk" (Pontiff, 2006). Thus, our survival analysis illustrates that there is a delicate balance between the horizon risk and holding risk. Our survival analysis clearly shows that if one is going to reduce the horizon risk; one may have to incur more holding risk; therefore, the convergence trade is far from being risk-free in an operational sense. To the best of our knowledge, the trade-off between horizon risk and holding risk has not been discussed in the literature.

Third, in the previous sections, we have shown that the pairs with idiosyncratic news on at least one of the pairs at the time of divergence earns significantly lower returns. Table 11 reveals that at least part of the reason for the declines in profitability is the increase in the time-to-convergence. For instance, according to the estimates from Panel A and equation (18), for a holding horizon of ten trading days, the expected time-to-convergence for the pairs with news is about 28.40% ($= \exp[0.25 \times (1-0)] - 1 = 28.4\%$) longer than otherwise similar stocks without news (including both stocks without any media coverage, and the pairs with only coverage but not news); and according to the estimates from Panel B and equation (18), for a holding horizon of six months, the time-to-convergence for the pairs with news is about 52.20% ($= \exp[0.42 \times (1-0)] - 1 = 52.20\%$) longer than otherwise similar stocks without news. Clearly, the time-to-convergence difference due to stock level idiosyncratic news is both statistically significant and economically important.

Slow information diffusion hypothesis (Hong and Stein, 1999) suggests idiosyncratic information should cause permanent differences in the arbitrage spread and may even lead to the spread widening as the idiosyncratic information diffuses. To test this hypothesis, we create a binary variable, *NegativeNews*, which takes the value of one if on the divergence date the news associated with the stock is negative; and zero otherwise. Then we interact the News binary variable. If the time-to-convergence is positively related to the interaction term, $News \times NegativeNews$, then it indicates that drift effect is particularly pronounced for the "bad news" pairs, and the evidence is consistent with the "bad news" travel slowly story. In unreported regressions, we indeed find evidence consistent with the hypothesis. In all specifications, the interaction term is statistically significant at one percent level, and all other regression coefficients are qualitatively similar to those reported in Table 13. Moreover, the point estimates for the interaction terms are about 0.27 for the holding horizon of 10 days, and 0.29 for the holding horizon of 6 months. In contrast, the point estimates for the News variable are about 0.20 for the holding horizon of 10 days, and 0.36 for the holding horizon of 6 months. A simple back of the envelope calculation shows that the expected time-till-convergence for the pairs with "bad news" is about 31% to 33% longer than otherwise similar stocks with "good news" depending on the holding horizon we examine.

Fourth, the institutional ownership structure ($Common\_Holding$) and the information intermediary structure ($Common\_Analyst$) are related to the time-to-convergence. Columns (3) and (4) from Panel A and Panel B show that, if there are few institutions holding both stocks in a pair, the expected time-till-convegence decreases by at least 25% ($= \exp[0.223 \times (1-0)] - 1 = 25.00\%$). Similarly, if there are few analysts from the same brokerage house covering both stocks of a pair, the expected time-till-convegence decreases by at least 8.33% ($= \exp[0.08 \times (1-0)] - 1 = 8.33\%$).

Finally, Panel C and Panel D of Table 13 show that the larger the industry information diffusion measure, i.e., the larger the difference between the stock's speed of adjustment to the common information, the shorter the expected time-to-convergence. Such effect is especially strong among small stocks, less liquid stocks, stocks with few common institutional holding, and stocks with few common analyst coverage.

6.4 Divergence Risk

A typical arbitrager engaged in convergence trading like pairs trading faces the possible widening of arbitrage spreads. Widening arbitrage spreads expose the arbitragers to margin calls, which may require partial or complete liquidation of his position or additional capital infusion. In either case, the profits from convergence trading decreases. Following common practice, we will refer to this type of risk as "divergence risk". For a given arbitrager, when there spreads widen, the arbitrager has the choice of complete or partial liquidation or capital infusion; and the arbitrager also has the choice of liquidating some assets instead of others. Without detailed assumptions on the cost of arbitrage capital and capital constraints, it is difficult to quantify directly divergence risk impact on the total return accrued to the convergence trade. To avoid such potentially *ad hoc* assumptions, we choose to model the occurrence of spread-widening events during the path of convergence trades, and treat such occurrence as an measure of divergence risk.

For any particular pair at the end of trading day t, when the arbitrage spreads $x(t)$ increase compared with the prior maximum spreads since the establishment of the position, i.e., $x(t) > \max[x(1), x(2), ..., x(t-1)]$, we define a spread-widening event occurs. Statistically, the occurrence of spreads widening events is a set of non-negative discrete random variable. To accommodate the data feature, we use the zero-inflated negative binomial regression model is a natural candidate. A brief description of the statistical event count model is presented in Appendix B, and the reader is referred to Cameron and Trivedi (1999) for more details.

Table 14 reports the estimates of the zero-inflated negative binomial regression. The dependent variable is the count of spreads-widening events during six-month holding horizon. The independent variables in the *zero-inflation equations* include a constant term, a binary indicator variable taking value of one if the pair converges in ten days; and zero otherwise, and the change of average pairwise proportional effective spreads prior to convergence. The independent variables in the *main equations* are similar to those used in the survival analysis regressions outlined in the previous section.

To determine the model specifications, we tested several alternative models by including several variables characterizing a pair in the auxiliary zero-inflation equation. However, in the presence of the binary indicator variable describing whether the convergence happens in ten days or not, most of the other variables are statistically insignificant at conventional level; and inclusion of these additional variables do not significantly change the estimates from the main regression, so we choose the simpler model. The binary indicator variable describing whether the convergence happens in ten days or not is the only variable that is always highly significant in the zero-inflation equation. Since the divergence risk are larger for those pairs which do not converge in a relatively short period, this is not entirely surprising.

In unreported tests, we find that Vuong test statistics (Vuong, 1989) indicate that the zero-inflated negative binomial regression model in preferred to the negative binomial regression model against. We also find that the likelihood ratio test statistics reject the dispersion parameter $\alpha = 0$ at one-percent significance level, which indicates that the negative binomial regression model is preferred to the Poisson regression model.

Conceptually, divergence risk and horizon risk describe different aspects of the arbitrage risks associated with the convergence trade. For the convergence trades with fixed time-till-convergence, i.e., no horizon risk, there still could be substantial divergence risk before the convergence (see Liu and Longstaff, 2006). However, divergence risk and horizon risk are not unrelated. As can be seen from Table 14, divergence risk is lower for pairs with higher average proportional effective spreads, lower turnover, higher past twelve-month or three-year cumulative returns, higher idiosyncratic volatilities, larger difference in the speed of adjustment to common industry information, less common holding by institutions, less common coverage by the sell-side analysts, and among pairs without news. Further, when the stocks within the pair are less liquid, or less likely to be held by the same institution, the speed of adjustment to common industry information impact on the divergence risk are stronger.

7 Robustness Check

7.1 Default Risk and Pairs Trading

There is some concerns that unrealized bankruptcy risk may be related to the profits from pairs trading. We use the Expected Default Frequency (EDF^{TM}) produced by Moody's-KMV as a proxy for default risk at individual stock level to explicitly examine such possibility.[11] We compute the default risk at the pair level by averaging individual stock's EDFs, or taking the maximum of individual stock's EDFs within the pair. Then we apply the pairwise EDF in the cross-sectional regressions which relate the pair's characteristics and pair's total returns. As Da and Gao (2008) suggest, a sudden increase in the default likelihood could induce a clientele change and consequently a short-term return reversal effect. To avoid such confounding effect, we use the pair's EDF value one month before the divergence month. Consistent with findings in Gatev, Goetzmann and

[11] We thank Moody's-KMV for making the EDF^{TM} data available to us.

Rouwenhorst (2006), the regression results show that the pair's EDF is not related to the pairs trading profits for both ten-day and six-month horizons (not reported).

7.2 Short-sale Constraints and Pairs Trading

Short-sale constraints, such as prohibitively expensive short-rebate rates, could make pairs trading not implementable. To examine such possibility, following the identification proposed by Asquith, Pathak and Ritter (2005), Chen, Hong and Stein (2002) and Nagel (2005), we consider (*i*) the minimal institutional holding of the constituents of the pair, and (*ii*) zero holding of institutional holding of the constituents of the pair as proxy variables for the short-sale constraints. However, we did not find either of the proxy for short-sale constraints are related to the return and risk of pairs trading. This is likely due to the fact that the stocks in our sample are generally large (the average market capitalization of the stocks are about 60th percentile in terms of NYSE size percentile breakpoints), and short-sale constraints are not a major friction. For example, D'Avolio (2002) documents that hard-to-borrow stocks almost exclusively concentrate among the smallest size decile (based on NYSE size decile breakpoints) or low priced (less than five dollars).

We also consider the stocks of the pair have traded options or not. In practice, instead of directly borrowing shares, one can construct so-called "synthetic shorts" using options (Battalio and Schultz, 2006). Using the Ivy OptionMetric database, we find that more about 99.76% of the pairs positions opened, there are options traded on the organized exchange for both stocks of the pair.

8 Summary and Concluding Comments

This paper investigates the source of profits from pairs trading. The following table summarizes the main results from our empirical analysis by describing how increases in the values of certain variables affect total pairs trading profits, the probability that a pair will open, the horizon risk, the divergence risk, and the convergence speed. In this table, "+" denotes that certain variables relate *positively* to trading profits, the probability that a pair will open, the horizon risk, the divergence risk, the convergence speed and arbitrage risk; "−" denotes a *negative* relationship; and "*n.s.*" denotes a *statistically insignificant* relationship. Several interesting findings emerge.

First, when there is idiosyncratic news about at least one stock within the pair, the total profits from pairs trading decreases even though the news creates potential opportunities for pairs trading since it is more likely that the pair may diverge. While idiosyncratic news events are more likely to make pairs diverge, they increase the horizon risk and divergence risk to risk to the arbitrageur and slow the speed of pair convergence.

Second, the level of liquidity and short-term changes in liquidity ("liquidity shock") proxied by $PESPR$ and $\Delta PESPR$ contribute positively to the total profits - which arises because of an increase in opening probability, a decrease in horizon and divergence risk, and an increase in convergence speed. However, they are also associated with increases in arbitrage risk.

Variable (↑)	Total Profits	Opening Probability	Horizon Risk	Divergence Risk
PESPR	+	+	−	−
ΔPESPR	+	+	−	*n.s.*
Turnover	−	−	+	+
ΔTurnover	*n.s.*	+	+	*n.s.*
News	−	+	+	+
B/M	−	−	+	+
Size	−	−	+	+
Volatilities	+	+	−	−
Common_Holding	−	−	+	+
Common_Analyst	*n.s.*	−	+	*n.s.*
DIF_F12_D3	+	+	−	−

Third, the difference in the relative speed of adjustment to common industry information is strongly related to pairs trading profits. Large differences in pairwise information diffusion rates contribute to the return because it creates trading opportunity, decreases the horizon risk and divergence risk, and also increases the speed of convergence. However, these are the situations when arbitrage risk - in particular liquidity and price impact - may be high.

Fourth, the information diffusion rates interact with size, liquidity level, and the underlying institutional ownership and information intermediary in a predictable way. The impact of the information diffusion rates are stronger among small, less liquid stocks, which are less likely to be held simultaneously or covered simultaneously by the same institution or sell-side analyst.

Taken together, we have documented that the profitability from pairs trading is strongly related to the way information diffuses across the stocks in the pair and the frictions which stifle this information flow. We have also highlighted the importance of identifying a variety of risks that an arbitrageur faces when he executes a pairs trading strategy. What is particularly interesting is that the table indicates arbitrage risk - including execution risk and holding risk - seems to move in the opposite direction as horizon risk and divergence risk. This suggests an arbitrageur may face difficult trade-offs when executing the pairs trading strategy. The interaction between these risk types and the optimal investment behavior of the arbitrageur when facing different dimensions of risk appears to be an interesting direction for future research.

References

Acharya, V.V. and L.H. Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Ali, Ashiq, Lee-Seok Hwang, and Mark A. Trombley, 2003, Arbitrage Risk and the Book-to-Market Anomaly, *Journal of Financial Economics* 69, 355-373.

Amihud, Yahov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5, 31-56.

Amihud, Yakov and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Antweiler, Werner, and Murray Z. Frank, 2004, The market impact of corporate news stories, *Working Paper*, Carlson School of Management, University of Minnesota.

Asness Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2008, Value and Momentum Everywhere, *Working Paper*, AQR Asset Management.

Asquith, Paul, Parag A. Pathak and Jay R. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243-276.

Avramov, Doron, Tarun Chordia and Amit Goyal, 2005, Liquidity and autocorrelations in individual stock returns, *Journal of Finance* 61, 2365-2394.

Avramov, Doron, Tarun Chordia and Amit Goyal, 2006, The Impact of Trades on Daily Volatility, *Review of Financial Studies* 19, 1241-1277.

Andrade, S., Di Pietro, V., Seasholes, M., 2005, Understanding the profitability of pairs trading, *Working Paper*, Northwestern University.

Ball, Ray, and Philip Brown, 1968, An Empirical Evaluation of Accounting Income Numbers, *Journal of Accounting Research* 6, 159-178.

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

Battalio Robert H., and Richard R. Mendenhall, 2007, Post-Earnings Announcement Drift: Timing and Liquidity Costs, *Working Paper*, University of Notre Dame.

Battalio Robert H., and Paul Schultz, 2006, Options and Bubble, *Journal of Finance* 61, 2071 - 2102.

Berk, Jonathan, Richard Green, and Vasant Naik, 1999, Optimal Investment, Growth Options and Security Returns, *Journal of Finance* 54, 1153-1607.

Bernard, V.L., Thomas, J.K., 1989, Post-earnings announcement drift: delayed price response or risk premium?, *Journal of Accounting Research* 27, 1–36.

Bessembinder, Hendrik, 2003, Issues in Assessing Trade Execution Costs, *Journal of Financial Markets* 6, 233-257.

Bossaerts, Peter, 1988, Common Nonstationary Components of Asset Prices, *Journal of Economic Dynamics and Control* 12, 347-364.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns, *Journal of Financial Economics* 49, 345-373.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2008, Market Liquidity and Funding Liquidity, *Review of Financial Studies*, Forthcoming.

Brunnermeier, Markus K., Stefan Nagel, and Lasse Heje Pedersen, 2008, Carry Trades and Currency Crashes, *NBER Macro Annual*, Forthcoming.

Campbell, J. Y., S. J. Grossman, and J. Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

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

Cameron, A. Colin, and Pravin K. Trivedi, 1998, Regression Analysis of Count Data (Econometric Society Mongraphs), Cambridge University Press.

Chan, Wesley S., 2003, Stock price reaction to news and no-news: drift and reversal after headlines, Journal of Financial Economics 70, 223-260.

Chen, Joseph, Harrison Hong and Jeremy C. Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171-205.

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.

Chordia, Tarun and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485 - 518.

Chordia, Tarun, and Bhaskaran Swaminathan, 2000, Trading Volume and Cross-Autocorrelations in Stock Returns, *The Journal of Finance* 55, 913-935.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic Links and Predictable Returns, *Journal of Finance* 63, 1977 - 2011.

Conrad, J. S., A. Hameed and C. Niden, 1994, Volume and autocovariances in short-horizon individual security returns, *Journal of Finance* 49, 1305-1329.

1999, Filter Rules Based on Price and Volume in Individual Security Overreaction, *The Review of Financial Studies* 12, 901-935.

Cox, D.R., and David Oakes, 1984, Analysis of Survival Data, Chapman & Hall, New York.

Da, Zhi and Pengjie Gao, 2008, Clientele Change, Liquidity shock, and the Return on Financially Distressed Stocks, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic based benchmarks, *Journal of Finance* 52, 1035–1058.

D'Avolio, Gene, 2002, The market for borrowing stock, *Journal of Financial Economics* 66, 271-306.

D'Avolio, Gene Michael, 2003, Essays in financial economics, unpublished Ph.D. dissertation, Harvard University.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldman, 1990, Noise Trader Risk in Financial Markets, *Journal of Political Economy.*

Elliott, Robert J., John Van Der Hoek and William P. Malcolm, 2005, Pairs Trading, *Quantitative Finance* 5, 271-276.

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

Fama, Eugene F., Kenneth R. French, 1996. Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene, Kenneth French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *The Journal of Political Economy* 81, 607-636.

Fang, Lily H., and Joel Peress, 2008, Media Coverage and the Cross-Section of Stock Returns, *Working Paper*, INSEAD.

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

Gervais, Simon, Ron Kaniel and Dan H. Mingelgrin, 2001, The High-Volume Return Premium, *The Journal of Finance* 56, 877-919.

Gourieroux, Christian and Joann Jasiak, 2001, Financial Econometrics: Problems, Models, and Methods, Princeton University Press.

Grossman, S., and J. Stiglitz, 1980, On the impossibility of informational efficient markets, *American Economic Review* 70, 393-408.

Hanna, J. D., and Mark J. Ready, 2005, Profitable Predictability in the Cross Section of Stock Returns, *Journal of Financial Economics*, 78, 463-505.

Harris, Larry, 2002, Trading and Exchanges: Market Microstructure for Practitioners, Oxford University Press.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the Determinants of Expected Stock Returns, *Journal of Financial Economics*, 41, 401-439.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance*, Vol 54, 2143-2184.

Hong, Harrison, Terrance Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265-295.

Hong, Harrison, Walter Torous and Rossen Valkanov, 2007, Do industries lead stock markets?, *Journal of Financial Economics* 83, 367-396.

Hou, K., 2006, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20, 1113-1138.

Hou, K., Moskowitz, T., 2005. Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981-1020.

Huang, Roger. and Hans R. Stoll, 1996, Dealer versus Auction markets: a paired comparison of execution costs on Nasdaq and the NYSE, *Journal of Financial Economics* 41, 313-358.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, Journal of Financial Economics 39, 181-208.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of Momentum Strategies: An Evaluation of Alternative Explanations, The Journal of Finance 56, 699-720.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Jurek, Jakub W., and Halla Yang, 2005, Dynamic Portfolio Selection in Arbitrage, *Working Paper*, Harvard University.

Kalbfleisch, J. D., and R. L. Prentice, 2002, The statistical analysis of failure time data, 2nd edition, New York, Wiley.

Kavajecz, Kenneth A. and Elizabeth Odders-White, 2004, Technical Analysis and Liquidity Provision, *Review of Financial Studies* 17, 1043-1071.

Khandani, Amir E., and Andrewq W. Lo, 2007, What Happened to the Quants in August 2007, *Working Paper*, MIT.

Korajczyk, Rorbert A. and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039–1082.

Kondor, Peter, 2008, Risk in Dynamic Arbitrage: The Price Effects of Convergence, *Journal of Finance*, Forthcoming.

Kovajecz, Kenneth A., and Elizabeth R. Odders-White, 2004, Technical Analysis and Liquidity Provision, *Review of Financial Studies* 17, 1043-1071.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2008, The aggregate demand for treasury debt, *Working Paper*, Northwestern University.

Lehmann, Bruce, 1990, Fads, Martingales and Market Efficiency, *Quarterly Journal of Economics* 105, 1–28.

Lesmond, David A., Michael J. Schill and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349–380

Liu, Jun and Francis A. Longstaff, 2004, Losing Money on Arbitrage: Optimal Dynamic Portfolio Choices in Markets with Arbitrage Opportunities, *Review of Financial Studies* 17, 611-641.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, *Review of Financial Studies* 15, 1005 - 1047.

Lo, Andrew W., Craig MacKinlay and June Zhang, 2002, Econometric models of limit-order executions, *Journal of Financial Economics* 65, 31 - 71.

Lo, AW, and AC MacKinlay, 1990, When are Contrarian Profits due to Stock Market Overreaction? *Review of Financial Studies* 3, 175-205.

Loughran, Tim, and Jay Ritter, 1997, The Operating Performance of Firms Conducting Seasoned Equity Offerings, *Journal of Finance* 52, 1823-1850.

Loughran, Tim, and Jay Ritter, 1995, The new issue puzzle, *Journal of Finance* 50, 23 - 51.

Mech, Timothy S., 1993, Portfolio Return Autocorrelation, *Journal of Financial Economics* 34, 307-344.

Mendenhall, Richard R., 2004, Arbitrage risk and post-earnings-announcement drift, *Journal of Business* 77, 875–894.

Menzly, Lior and Oguzhan Ozbas, 2006, Cross-Industry Momentum, *Working Paper*, Marshall School of Business, University of Southern Carlifornia.

Michaely, Roni, Richard H. Thaler, and Kent Womack, 1995, Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift?, *Journal of Finance* 50, 573–608.

Mitchell, Mark, Todd Pulvino, and Erik Stafford, 2002, Limited Arbitrage in Equity Markets, *The Journal of Finance* 57, 551-584.

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of Risk and Return in Risk Arbitrage, *The Journal of Finance* 56, 2135-2175.

Mullainathan, Sendhil, 2000, Thinking through Categories, *Working Paper*, Harvard University.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277-309.

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

Petersen, Mitchell A., 2008, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies*, forthcoming.

Pontiff, Jeffrey, 2006, Costly Arbitrage and the Myth of Idiosyncratic Risk, Journal of Accounting and Economics, 42, 35-52.

Pontiff, Jeffrey, 1996, Costly Arbitrage: Evidence from Closed-End Funds, *The Quarterly Journal of Economics* 111, 1135-1151.

Protter, Philip E., 2005, Stochastic Integration and Differential Equations, 2nd Edition, Springer.

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

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

Schultz, Paul, and Sophie Shive, 2008, Where are the limits to arbitrage? An application to dual-class shares, *Working Paper*, University of Notre Dame

Smith, Brain F. and Ben Amoako-Adu, 1995, Relative Prices of Dual Class Shares, *Journal of Financial and Quantitative Analysis* 30, 223-239.

Tetlock, Paul C., All the News That's Fit to Reprint: Do Investors React to Stale Information?, *Working Paper*, Columbia Business School.

Tetlock, Paul C., 2007, Giving content to investor sentiment: the role of media in the stock market, Journal of Finance 62, 1139-1168.

Vega, Clara, 2006, Stock price reaction to Public and private information, *Journal of Financial Economics* 82, 103-133.

Wurgler, J. and K. Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks?, *Journal of Business* 75, 583-608.

Vuong, Quang, 1989, Likelihood ratio tests for model selection and nonnested hypotheses, *Econometrica* 57, 307- 333.

Xiong, Wei, 2001, Convergence Trading with Wealth Effects, *Journal of Financial Economics* 62, 247-292.

Zingales, Luigi, 1995, What determines the value of corporate votes, *Quarterly Journal of Economics* 110, 1047 - 1073.

Appendix A: A Brief Review of Survival Analysis

In this section, we first present a brief review of survival analysis, which draws heavily from Lo, MacKinlay and Zhang (2002). For a more detailed treatment, see Cox and Oakes (1984), and Kalbfleisch and Prentice (2002).

Let T denote a nonnegative random variable that represents the time until the convergence of a pair - the state of non-convergence. Let $f(t)$ and $F(t)$ denote the probability density function (PDF) and cumulative distribution function (CDF), respectively, of T, where t is the realization of T. Then the survival function is defined as

$$S(t) = 1 - F(t),$$

which is the probably that the non-convergence state starts at time $t = 0$ and is still going on at time t. The probability that the state of non-convergence ends between time t and time $t + \Delta t$, given convergence has not occurred at time t, for small Δt is given by:

$$\lim_{\Delta t \to 0} \Pr(t < T \leq t + \Delta t | T \geq t) = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{S(t)} = \frac{f(t)}{S(t)}. \tag{13}$$

The conditional probability in (13) is usually referred to as the instantaneous hazard rate of T at time t, denoted as $h(t)$, or $h(t) \equiv \frac{f(t)}{S(t)}$.

After assuming a specific parametric distribution of the time till convergence, we can adopt the parametric survival analysis, which allows the maximum likelihood estimation of parameters of interest. Let $(t_1, t_2, ..., t_n)$ denote a sequence of realizations of random variable T, with possible right-censoring. In our context, we know which observations have been right-censored, because the pairs trade have not converged by the end of trading horizon (but they may converge in the future beyond the trading horizon).. Thus we further let $(\delta_1, \delta_2, ..., \delta_n)$ denotes the sequence of censoring indicators, which take the value one if the observation is censored; and zero otherwise. We assume that (t_i, δ_i) are statistically independent of each other given the explanatory variables $\mathbf{X}_i$ of pair characteristics. Then, for given pairs (t_i, δ_i), conditional on a vector of explanatory variables $\mathbf{X}_i$, we can write the maximum likelihood function as

$$L = \prod_{i=1}^{n} f(t_i; \mathbf{X}_i)^{\delta_i} S(t_i; \mathbf{X}_i)^{1-\delta_i} = \prod_{i \in U} f(t_i; \mathbf{X}_i) \prod_{i \in C} S(t_i; \mathbf{X}_i) \tag{14}$$

where U and C denotes the set of uncensored and censored observations.

The dependence of the failure time on the explanatory variables is accommodated by the *accelerated failure time* (AFT) model, which essentially rescales the time. Specifically, the accelerated failure time model takes the form

$$T = T_0 e^{\mathbf{X}\boldsymbol{\beta}} \tag{15}$$

where T is the time till convergence, T_0 is the *baseline failure time* and it follows the *baseline distribution*, $\mathbf{X}$ is a vector of explanatory variables, and $\boldsymbol{\beta}$ is the parameter vector.

The final step is to specify the baseline distribution. Many choices are available, including exponential, Weibull, gamma, lognormal and inverse Gaussian. We choose the generalized gamma distribution because it nests a set of other distributions as a special case. The generalized gamma distribution function has the following probability density function:

$$f(t) = \frac{\lambda |p| \kappa^{\kappa} (\lambda t)^{p\kappa - 1} \exp\left[-(\lambda t)^p \kappa\right]}{\Gamma(\kappa)} \tag{16}$$

and the corresponding survival function:

$$S(t) = \begin{cases} \frac{\Gamma(\kappa, (\lambda t)^p \kappa)}{\Gamma(\kappa)} & \text{if } p < 0 \\ 1 - \frac{\Gamma(\kappa, (\lambda t)^p \kappa)}{\Gamma(\kappa)} & \text{if } p > 0 \end{cases} \tag{17}$$

where $\Gamma(a, b)$ denotes the incomplete gamma function and $\Gamma(a)$ denotes the complete gamma function. When $\kappa = 1$, the generalized gamma distribution degenerates to a Weibull distribution with the probability density function of the form

$$f(t) = \lambda |p| (\lambda t)^{p-1} \exp\left[-(\lambda t)^p\right];$$

and when $\kappa = 1$ and $p = 1$, the generalized gamma distribution degenerates to an exponential distribution with the probability density function of the form

$$f(t) = \lambda \exp(-\lambda t);$$

and when $\kappa = 0$, it degenerates to a lognormal distribution with the following probability distribution function

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left[-\frac{1}{2}\left(\frac{\log(t) - \mu}{\sigma}\right)\right].$$

Combining (**??**), (15) and (16), replacing the scale parameter λ with $\exp(-\mathbf{X}\boldsymbol{\beta})$, we obtain the density function

$$f(t) = \frac{\exp(-\mathbf{X}\boldsymbol{\beta}) |p| \kappa^{\kappa} (\exp(-\mathbf{X}\boldsymbol{\beta}) t)^{p\kappa - 1} \exp\left[-(\exp(-\mathbf{X}\boldsymbol{\beta}) t)^p \kappa\right]}{\Gamma(\kappa)}.$$

It is easy to see that the accelerated failure time (AFT) model assumes that the effect of explanatory variables on the time till convergence is to rescale the failure time itself. The sign and estimates of the coefficient of an individual variable indicate the direction and magnitude of the partial effect of the variable on the conditional probability of pairs convergence. Finally, the conditional expectation of failure time is exponential-linear in the product term of covariates and coefficients. Thus, the ratio of the conditional expectations based on different realizations of the covariates can be expressed as

$$E[T|\mathbf{X}_1]/E[T|\mathbf{X}_2] = \exp\left[(\mathbf{X}_1 - \mathbf{X}_2)^T \boldsymbol{\beta}\right], \tag{18}$$

which will be used in interpreting some of the later results.

Appendix B: A Brief Review of Event Count Model

Assume a discrete random variable Y - in our application, it is the number of times the pair's spread widens compared to previous maximum spread before eventually converging within 6 months (or 10 days for the cream skimming strategy), follows the negative binomial distribution. Then its probability distribution function can be written as

$$\Pr(Y=y) \equiv NB(\alpha,u) = \frac{\Gamma\left(y+\alpha^{-1}\right)}{y!\Gamma\left(\alpha^{-1}\right)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+u}\right)^{\alpha^{-1}}\left(\frac{u}{\alpha^{-1}+u}\right)^{y}, y=0,1,2,... \tag{19}$$

where $\Gamma(.)$ is the Gamma probability distribution, α is the dispersion parameter. The expected number of spread widening events for this distribution is u. , EAs the dispersion parameter α increases, the variance of the negative binomial distribution also increases; and as the dispersion parameter α decreases to zero, the negative binomial distribution degenerates to the familiar Poisson distribution. One can test whether the data come from the Poisson process or the negative binomial process using a likelihood ratio test. The negative binomial regression model incorporates the observed and unobserved heterogeneity in the conditional means, i.e., conditional mean of pair i being different from that of pair j via an exponential mean function

$$u_i(\boldsymbol{\beta}) \equiv \exp\left(\mathbf{X}_i\boldsymbol{\beta}+\epsilon_i\right) \tag{20}$$

which makes use of the linear index function $\mathbf{X}_i\boldsymbol{\beta}$ to take into account the observed heterogeneity and ϵ_i to take into account the unobserved heterogeneity. If one has reasons to believe the excessive amount of zeros of the distribution results from a different data generating process, the negative binomial distribution regression model can be modified into so-called "zero-inflated" negative binomial regression, which allows one to model each of the data generating processes separately. Specifically, the zero-inflated negative binomial regression take the following form

$$y = \begin{cases} 0 & \text{with probability } q \\ NB(\alpha,u) & \text{with probability } (1-q) \end{cases} \tag{21}$$

where the probability q_i, the number of spread widening events taking the value of 0 for the i'th pair is described as a logistic distribution function

$$q_i(\boldsymbol{\theta}) = \Lambda_i(\boldsymbol{\theta}) = \frac{\exp\left(\mathbf{Z}_i\boldsymbol{\theta}\right)}{1+\exp\left(\mathbf{Z}_i\boldsymbol{\theta}\right)} \tag{22}$$

where $\mathbf{Z}_i$ is the set of attributes which may or may not overlap the set of attributes $\mathbf{X}_i$. A Vuong test (Vuong, 1989) can be applied to test the negative binomial regression model against the zero-inflated negative binomial regression model. The expected number of spread widening events for the zero-inflated negative binomial distribution is $(1-q_i(\boldsymbol{\theta})u_i(\beta)$.

Figure 1. Pairs Trading Profitability in Event Time

The lines plots mean returns in event time from a pairs trading strategy. The pairs trading strategy involves matching stock pairs based on normalized price difference over a one year estimation period. Then, during the following year, the strategy looks for instances in which the price of the two stocks in the pair diverge by more than two standard deviations of the price difference established during the estimation period. This is called divergence (convergence is the event when, after divergence, the pairs have no difference in normalized price). When there is divergence the strategy buys the stock that went up and shorts the stock that went down. Event day 0 is one day after this divergence and is meant to control for bid-ask bounce. The dashed line plots the mean return in event time from a pairs trading strategy and the solid line plots the corresponding 5 day moving average.

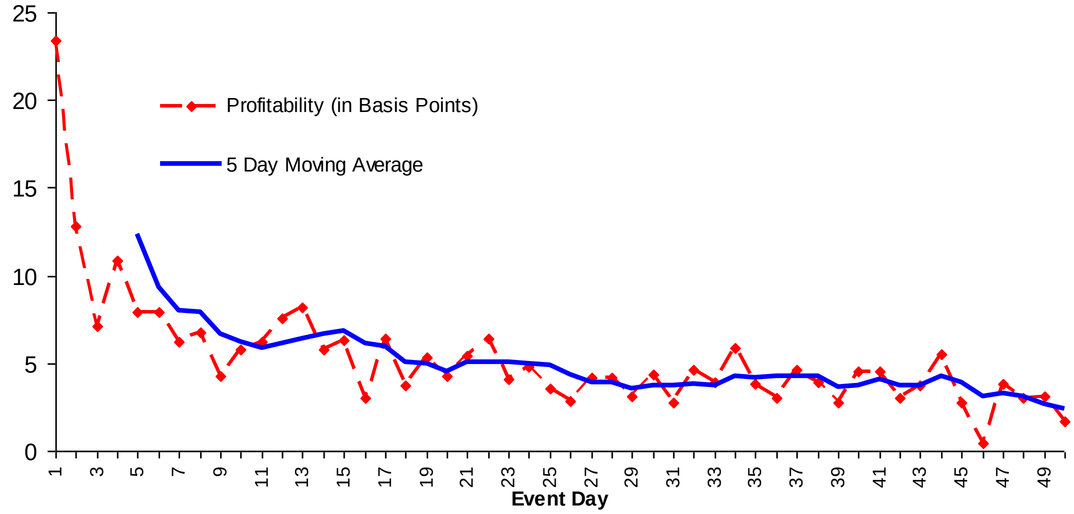
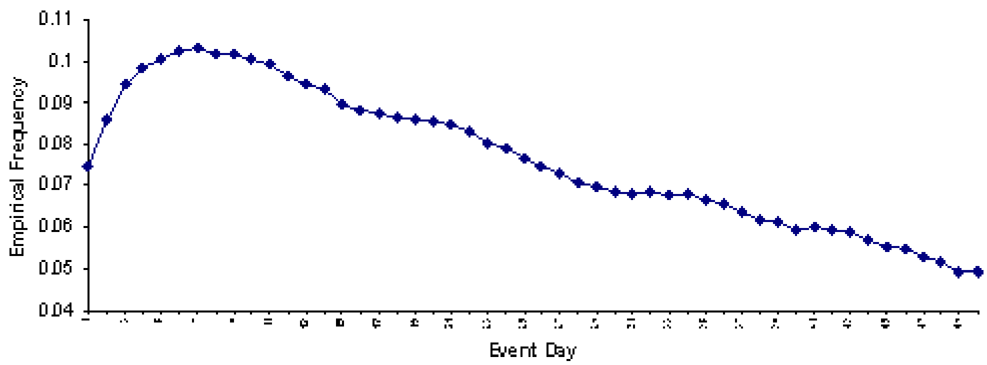


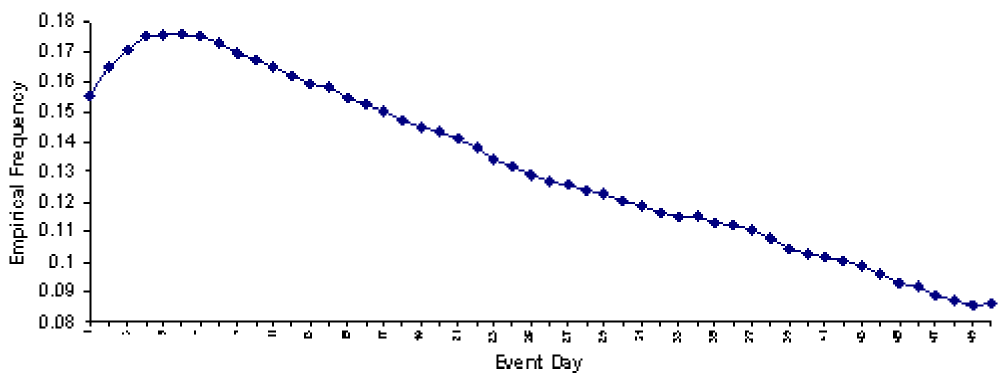
Figure 2. Convergence Probabilities in Event Time

See figure 1 for definitions of pairs trading, event day and convergence. Panel A (B, C) plots the frequency of convergence within 5 (10, 20) days after divergence.

Panel A: Frequency of Convergence within 5 Days



Panel B: Frequency of Convergence within 10 Days



Panel C: Frequency of Convergence within 20 Days

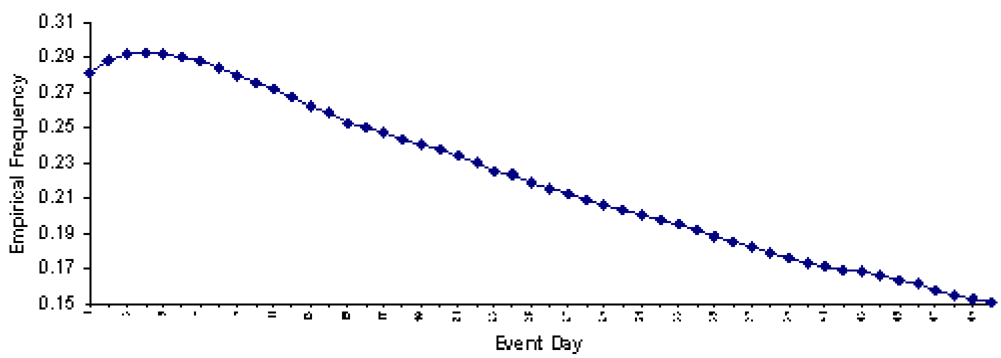


Figure 3. Distribution of Pairs Convergence

See figure 1 for definitions of pairs trading, event day and convergence. The bars of the figure plot the empirical distribution of days until convergence after a pair diverges. The blue line is the kernel density with a uniform kernel and a bandwidth chosen using Silverman's rule of thumb.

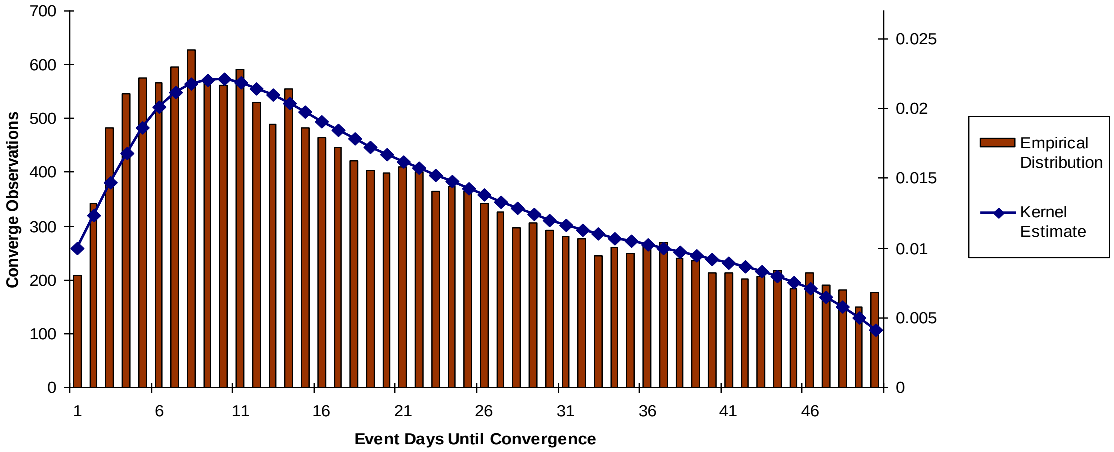


Figure 4. Distribution of the Number of Pairs Invested In Calendar Time – 10 Day Cream Skimming Strategy

Figure 4 illustrates the distribution of the number of total pairs and number of unique pairs invested in calendar time for the pairs trading strategy with ten day maximum holding horizon. See figure 1 for definitions of pairs trading, event day and convergence.

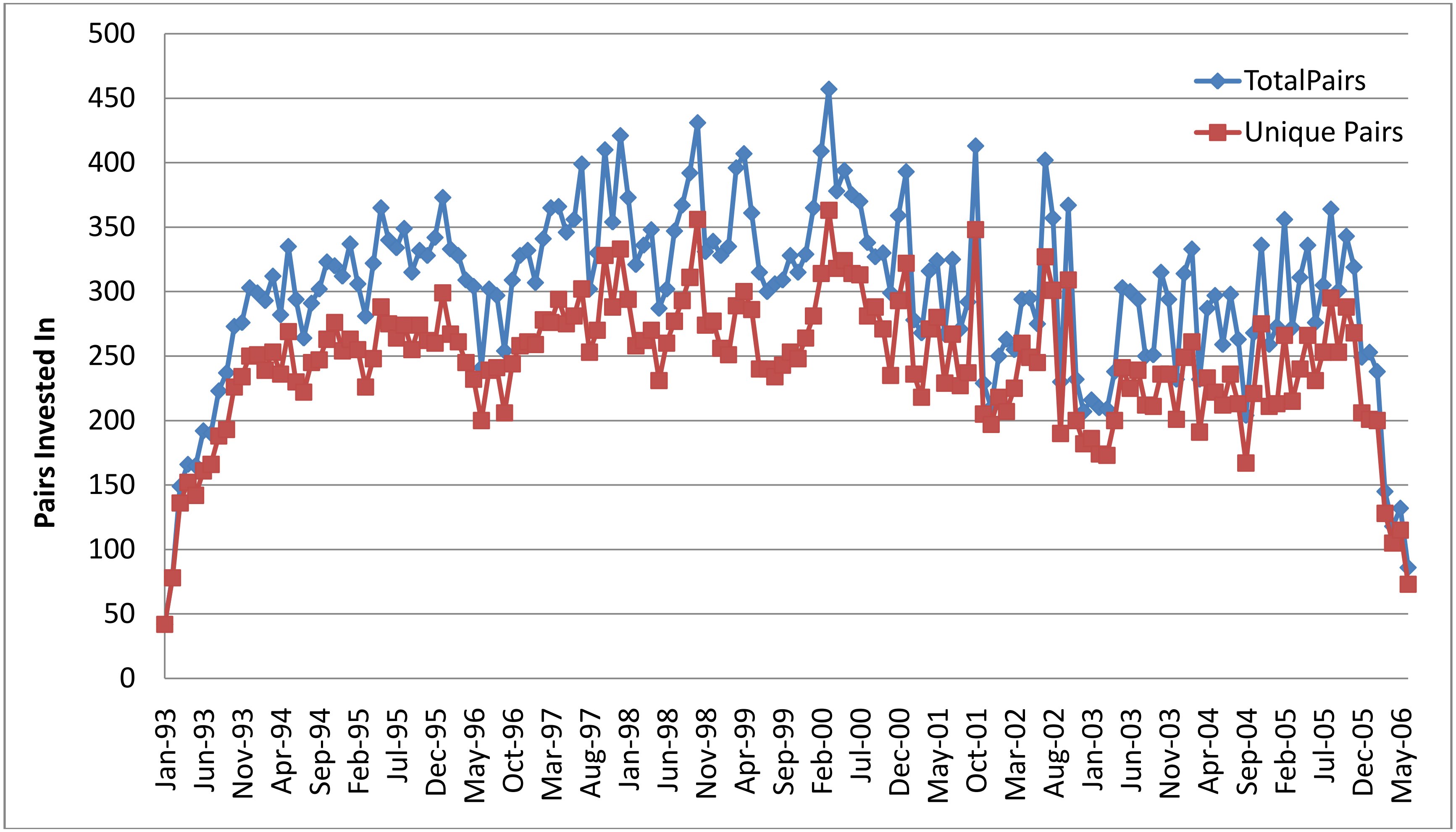


Figure 5. Distribution of the Number of Pairs Invested In Calendar Time – 6 m Strategy

Figure 5 illustrates the distribution of the number of total pairs and number of unique pairs invested in calendar time for the pairs trading strategy with six month maximum holding horizon. See figure 1 for definitions of pairs trading, event day and convergence.

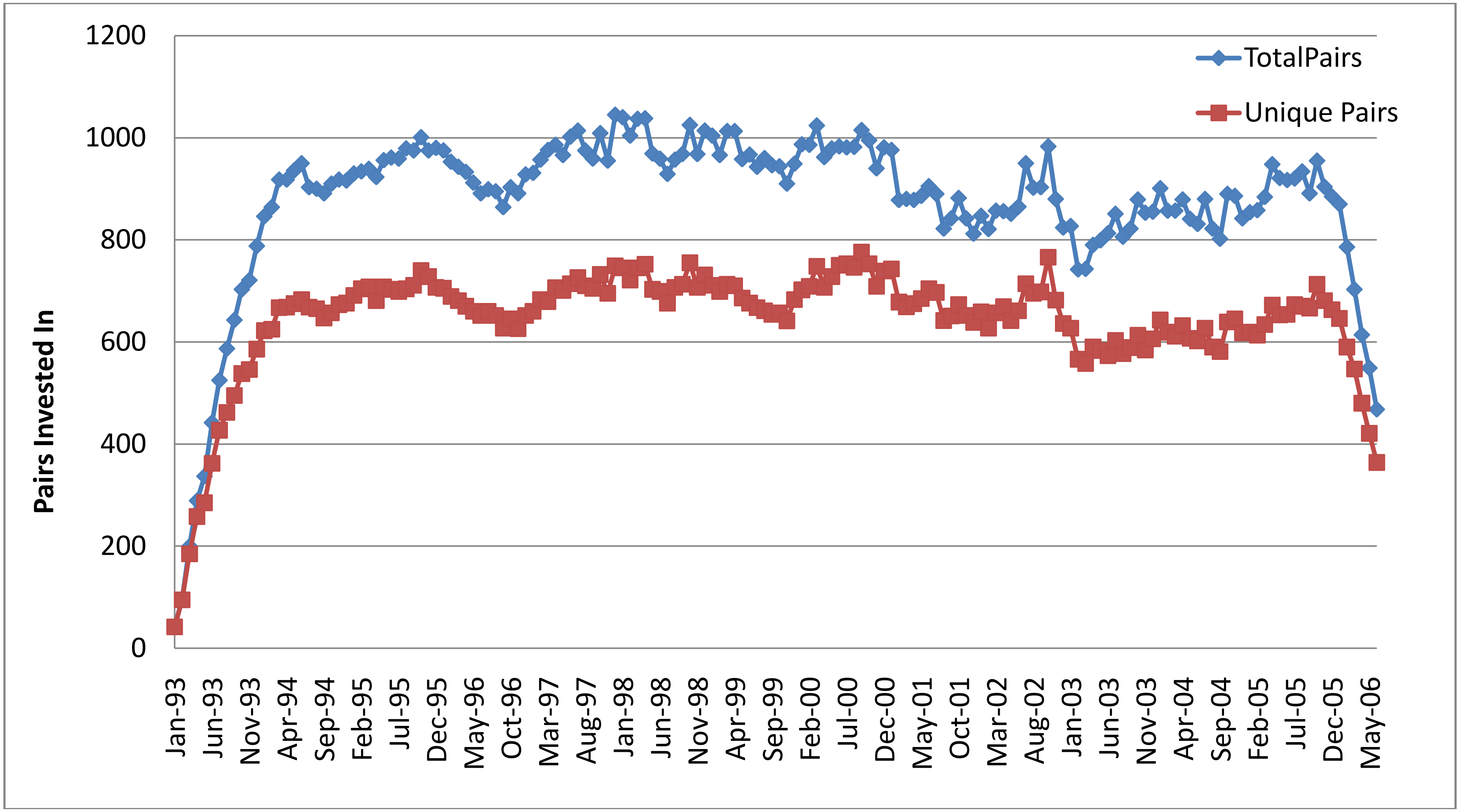


Figure 6. Distribution of the Number of Pairs Invested In Event Time – 10 d Cream Skimming & 6 m Strategies

Figure 6 illustrates the distribution of the number of total invested in event time for the pairs trading strategy with six month maximum holding horizon (upper line) and ten day maximum holding horizon (lower line).

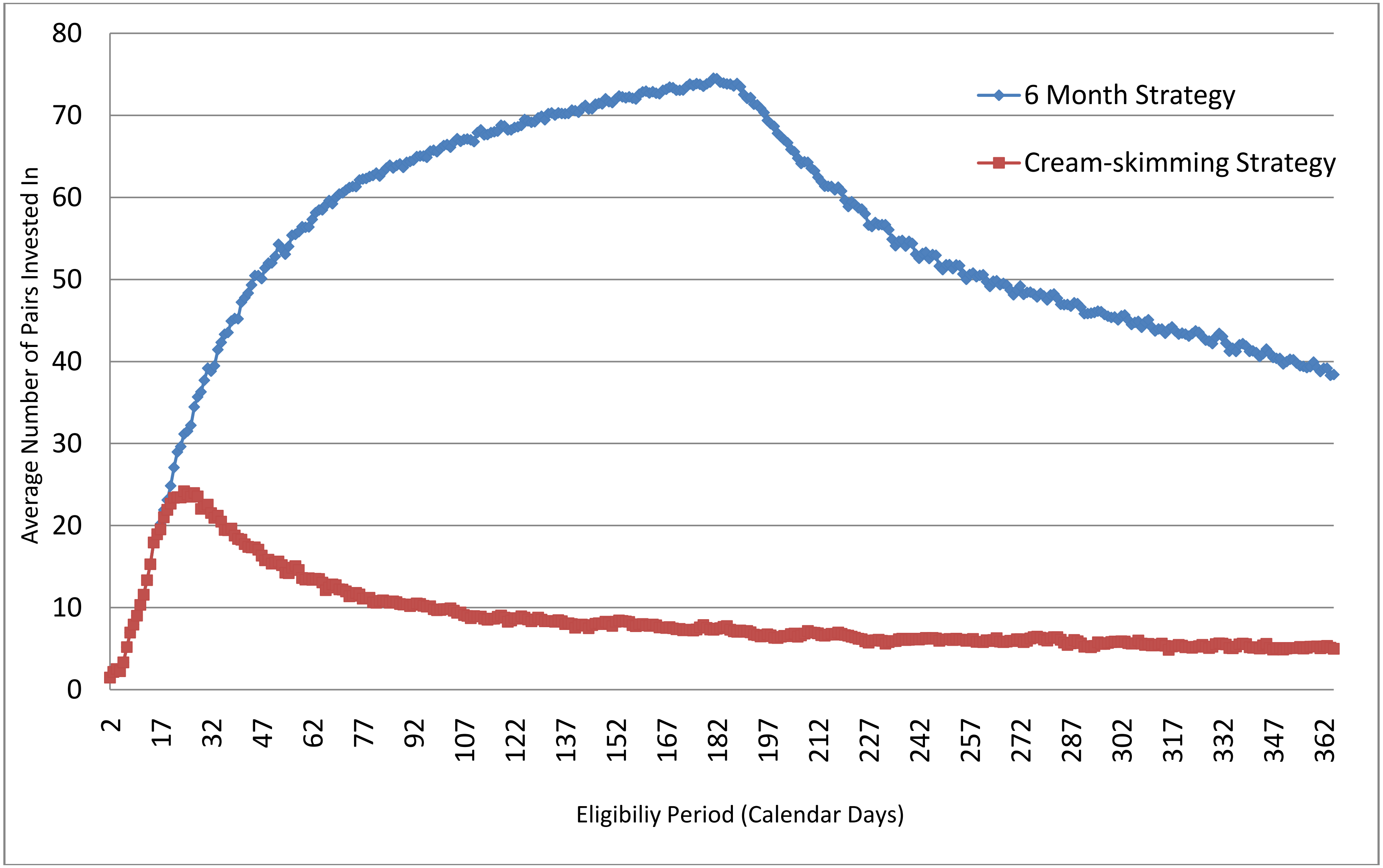


Figure 7. Frequency of Spreads Widening Events before Convergence

This figure plots the distribution (in percentage term) of spread-widening events among all opened pairs during the maximum six month holding horizon. The spread widening event is defined as the event such that the spreads on day (t) further widen compared with the maximum spread occuring in the window of [1, t-1], i.e., all prior trading day's maximum spread since the pair opened.

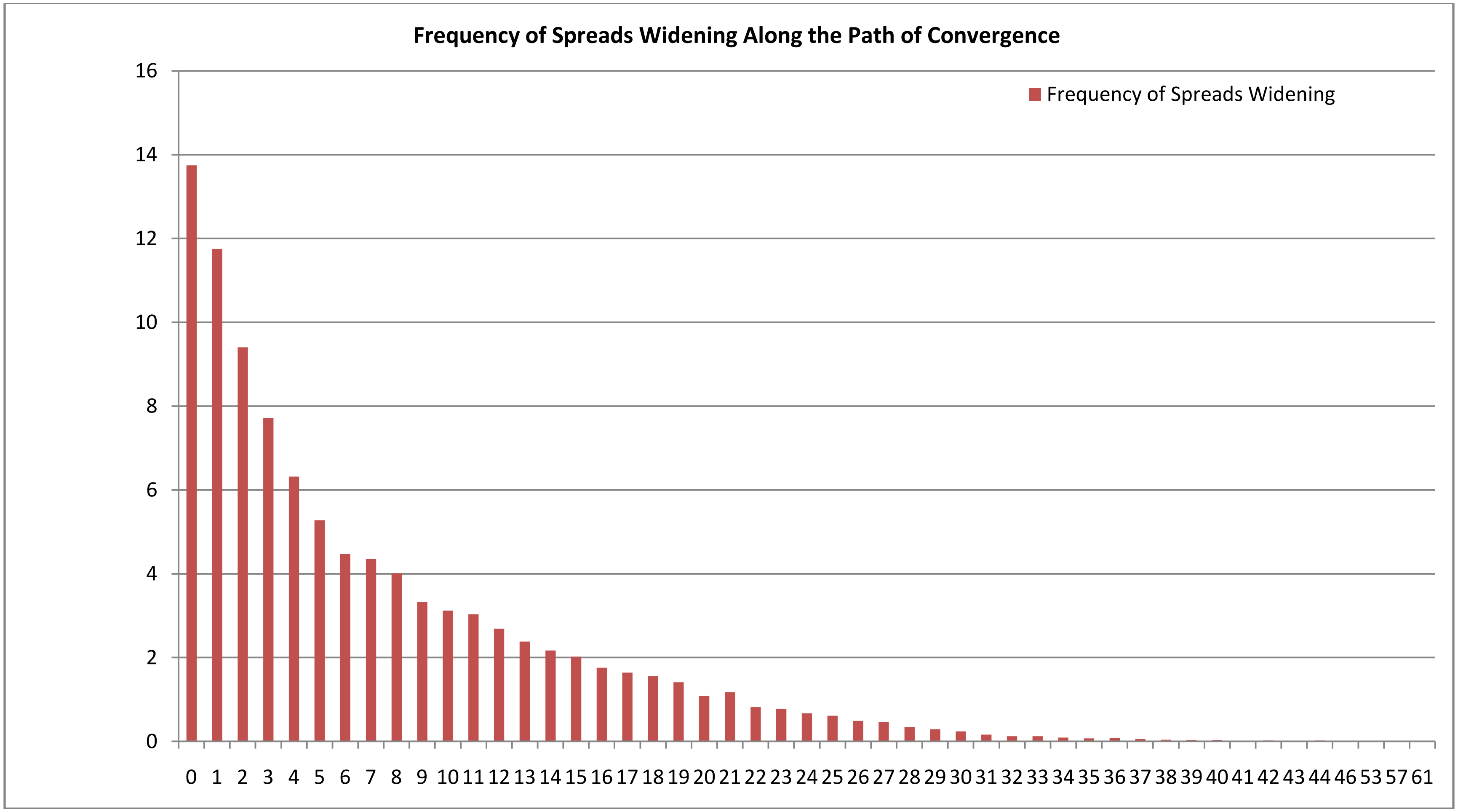


Table 1: Summary Statistics

Panel A reports a set of pairwise characteristics. The definition of the variables are provided in Section 4 of the text. Panel B reports the returns from pairs trading with 10-day and 6-month maximum holding horizon, sorted on the Fama-French 12-industry classification. Panel C reports the returns from pairs trading with 10-day and 6-month holding horizon, as well as the pairwise characteristics, sorted on whether the constituent stocks of the pairs trade on the NYSE, AMEX, NASDAQ (the same exchanges) or trading on different stock exchanges (mixed exchanges). To test the difference in returns or pairwise characteristics of pairs traded on the same exchange versus the mixed exchange, we use both asymptotic t-test and the Kolmogorov-Smirnov nonparametric test. In all cases, the equality of sample mean was rejected at 1% level.

Panel A: Summary Statistics of Variables

	Q1	Mean	Std	Median	Q3
Avg_PESPR	0.003	0.007	0.006	0.005	0.009
Avg_PESPR_Change	-0.001	0.000	0.006	0.000	0.001
Avg_Depth (in round lots)	7	19	43	11	21
Avg_Turn	0.319	0.559	0.594	0.453	0.644
Avg_dTurn_Change	-0.451	0.111	1.625	0.060	0.585
Avg_Ret_pst1mth	-0.029	0.011	0.075	0.010	0.051
Avg_Ret_pst12mth	-0.042	0.103	0.228	0.099	0.237
Avg_Ret_pst36mth	0.056	0.332	0.543	0.262	0.516
Avg_BM	0.512	0.737	0.331	0.691	0.902
Avg_MktCap (in millions)	847	6,537	15,163	2,276	5,862
Avg_SizeRank (percentiles)	45.0	62.8	22.6	65.0	82.5
Avg_Price	25.9	39.7	28.0	33.9	45.5
Avg_mRetVola	0.046	0.064	0.032	0.057	0.072
DIF_FF12_D3	0.166	0.487	0.426	0.366	0.694
Common_Holding_Ratio	0.220	0.349	0.166	0.350	0.471
Common_Analyst_Ratio	0.000	0.240	0.244	0.182	0.391
Pair_Holders	27	88	97	61	115
Pair_Analysts	0	4	5	2	6
Avg_EDF (x100)	0.05	0.26	0.63	0.10	0.24
Max_EDF (x 100)	0.06	0.39	1.08	0.15	0.34
minCumReturn126	-13.80%	-9.11%	10.65%	-5.91%	-1.50%
maxCumReturn126	4.62%	7.57%	5.06%	7.33%	10.11%
Time-till-convergence, unconditional	19	67	49	55	126
Time-till-convergence, conditional on converging in 10 days	4	6	3	6	8
Time-till-convergence, conditional on converging in 6 months	13	38	31	28	56

Panel B: Returns from pairs trading by Industry, 10-day and 6-month holding horizons

Industry Code	Industry Description	N	Percentage (%)	Return (10 day) Mean	Return (6 month) Mean	Return (10 day) Median	Return (6 month) Median
1	Consumer Nondurables	1,822	6.58	0.55%***	2.53%***	0.18%***	7.54%***
2	Consumer Durables	275	0.99	1.47%***	2.47%***	1.32%***	8.04%***
3	Manufacturing	3,867	13.96	0.68%***	1.99%***	0.55%***	7.75%***
4	Energy	1,016	3.67	0.33%**	0.56%	0.21%**	6.39%***
5	Chemicals and Allied Products	785	2.83	0.45%***	1.52%***	0.28%*	6.84%***
6	Business Equipment	401	1.45	2.43%***	4.32%***	1.27%***	8.65%***
7	Telecom	235	0.85	0.34%	0.94%*	0.44%***	1.27%***
8	Utilities	6,239	22.52	0.52%***	1.28%***	0.43%***	5.13%***
9	Whole Sales and Retails	699	2.52	-0.39%	0.20%	0.14%	7.04%***
10	Healthcare, Medical Equipment, and Drugs	50	0.18	3.44%***	6.87%***	2.87%***	9.73%***
11	Finance	12,294	44.38	1.13%***	2.83%***	0.85%***	7.26%***
12	Others, non-classified industries	20	0.07	1.23%**	3.13%***	0.96%*	3.53%***
All	All industries	27,703	100.00	0.82%***	2.17%***	0.58%***	6.76%***

Panel C: Returns from pairs trading by stock exchange, 10-day and 6-month holding horizons

Exchange	N	Percentage (%)	Return (10 day) Mean	Return (6 month) Mean	Return (10 day) Median	Return (6 month) Median
NYSE	16,158	58.37	0.48%***	1.47%***	0.40%***	6.22%***
Amex	121	0.44	1.88%***	2.49%***	1.21%***	2.32%***
NASDAQ	4,303	15.54	1.62%***	3.53%***	1.26%***	7.76%***
Same Exchange	20,582	74.35	0.72%***	1.91%***	0.53%***	6.48%***
Mixed Exchange	7,121	25.65	1.10%***	2.92%***	0.77%***	7.63%***
Difference, Mixed - Same			0.37%***	1.01%***		

Panel C, continued

Exchange	Avg_MktCap	Avg_SizeRank	DIF_FF12_D3	Common_Holding_Ratio	Common_Analyst_Ratio	Pair_Holders	Pair_Analysts
NYSE	3,901	75.0	0.334	0.396	0.279	94	4
Amex	3,481	58.1	0.397	0.246	0.530	34	10
NASDAQ	1,319	39.4	0.468	0.332	0.153	30	1
Same Exchange	7,380	65.8	0.457	0.377	0.280	102	5
Mixed Exchange	4,104	54.0	0.574	0.271	0.127	47	2
Difference, Mixed - Same	-3,276***	-11.8***	0.117***	-0.105***	-0.153***	-55***	-3***

Table 2. Distribution of Select Corporate Events around Divergence Dates

First four columns in Panel A report the distribution of select corporate events (quarterly earnings announcements, seasoned equity offerings, mergers and acquisitions, debt issuance) within [t-1, t] two-day event window leading up to the date of pair divergence. Zero stands for none of the constituent stocks of the pair experiences any corresponding corporate events, and one stands for at least one of the constituent stocks of the pair experience the corresponding corporate events. The last column in Panel A, "All Events", counts multiple events happening within [t-1, t] two-day event window leading up to the date of pair divergence as one. Panel B is similar to Panel A, but Panel B only consider the pairs where there is at least one piece of news coverage identified from Dow Jones News Wire (DJNW) news database on the date of divergence.

	Earnings Announcement	SEO	Mergers & Acquisitions	Debt Issuance	All Events
Panel A: Number and percentage of select events within two-day window around date of pair divergence					
0	26035	27673	27663	27554	25843
1	1668	30	40	149	1860
All	27703	27703	27703	27703	27703
0	93.98%	99.89%	99.86%	99.46%	93.29%
1	6.02%	0.11%	0.14%	0.54%	6.71%
All	100.00%	100.00%	100.00%	100.00%	100.00%
Panel B: Number and percentage of select events within two-day window around date of pair divergence, conditional on news coverage					
0	5743	6237	6223	6185	5661
1	503	9	23	61	585
All	6246	6246	6246	6246	6246
0	91.95%	99.86%	99.63%	99.02%	90.63%
1	8.05%	0.14%	0.37%	0.98%	9.37%
All	100.00%	100.00%	100.00%	100.00%	100.00%

Table 3. Profitability with Different Holding Periods

See figure 1 for definitions of event day, convergence and divergence. Table 1 reports the results of a regression where the dependent variable is monthly return from a pairs trading strategy and the independent variables are various factor returns. Table 1 reports the results of a regression where the dependent variable is monthly return from a pairs trading strategy with a 6-month maximum holding period (see Table 1) and the independent variables are standard factor returns taken from Ken French's website: the value weighted market excess return (Mkt – Rf), a portfolio of small stocks minus big stocks (SMB), a portfolio of high book-to-market minus low book-to-market stocks (HML), a portfolio of year-long winners minus year-long losers (MOM) and a portfolio of last month losers minus last month winners (ST_REV). "6-Month Maximum" means we close our position in a pair if it has not converged after 126 trading days. "10-Day Maximum" means we close our position in a pair if it has not converged after 10 trading days. Daily returns for the strategy are weighted by the cumulative returns of the component pairs. Daily returns are compounded to calculate monthly returns. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level. Factor series and details on construction of these factor series can be found from Ken French's website at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

	6-Month Maximum	10-Day Maximum
Intercept	0.007***	0.017***
	(0.001)	(0.001)
Mkt – Rf	-0.004	0.004
	(0.023)	(0.039)
SMB	0.043	0.013
	(0.027)	(0.040)
HML	-0.005	-0.016
	(0.032)	(0.053)
MOM	-0.082***	-0.056**
	(0.017)	(0.027)
St_Rev	0.045**	0.055
	(0.021)	(0.035)
Observations	162	162
R-Square	0.2992	0.07423

Table 4. Size Sorted Profitability with Differential Liquidity

See figure 1 for definitions of event day, convergence and divergence and Table 2 for factor definitions. Portfolios are first sorted into above (big) and below (small) median size portfolios and then sorted by terciles based on the proportional effective spread (PESPR) during the estimation period or the change in PESPR during the 5 days before divergence. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

Panel A: Size Sorted Profitability with Differential Liquidity, maximum 10 day holding period.

	Average PESPR (Big Firms)				Average PESPR Change (Big Firms)			
	Liquid		Illiquid	Illiquid- Liquid	Small Change		Large Change	Large Change– Small Change
Intercept	0.009***	0.013***	0.015***	0.007*	0.011***	0.010***	0.014***	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)
Mkt – Rf	0.107*	0.089	0.111	0.005	0.137**	0.158**	-0.010	-0.146
	(0.055)	(0.057)	(0.101)	(0.108)	(0.069)	(0.062)	(0.074)	(0.102)
SMB	-0.008	0.010	0.040	0.049	0.133*	-0.053	0.018	-0.115
	(0.058)	(0.065)	(0.067)	(0.084)	(0.068)	(0.063)	(0.073)	(0.070)
HML	0.164**	0.026	0.169	0.005	0.196**	0.160**	-0.047	-0.243**
	(0.071)	(0.080)	(0.109)	(0.122)	(0.087)	(0.080)	(0.088)	(0.107)
MOM	-0.079**	-0.100***	-0.052	0.027	-0.103*	-0.045	-0.113***	-0.010
	(0.037)	(0.037)	(0.060)	(0.068)	(0.053)	(0.047)	(0.033)	(0.061)
St_Rev	0.020	0.064	0.119	0.099	0.104	-0.022	0.140***	0.036
	(0.054)	(0.052)	(0.074)	(0.093)	(0.073)	(0.065)	(0.040)	(0.087)
Observations	156	156	156	156	156	156	156	156
R-Square	0.1007	0.1	0.06373	0.01403	0.1554	0.08707	0.1082	0.03839

	Average PESPR (Small Firms)				Average PESPR Change (Small Firms)			
	Liquid		Illiquid	Illiquid- Liquid	Small Change		Large Change	Large Change– Small Change
Intercept	0.016***	0.026***	0.024***	0.008**	0.018***	0.020***	0.029***	0.011***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Mkt – Rf	-0.032	0.025	-0.103	-0.072	-0.090	0.072	-0.063	0.027
	(0.078)	(0.098)	(0.091)	(0.115)	(0.073)	(0.098)	(0.097)	(0.112)
SMB	0.058	0.019	0.124	0.065	0.069	0.053	0.110	0.041
	(0.058)	(0.100)	(0.083)	(0.079)	(0.090)	(0.076)	(0.079)	(0.091)
HML	-0.051	-0.030	-0.088	-0.037	-0.021	-0.001	-0.158	-0.136
	(0.082)	(0.124)	(0.121)	(0.112)	(0.096)	(0.120)	(0.110)	(0.120)
MOM	0.079*	-0.030	-0.024	-0.102**	-0.031	0.057	0.030	0.061
	(0.045)	(0.062)	(0.066)	(0.049)	(0.058)	(0.056)	(0.055)	(0.079)
St_Rev	0.054	0.011	0.054	0.000	0.111	0.010	0.020	-0.091
	(0.054)	(0.097)	(0.082)	(0.058)	(0.088)	(0.069)	(0.080)	(0.094)
Observations	156	156	156	156	156	156	156	156
R-Square	0.04523	0.006716	0.03012	0.02178	0.03748	0.02389	0.04433	0.03867

Pane B: Size Sorted Profitability with Differential Liquidity, maximum 6 month holding period

	Average PESPR (Big Firms)				Average PESPR Change (Big Firms)			
	Liquid		Illiquid	Illiquid- Liquid	Small Change		Large Change	Large Change– Small Change
Intercept	0.005***	0.005***	0.007***	0.002	0.006***	0.004***	0.005***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mkt – Rf	0.074***	0.063**	-0.002	-0.076**	0.023	0.098***	0.005	-0.023
	(0.028)	(0.030)	(0.034)	(0.037)	(0.027)	(0.026)	(0.031)	(0.034)
SMB	0.030	0.041	0.043	0.013	0.043	0.021	0.042	-0.005
	(0.025)	(0.030)	(0.041)	(0.047)	(0.028)	(0.029)	(0.027)	(0.025)
HML	0.103***	0.084*	0.018	-0.085	0.047	0.095***	0.049	0.000
	(0.032)	(0.044)	(0.053)	(0.055)	(0.037)	(0.036)	(0.040)	(0.042)
MOM	-0.115***	-0.078***	-0.076***	0.039	-0.089***	-0.064***	-0.128***	-0.044*
	(0.020)	(0.025)	(0.028)	(0.031)	(0.022)	(0.020)	(0.026)	(0.023)
St_Rev	-0.007	0.046	0.120***	0.128***	0.066**	0.019	0.054**	-0.014
	(0.025)	(0.030)	(0.034)	(0.040)	(0.029)	(0.028)	(0.027)	(0.031)
Observations	162	161	161	161	160	161	162	161
R-Square	0.2691	0.1983	0.2009	0.1169	0.2141	0.1982	0.2798	0.02719
	Average PESPR (Small Firms)				**Average PESPR Change (Small Firms)**			
	Liquid		Illiquid	Illiquid- Liquid	Small Change		Large Change	Large Change– Small Change
Intercept	0.006***	0.010***	0.011***	0.005***	0.010***	0.008***	0.009***	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Mkt – Rf	0.035	-0.042	-0.057	-0.092	-0.017	-0.009	-0.024	-0.007
	(0.048)	(0.042)	(0.041)	(0.056)	(0.043)	(0.050)	(0.041)	(0.042)
SMB	0.059*	0.063	0.057	-0.001	0.060	0.056	0.102**	0.041
	(0.033)	(0.039)	(0.049)	(0.040)	(0.045)	(0.043)	(0.043)	(0.038)
HML	0.015	-0.046	-0.036	-0.051	0.000	-0.035	-0.012	-0.011
	(0.050)	(0.050)	(0.061)	(0.054)	(0.058)	(0.063)	(0.054)	(0.055)
MOM	-0.027	-0.055**	-0.065**	-0.038	-0.055*	-0.030	-0.051**	0.004
	(0.024)	(0.026)	(0.033)	(0.030)	(0.029)	(0.032)	(0.022)	(0.030)
St_Rev	-0.003	0.057	0.080*	0.083***	0.066*	0.006	0.065**	-0.001
	(0.025)	(0.035)	(0.041)	(0.032)	(0.034)	(0.035)	(0.028)	(0.036)
Observations	161	162	161	161	162	161	161	161
R-Square	0.04736	0.1033	0.1064	0.06971	0.09001	0.04404	0.1231	0.00974

Table 5. Profitability of Pairs Trading Strategy with News and No News

See figure 1 for definitions of event day, convergence and divergence and Table 2 for factor descriptions. Big (small) stocks are those above (below) the median in our sample in the month of divergence. "No Abnormal Return" means neither stock in the pair had an absolute excess return on the day of divergence that was greater than two historical standard deviations. Standard deviation is calculated over the prior 21 trading days (one month). "News" means that at least one stock in the pair had an abnormal return on the day of divergence and had a story in the Dow Jones News Service. "No News" means that at least one stock in the pair had an abnormal return on the day of divergence but that (or those) stocks did not have a story in the Dow Jones News Service. Daily returns for the strategy are weighted by the cumulative returns of the component pairs. Daily returns are compounded to calculate monthly returns. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	Big Stocks				Small Stocks			
	No Abnormal Return	News	No News	No News - News	No Abnormal Return	News	No News	No News - News
Intercept	0.004***	0.004***	0.007***	0.003***	0.007***	0.007***	0.010***	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Mkt – Rf	0.056**	0.063**	0.032	-0.032	-0.057	-0.019	-0.067	-0.048
	(0.025)	(0.032)	(0.030)	(0.044)	(0.037)	(0.055)	(0.048)	(0.053)
SMB	0.004	0.064*	0.033	-0.032	0.067	0.194***	0.016	-0.179***
	(0.023)	(0.037)	(0.030)	(0.041)	(0.052)	(0.050)	(0.043)	(0.038)
HML	0.056*	0.085*	0.059	-0.025	-0.063	-0.061	-0.095	-0.034
	(0.033)	(0.049)	(0.048)	(0.073)	(0.057)	(0.064)	(0.076)	(0.069)
MOM	-0.085***	-0.124***	-0.063***	0.061**	-0.050	-0.034	-0.084***	-0.050
	(0.024)	(0.024)	(0.022)	(0.029)	(0.032)	(0.036)	(0.027)	(0.031)
St_Rev	-0.012	0.042	0.052*	0.009	0.074*	-0.040	0.062**	0.101**
	(0.028)	(0.031)	(0.027)	(0.038)	(0.041)	(0.047)	(0.026)	(0.042)
Observations	162	161	162	161	162	161	162	161
R-Square	0.1942	0.2803	0.1632	0.05001	0.1296	0.1668	0.1225	0.1611

Table 6. Profitability of Pairs Trading Strategy with Industry Information Diffusion Rates

See Equation 5.5 in the text for the definition of industry diffusion rate. Portfolios are first sorted into above (big) and below (small) median size portfolios and then sorted by terciles based on the difference in industry diffusion rates. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

Panel A: Difference in Industry Diffusion Rate (All Firms)

	Small		Big	Big - Small
Intercept	0.005***	0.006***	0.009***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Mkt – Rf	0.031	0.017	0.013	-0.018
	(0.023)	(0.024)	(0.040)	(0.033)
SMB	0.047**	0.055**	0.033	-0.014
	(0.022)	(0.026)	(0.036)	(0.025)
HML	0.034	0.023	0.039	0.004
	(0.030)	(0.031)	(0.051)	(0.037)
MOM	-0.063***	-0.098***	-0.041	0.022
	(0.019)	(0.017)	(0.026)	(0.022)
St_Rev	0.038	0.027	0.071***	0.034*
	(0.025)	(0.020)	(0.025)	(0.019)
Observations	162	161	161	161
R-Square	0.1947	0.3055	0.08476	0.0182

Panel B: Difference in Industry Diffusion Rate (Big Firms)

	Small		Big	Big - Small
Intercept	0.003***	0.005***	0.008***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Mkt – Rf	0.067***	0.022	0.026	-0.041
	(0.024)	(0.026)	(0.036)	(0.033)
SMB	0.031	0.022	0.031	0.001
	(0.025)	(0.026)	(0.035)	(0.028)
HML	0.067**	0.064*	0.048	-0.020
	(0.031)	(0.035)	(0.046)	(0.042)
MOM	-0.082***	-0.124***	-0.077***	0.005
	(0.019)	(0.025)	(0.026)	(0.022)
St_Rev	0.037	0.035	0.067**	0.030
Observations	162	161	161	161
R-Square	0.2486	0.2245	0.1568	0.01802

Panel C: Difference in Industry Diffusion Rate (Small Firms)

	Small		Big	Big - Small
Intercept	0.009***	0.008***	0.010***	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Mkt – Rf	-0.022	0.001	-0.020	0.002
	(0.041)	(0.042)	(0.042)	(0.040)
SMB	0.055	0.080**	0.038	-0.017
	(0.037)	(0.038)	(0.046)	(0.030)
HML	-0.027	-0.018	-0.006	0.022
	(0.050)	(0.060)	(0.059)	(0.042)
MOM	-0.027	-0.075***	-0.028	-0.001
	(0.028)	(0.025)	(0.029)	(0.029)
St_Rev	0.055	0.026	0.048	-0.007
Observations	162	161	161	161
R-Square	0.07094	0.1158	0.03115	0.006289

Table 7. Profitability with Differential Coverage/Holdings of Pairs

See figure 1 for definitions of event day, convergence and divergence and Table 2 for factor definitions. Each quarter, we examine whether the analysts from the same brokerage house actively follow both stocks from the pair by issuing at least one earning forecast (regardless of forecast horizon). If the brokerage covers both stocks from the pair, we call that brokerage "Pair Analyst". The number of unique brokerage house covering both stocks from the pair is called the total number of Analysts that Cover Pairs. Each quarter, we examine whether a financial institution holds the both stocks in the pair within its portfolio. If that institution holds both stocks from the pair, we call that institution a "Pair Holder". Among all financial institutions filing form S34, we count how many unique institutions holding both stocks within a pair. The number of unique institutions holding both stocks from the pair is called the total number of Institutions that Hold Pairs. Portfolios are sorted on terciles based on the day of divergence. Tercile cutoffs are calculated monthly. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	Institutions that Hold Pairs				Analysts that Cover Pairs			
	Few		Many	Many - Few	Few		Many	Many - Few
Intercept	0.009***	0.008***	0.004***	-0.004***	0.009***	0.008***	0.004***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mkt – Rf	-0.040	0.054	0.062***	0.102***	-0.052	0.022	0.084***	0.136***
	(0.036)	(0.035)	(0.021)	(0.037)	(0.040)	(0.035)	(0.022)	(0.041)
SMB	0.044	0.066*	0.038*	-0.006	0.043	0.085**	0.028	-0.015
	(0.036)	(0.037)	(0.022)	(0.040)	(0.039)	(0.039)	(0.019)	(0.042)
HML	-0.024	0.044	0.083***	0.107**	-0.044	0.054	0.092***	0.136**
	(0.046)	(0.056)	(0.028)	(0.052)	(0.057)	(0.049)	(0.030)	(0.059)
MOM	-0.055***	-0.048*	-0.106***	-0.051**	-0.044*	-0.051**	-0.101***	-0.058*
	(0.019)	(0.029)	(0.017)	(0.022)	(0.026)	(0.024)	(0.019)	(0.031)
St_Rev	0.073***	0.027	0.029	-0.044	0.085***	0.039	0.024	-0.061*
	(0.026)	(0.030)	(0.023)	(0.031)	(0.031)	(0.028)	(0.022)	(0.033)
Observations	162	161	162	162	162	161	162	162
R-Square	0.1628	0.1108	0.3294	0.1247	0.1135	0.1527	0.3258	0.1419

Table 8. Size Sorted Profitability with Differential Coverage/Holdings of Pairs

See Table 7 for definitions. Portfolios are first sorted into above (big) and below (small) median size portfolios and then sorted by terciles as in Table 7. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	Institutions that Hold Pairs (Big Firms)				Institutions that Hold Pairs (Small Firms)			
	Few		Many	Many - Few	Few		Many	Many - Few
Intercept	0.007***	0.004***	0.005***	-0.002*	0.010***	0.009***	0.008***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Mkt – Rf	-0.001	0.072**	0.066**	0.066	-0.042	-0.021	0.021	0.063
	(0.041)	(0.028)	(0.026)	(0.043)	(0.042)	(0.052)	(0.037)	(0.047)
SMB	0.063	0.018	0.032	-0.030	-0.001	0.104**	0.098**	0.098**
	(0.044)	(0.025)	(0.027)	(0.048)	(0.039)	(0.046)	(0.047)	(0.042)
HML	0.053	0.083**	0.076**	0.023	-0.058	0.006	0.014	0.072
	(0.062)	(0.040)	(0.034)	(0.069)	(0.056)	(0.073)	(0.047)	(0.057)
MOM	-0.051	-0.073***	-0.145***	-0.094**	-0.026	-0.064**	-0.041	-0.015
	(0.031)	(0.022)	(0.024)	(0.037)	(0.029)	(0.027)	(0.027)	(0.033)
St_Rev	0.085**	0.064***	0.004	-0.080*	0.067**	0.057*	0.006	-0.062*
	(0.036)	(0.022)	(0.033)	(0.043)	(0.032)	(0.034)	(0.035)	(0.033)
Observations	161	162	161	161	162	161	161	161
R-Square	0.118	0.2071	0.3265	0.107	0.05143	0.0987	0.1042	0.057

	Analysts that Cover Pairs (Big Firms)				Analysts that Cover Pairs (Small Firms)			
	Few		Many	Many - Few	Few		Many	Many - Few
Intercept	0.007***	0.004***	0.004***	-0.003**	0.008***	0.011***	0.009***	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Mkt – Rf	0.009	0.084***	0.073***	0.064*	-0.036	0.009	-0.011	0.025
	(0.036)	(0.029)	(0.024)	(0.038)	(0.039)	(0.079)	(0.052)	(0.040)
SMB	0.033	0.052	0.036	0.004	0.048	-0.046	0.081*	0.033
	(0.037)	(0.032)	(0.025)	(0.042)	(0.041)	(0.091)	(0.042)	(0.035)
HML	0.018	0.114***	0.107***	0.088	-0.034	0.100	-0.057	-0.023
	(0.057)	(0.042)	(0.032)	(0.056)	(0.052)	(0.099)	(0.064)	(0.051)
MOM	-0.053*	-0.114***	-0.107***	-0.054*	-0.057**	-0.042	-0.027	0.031
	(0.030)	(0.021)	(0.019)	(0.030)	(0.022)	(0.051)	(0.036)	(0.032)
St_Rev	0.083***	0.058**	-0.001	-0.084**	0.058**	0.039	0.029	-0.029
	(0.031)	(0.029)	(0.028)	(0.033)	(0.028)	(0.051)	(0.039)	(0.033)
Observations	161	162	161	161	162	152	161	161
R-Square	0.1184	0.3114	0.2612	0.08887	0.1022	0.04431	0.07114	0.03366

Table 9: Factor Regression of Monthly Pairs Trading Strategy Returns

This table reports the factor regression of pairs trading strategy monthly returns. In Panel A, pairs are closed out by the tenth day; and in Panel B, pairs are closed out by the end of the sixth month. MKTRF, SMB, HML, MOM and LIQ are the market factor, small-minus-big factor, high-minus-low factor, momentum factor, and liquidity factor. In regressions (1) and (2) of each panel, the liquidity factors are the value-weighted version and equally-weighted version of Pastor-Stambaugh liquidity factor (Pastor and Stambaugh, 2003), respectively; in regressions (3) and (4) of each panel, the liquidity factors are the fixed-cost and variable-cost components of the spreads liquidity risk factors constructed by Sadka (2006). Due to availability of liquidity risk factors, the sample period for regressions (1) and (2) is from January, 1993 to December, 2004; and the sample period for regressions (3) and (4) is from January, 1993 to December, 2005.

	Panel A: 10-day strategy monthly return				Panel B: 6-month strategy monthly return			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Intercept	0.019***	0.019***	0.018***	0.018***	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
MKTRF	0.018	0.006	0.012	0.020	0.005	-0.001	0.002	0.004
	(0.038)	(0.040)	(0.039)	(0.042)	(0.024)	(0.024)	(0.023)	(0.025)
SMB	0.021	0.017	0.022	0.025	0.050*	0.048	0.051*	0.051
	(0.039)	(0.042)	(0.042)	(0.041)	(0.030)	(0.032)	(0.031)	(0.031)
HML	-0.033	-0.034	-0.009	0.001	-0.013	-0.013	0.002	0.002
	(0.054)	(0.058)	(0.055)	(0.060)	(0.033)	(0.034)	(0.033)	(0.038)
MOM	-0.037	-0.046	-0.067**	-0.056**	-0.077***	-0.082***	-0.092***	-0.089***
	(0.029)	(0.037)	(0.028)	(0.028)	(0.019)	(0.019)	(0.016)	(0.018)
LIQ	-0.049**	-0.032	-0.536	-0.797*	-0.026	-0.017	-0.389	-0.288
	(0.024)	(0.036)	(0.887)	(0.443)	(0.017)	(0.020)	(0.665)	(0.276)
Observations	144	144	156	156	144	144	156	156
R-Squared	0.077	0.06088	0.06117	0.09048	0.2927	0.2807	0.2831	0.2899

Table 10: Macroeconomic and Liquidity Risk Exposures

This table reports results from time-series regressions of the monthly returns of pairs trading strategies with holding horizon of ten days (Panel A) and six months (Panel B), on various measures of macroeconomic and liquidity risks. The macroeconomic variables are the long-run consumption growth rates, default spreads measured as the spreads between Moody's BAA and Moody's AAA corporate bond rates. The macro liquidity risk proxy variables include the US TED spread and AAA/T-Bill spreads. US TED Spreads is the average daily spread between 3-month LIBOR rates and 3-month Treasury Bill rates in the US over the month. AAA/T-Bill spreads is the average daily spread between Moody's AAA corporate bond rates and 3-month Treasury Bill rates in the US over the month. The sampling period of regression (1) in Panel A and Panel B is from January, 1993 to June, 2006; and the sampling period of regressions (2) and (3) in Panel A and Panel B is from January, 1993 to March, 2005 due to lack of data on long-run consumption growth.

	Panel A: Convergence Strategy in 10-day			Panel B: Convergence Strategy in 6-month		
	(1)	(2)	(3)	(1)	(2)	(3)
US TED Spreads	2.278***	2.279***	2.286***	0.980**	1.000**	1.001**
	(0.707)	(0.715)	(0.721)	(0.424)	(0.430)	(0.430)
BAA/AAA Spreads	0.387	0.426	0.551	-0.012	0.074	0.077
	(0.784)	(0.845)	(0.884)	(0.498)	(0.536)	(0.525)
AAA/T-Bill Spreads	0.541	0.522	0.514	0.395*	0.359	0.359
	(0.354)	(0.375)	(0.376)	(0.212)	(0.223)	(0.222)
Long-run Consumption Growth			-0.049			-0.001
			(0.083)			(0.051)
Observations	162	147	147	162	147	147
R-Square	10.3%	10.3%	10.5%	6.7%	6.6%	6.6%

Table 11. Time-series Cross-Sectional Regressions of Pairs Returns on Pairs Characteristics

This table reports the time-series cross-sectional regression of individual pair's profits (dependent variable) on the pair's characteristics (independent variables). In Panel A and Panel C, the profits are computed from a strategy of maximum holding horizon of 10 days. In Panel B and Panel D, the profits are computed from a strategy of maximum holding horizon of 6 months. Panel A (Panel C) and Panel B (Panel D) differ in terms of independent variables. Compared with Panel A (Panel C), Panel B (Panel D) include industry information diffusion measure (DIF_FF12_D3) as well as its interaction with Size, Liquidity, Common Analyst Coverage and Common Institutional Holding variables. Several pairs characteristics control variables, including Avg_Ret_pst1mth, Avg_Ret_pst12mth, Avg_Ret_pst36mth, Avg_BM, Log_Avg_MktCap, Avg_mRetVola, though included in the regressions of Panel C and Panel D, are not reported to preserve brevity. Avg_PESPR is the pair's average proportional effective spreads, measured during the pair formation period. Avg_PESPR_Change is the change of the average of the pair's proportional effective spreads, measured in the previous five days leading to the event day minus the pair's average proportional effective spreads, measured during the pair formation period. Avg_Turn is the pair's average daily turnover ratio, measured during the pair formation period. Avg_dTurn_Change is the change of the average of the pair's daily turnover ratio, measured in the previous five days leading to the event day; minus the pair's average daily turnover ratio, measured during the pair formation period. News is defined in Table 5 and Coverage is "No News" from Table 2. Avg_Ret_pst1mth is the pair's average cumulative returns over the one month prior to the event month (event month is the month when the event date occurs). Avg_Ret_pst12mth is the pair's average cumulative return over the eleven months prior to the second month to the event month. Avg_Ret_pst36mth is the pair's average cumulative return over the 24 months prior to the twelve month to the event month. Avg_BM is the pair's average book to market equity ratios measured using the most recently available book equity value, and the market equity values during the month ending at the beginning of the previous month. Log_Avg_MktCap is the logarithm of market capitalization of firms in billion dollars using last available market capitalization t during the estimation period. Avg_mRetVola is the average of the pair's monthly return volatilities during estimation period. Common_Holding_Ratio is the number of institutions holding both stocks in the pair during the quarter prior to the event quarter (the quarter the event date occurs), divided by the maximum number of institutions holding stock one or stock two of the pair during the same quarter. If the number of institutions holding two stocks of the pair is less than fifty, the Common_Holding indicator variable takes the value of one; and zero otherwise. Common_Coverage_Ratio is the number of brokerage houses (as identified by the brokerage code in I/B/E/S), divided by the maximum number of brokerage houses covering stock one or stock two of the pair during the same quarter. If the number of brokerage houses covering two stocks of the pair is less than or equal to two, the Common_Coverage indicator variable takes the value of one; and zero otherwise. Size_Rank takes the value of one if the average size percentile of the pair is below 50-th, and zero otherwise. In regressions 3 of both Panel A and Panel B, the common institutional holding and common analyst coverage variables are continuous variables. In regressions 4 and 5 of both Panel A and Panel B, the common institutional holding and common analyst coverage variables are binary dummy variables, which take the value of one if the value of the variable is below sample median, and zero otherwise. All regressions compute the clustered standard errors, where the cluster is defined by year, month and industry. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	Panel A: Returns from Convergence Strategy in 10 days				Panel B: Returns from Convergence Strategy in 6 months			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Intercept	0.023***	0.024***	0.008*	0.008*	0.046***	0.049***	0.016	0.016
	(0.004)	(0.004)	(0.005)	(0.005)	(0.009)	(0.010)	(0.011)	(0.012)
Avg_PESPR	0.089	0.069	0.043	0.042	0.447**	0.387**	0.360**	0.361**
	(0.091)	(0.090)	(0.088)	(0.089)	(0.176)	(0.175)	(0.175)	(0.177)
Avg_PESPR_Change	0.176***	0.173***	0.170**	0.170**	-0.068	-0.076	-0.075	-0.074
	(0.066)	(0.066)	(0.066)	(0.066)	(0.148)	(0.148)	(0.148)	(0.148)
Avg_Turn	-0.001	-0.001	0.000	0.000	-0.006**	-0.005**	-0.004*	-0.004*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Avg_dTurn_Change	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
News	-0.004**	-0.004**	-0.004**	-0.004**	-0.013***	-0.012***	-0.012**	-0.012**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
Coverage	0.000	0.000	0.000	0.000	-0.003	-0.003	-0.003	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Avg_Ret_pst1mth	-0.006	-0.007	-0.007	-0.007	-0.049***	-0.049***	-0.050***	-0.050***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.018)	(0.017)	(0.017)	(0.017)
Avg_Ret_pst12mth	0.004	0.004	0.003	0.003	0.004	0.004	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Avg_Ret_pst36mth	0.001	0.001	0.001	0.001	0.004**	0.004**	0.004**	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Avg_BM	-0.004***	-0.004***	-0.004**	-0.004**	-0.010***	-0.010***	-0.009***	-0.009***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Log_Avg_MktCap	-0.002***	-0.002***	-0.001	-0.001	-0.004***	-0.003***	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Avg_mRetVola	0.081***	0.079***	0.074***	0.074***	0.182***	0.176***	0.171***	0.171***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.052)	(0.052)	(0.052)	(0.052)
Common_Holding		-0.006**	0.007***	0.007***		-0.016**	0.008***	0.008***
		(0.003)	(0.001)	(0.001)		(0.007)	(0.003)	(0.003)
Common_Coverage		0.000	0.000	0.000		-0.004	0.006**	0.006**
		(0.002)	(0.001)	(0.001)		(0.005)	(0.002)	(0.002)
Size_Rank				0.000				0.000
				(0.001)				(0.003)
Observations	27703	27703	27703	27703	27703	27703	27703	27703
Clusters	1409	1409	1409	1409	1409	1409	1409	1409
R-Squared	0.60%	0.63%	0.76%	0.76%	0.73%	0.78%	0.84%	0.84%

	Panel C: Returns from Convergence Strategy in 10 days					Panel D: Returns from Convergence Strategy in 6 months				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	0.021***	0.021***	0.016***	0.011**	0.016***	0.038***	0.039***	0.024**	0.029***	0.034***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.010)	(0.009)	(0.010)	(0.010)	(0.011)
Avg_PESPR	0.088	0.025	0.074	0.051	0.069	0.441**	0.200	0.399**	0.406**	0.424**
	(0.091)	(0.101)	(0.090)	(0.088)	(0.090)	(0.176)	(0.192)	(0.176)	(0.177)	(0.179)
Avg_PESPR_Change	0.175***	0.174***	0.174***	0.170***	0.173***	-0.071	-0.076	-0.073	-0.075	-0.073
	(0.066)	(0.066)	(0.066)	(0.066)	(0.066)	(0.148)	(0.149)	(0.148)	(0.148)	(0.148)
News	-0.004**	-0.004**	-0.004**	-0.004**	-0.004**	-0.013***	-0.013***	-0.013***	-0.013***	-0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Coverage	0.000	0.000	0.000	0.000	0.000	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
DIF_FF12_D3	0.003***	0.002**	0.002**	0.002*	0.002**	0.010***	0.008***	0.009***	0.009***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
DIF_FF12_D3 X Liquidity		0.003					0.011***			
		(0.002)					(0.004)			
DIF_FF12_D3 X Common_Coverage			0.001*					0.004**		
			(0.001)					(0.002)		
DIF_FF12_D3 X Common_Holding				0.003***					0.003	
				(0.001)					(0.002)	
DIF_FF12_D3 X Size					0.001					0.001
					(0.001)					(0.002)
Observations	27703	27703	27703	27703	27703	27703	27703	27703	27703	27703
Clusters	1409	1409	1409	1409	1409	1409	1409	1409	1409	1409
R-Squared	0.006349	0.006442	0.006515	0.007023	0.006475	0.008337	0.00864	0.008684	0.008474	0.00836

Table 12. Analysis of Daily Opening Probability of Pairs

This table reports the time-series cross-sectional logistic regression analysis of pair's opening probability. The dependent variable is the status of the pair. If the pair opens, i.e., the normalized prices widen beyond two standard deviations of historical values, the dependent variable take the value of one; and zero otherwise. The independent variables are defined similar to Table 9. In columns (2), (3) and (4) of Panel A, Common_Holding and Common_Analyst variables are continuous variables; In columns (5) and (6), Common_Holding and Common_Analyst variables are binary variables. Regression specifications in Panel B are similar to Panel A, but Panel B include the industry information diffusion measure, as well as its interaction with Size, Liquidity, Common_Holding and Common_Coverage binary variables. Several pairs characteristics control variables, including Avg_Ret_pst1mth, Avg_Ret_pst12mth, Avg_Ret_pst36mth, Avg_BM, Log_Avg_MktCap, Avg_mRetVola, though included in the regressions of Panel B, are not reported for brevity. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

Panel A: Logistic Regressions of Daily Opening Probability of Pairs from the Pairs Trading Strategy with Pairs Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	-1.992***	-1.815***	-2.368***	-2.226***	-2.582***	-2.503***	-2.161***
	(0.083)	(0.084)	(0.081)	(0.086)	(0.088)	(0.095)	(0.096)
Avg_PESPR	5.733***	4.625***	4.967***	4.477***	4.848***	4.994***	5.457***
	(1.765)	(1.720)	(1.739)	(1.712)	(1.729)	(1.740)	(1.756)
Avg_PESPR_Change	5.309***	5.455***	5.403***	5.473***	5.351***	5.356***	5.310***
	(1.070)	(1.091)	(1.050)	(1.061)	(1.069)	(1.068)	(1.071)
Avg_Turn	-0.134***	-0.115***	-0.079***	-0.076***	-0.102***	-0.097***	-0.135***
	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
Avg_dTurn_Change	0.021***	0.020***	0.022***	0.022***	0.020***	0.020***	0.021***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
News	1.644***	1.641***	1.635***	1.634***	1.641***	1.641***	1.644***
	(0.046)	(0.046)	(0.045)	(0.045)	(0.046)	(0.046)	(0.046)
Coverage	0.022	0.022	0.046***	0.044***	0.024	0.024	0.022
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Avg_Ret_pst1mth	0.378*	0.376*	0.358*	0.359*	0.375*	0.375*	0.377*
	(0.201)	(0.198)	(0.196)	(0.194)	(0.198)	(0.198)	(0.202)
Avg_Ret_pst12mth	0.030	0.025	0.029	0.026	0.025	0.025	0.030
	(0.070)	(0.068)	(0.069)	(0.068)	(0.070)	(0.069)	(0.070)
Avg_Ret_pst36mth	0.040**	0.039**	0.030*	0.030*	0.037**	0.036**	0.041**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Avg_BM	-0.155***	-0.158***	-0.065**	-0.078***	-0.127***	-0.129***	-0.150***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Log_Avg_MktCap	-0.111***	-0.104***	-0.057***	-0.061***	-0.054***	-0.063***	-0.093***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
Avg_mRetVola	1.396***	1.079**	1.867***	1.677***	1.259***	1.238***	1.402***
	(0.438)	(0.442)	(0.419)	(0.420)	(0.436)	(0.436)	(0.438)
Common_Holding_Ratio		-0.576***		-0.305***	0.119***	0.135***	
		(0.051)		(0.058)	(0.024)	(0.025)	
Common_Analyst_Ratio			-0.608***	-0.530***	0.143***	0.157***	
			(0.037)	(0.043)	(0.021)	(0.020)	
Size_Rank						-0.057**	0.066***

						(0.025)	(0.023)
Observations	825,962	825,962	825,962	825,962	825962	825,962	825,962
Clusters	1587	1587	1587	1587	1587	1587	1587

Panel B: Logistic Regressions of Daily Opening Probability of Pairs from the Pairs Trading Strategy with Pairs Characteristics, with additional industry information diffusion measure and its interactions with Size, Liquidity, Common Institutional Ownership and Common Analyst Coverage

	(1)	(2)	(3)	(4)	(5)
Intercept	-2.114***	-2.118***	-2.118***	-2.554***	-2.293***
	(0.084)	(0.082)	(0.083)	(0.091)	(0.096)
Avg_PESPR	5.816***	3.529*	5.081**	5.072***	5.425***
	(1.740)	(1.814)	(2.067)	(1.717)	(1.732)
Avg_PESPR_Change	5.340***	5.433***	5.358***	5.358***	5.349***
	(1.062)	(1.065)	(1.062)	(1.062)	(1.063)
Avg_Turn	-0.128***	-0.127***	-0.128***	-0.095***	-0.123***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Avg_dTurn_Change	0.021***	0.021***	0.021***	0.021***	0.021***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
News	1.642***	1.642***	1.642***	1.640***	1.642***
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
Coverage	0.023	0.023	0.023	0.023	0.022
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
DIF_F12_D3	0.131***	0.102***	0.132***	0.088***	0.118***
	(0.016)	(0.018)	(0.017)	(0.016)	(0.017)
DIF_F12_D3 x Liquidity		0.137***			
		(0.029)			
DIF_F12_D3 x Common_Holding			0.058		
			(0.071)		
DIF_F12_D3 x Common_Analyst				0.131***	
				(0.014)	
DIF_F12_D3 x Size					0.050***
					(0.016)
Observations	825962	825962	825962	825962	825962
Clusters	1587	1587	1587	1587	1587

Table 13. Survival Analysis of Time-to-Convergence

This table reports the survival analysis of pair's time-to-convergence conditional on the pair opens. The survival analysis applies the accelerated failure time (AFT) model with the generalized gamma distribution as the baseline hazard function. The dependent variable is the time-tol-convergence with exogenous censoring at either the 10-th trading day since pair's opening (Panel A and Panel C), or at the end of the 6-th month after a pair's opening (Panel B and Panel D). The independent variables are defined similar to Table 9. In columns (2) of Panel A and Panel B, Common_Holding and Common_Analyst variables are continuous variables; In columns (3) and (4), Common_Holding and Common_Analyst variables are binary variables. Regression specifications in Panel C (Panel D) are similar to Panel A (Panel B), but Panel C (Panel D) includes the industry information diffusion measure, as well as its interaction with Size, Liquidity, Common_Holding and Common_Coverage binary variables. Several pairs characteristics control variables, including Avg_Ret_pst1mth, Avg_Ret_pst12mth, Avg_Ret_pst36mth, Avg_BM, Log_Avg_MktCap, Avg_mRetVola, though included in the regressions of Panel B, are not reported for brevity. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	Panel A: Convergence Strategy in 10 days				Panel B: Convergence Strategy in 6 months			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Intercept	3.214***	3.164***	3.887***	3.879***	3.191***	3.124***	3.849***	3.828***
	(0.093)	(0.100)	(0.123)	(0.131)	(0.084)	(0.091)	(0.108)	(0.115)
Avg_PESPR	-6.957***	-5.382***	-5.183***	-5.210***	-6.199***	-4.026**	-3.779**	-3.851**
	(1.910)	(1.921)	(1.924)	(1.929)	(1.829)	(1.821)	(1.819)	(1.825)
Avg_PESPR_Change	-4.352**	-4.206**	-4.209**	-4.214**	-1.730	-1.419	-1.387	-1.400
	(1.762)	(1.759)	(1.766)	(1.767)	(1.750)	(1.735)	(1.735)	(1.736)
Avg_Turn	0.073***	0.048**	0.040*	0.040*	0.094***	0.068***	0.062***	0.061***
	(0.023)	(0.023)	(0.022)	(0.023)	(0.020)	(0.020)	(0.020)	(0.021)
Avg_dTurn_Change	0.001	0.000	0.000	0.000	0.014**	0.014**	0.014**	0.014**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
News	0.250***	0.246***	0.243***	0.243***	0.419***	0.417***	0.413***	0.414***
	(0.054)	(0.054)	(0.054)	(0.054)	(0.043)	(0.043)	(0.043)	(0.043)
Coverage	-0.019	-0.025	-0.019	-0.019	0.001	-0.006	0.004	0.004
	(0.031)	(0.031)	(0.031)	(0.031)	(0.027)	(0.027)	(0.027)	(0.027)
Avg_Ret_pst1mth	-0.050	-0.044	-0.036	-0.036	-0.100	-0.091	-0.084	-0.083
	(0.136)	(0.135)	(0.135)	(0.135)	(0.126)	(0.126)	(0.126)	(0.126)
Avg_Ret_pst12mth	-0.186***	-0.176***	-0.172***	-0.172***	-0.246***	-0.229***	-0.227***	-0.227***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.044)	(0.043)	(0.043)	(0.043)
Avg_Ret_pst36mth	-0.011	-0.006	-0.005	-0.005	-0.063***	-0.056***	-0.056***	-0.055***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)
Avg_BM	0.025	0.006	-0.009	-0.009	0.070**	0.055*	0.037	0.038
	(0.035)	(0.035)	(0.035)	(0.035)	(0.031)	(0.031)	(0.031)	(0.031)
Log(Avg_MktCap)	0.081***	0.066***	0.015	0.016	0.090***	0.074***	0.024**	0.027**
	(0.009)	(0.010)	(0.012)	(0.013)	(0.008)	(0.009)	(0.010)	(0.011)
Avg_mRetVola	-3.462***	-3.324***	-3.256***	-3.256***	-3.992***	-3.823***	-3.771***	-3.768***
	(0.412)	(0.406)	(0.408)	(0.408)	(0.397)	(0.396)	(0.396)	(0.396)
Common_Holding		0.376***	-0.223***	-0.225***		0.436***	-0.225***	-0.231***
		(0.074)	(0.031)	(0.033)		(0.066)	(0.028)	(0.030)
Common_Analyst		0.155***	-0.082***	-0.083***		0.154***	-0.082***	-0.086***
		(0.058)	(0.029)	(0.031)		(0.051)	(0.026)	(0.027)
Size_Rank				0.007				0.017
				(0.036)				(0.032)
Scale Parameter	1.082***	1.051***	1.068***	1.068***	1.579***	1.576***	1.575***	1.575***
	(0.068)	(0.066)	(0.066)	(0.066)	(0.009)	(0.009)	(0.009)	(0.009)
Shape Parameter	0.129***	0.173***	0.145***	0.145***	-0.644***	-0.636***	-0.646***	-0.646***
	(0.102)	(0.099)	(0.099)	(0.099)	(0.028)	(0.028)	(0.028)	(0.028)
Observations	27703	27703	27703	27703	27703	27703	27703	27703

	Panel C: Convergence Strategy in 10 days					Panel D: Convergence Strategy in 6 months				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Intercept	3.278***	3.279***	3.652***	3.716***	3.612***	3.285***	3.282***	3.669***	3.690***	3.587***
	(0.095)	(0.095)	(0.117)	(0.117)	(0.124)	(0.086)	(0.086)	(0.103)	(0.102)	(0.109)
Avg_PESPR	-6.910***	-6.969***	-5.989***	-5.410***	-5.802***	-6.030***	-4.985**	-4.779***	-4.158**	-4.773***
	(1.906)	(2.190)	(1.920)	(1.931)	(1.928)	(1.825)	(2.122)	(1.824)	(1.826)	(1.838)
Avg_PESPR_Change	-4.304**	-4.304**	-4.260**	-4.208**	-4.197**	-1.627	-1.610	-1.476	-1.380	-1.441
	(1.757)	(1.757)	(1.764)	(1.772)	(1.762)	(1.748)	(1.747)	(1.742)	(1.737)	(1.744)
Avg_Turn	0.068***	0.069***	0.048**	0.047**	0.062***	0.089***	0.089***	0.068***	0.067***	0.084***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Avg_dTurn_Change	0.001	0.001	0.000	0.000	0.000	0.014**	0.014**	0.014**	0.013**	0.014**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
News	0.253***	0.252***	0.252***	0.252***	0.252***	0.422***	0.422***	0.420***	0.422***	0.421***
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Coverage	-0.020	-0.020	-0.019	-0.013	-0.018	0.000	0.000	0.002	0.008	0.003
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
DIF_F12_D3	-0.087***	-0.088***	-0.053**	-0.049*	-0.063**	-0.121***	-0.113***	-0.083***	-0.084***	-0.099***
	(0.025)	(0.027)	(0.026)	(0.025)	(0.025)	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)
DIF_F12_D3 x Liquidity		0.003					-0.044			
		(0.048)					(0.046)			
DIF_F12_D3 x Common Coverage			-0.109***					-0.116***		
			(0.019)					(0.017)		
DIF_F12_D3 x Common Holding				-0.129***					-0.128***	
				(0.019)					(0.018)	
DIF_F12_D3 x Size					-0.091***					-0.086***
					(0.021)					(0.019)
Scale Parameter	1.072***	1.072***	1.076***	1.076***	1.073***	1.578***	1.578***	1.576***	1.576***	1.577***
	(0.067)	(0.067)	(0.067)	(0.066)	(0.067)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Shape Parameter	0.143***	0.143***	0.135***	0.134***	0.140***	-0.637***	-0.636***	-0.640***	-0.642***	-0.640***
	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)

Table 14. Analysis of Divergence Risks of Pairs Trading Strategy

This table reports the zero-inflated negative binomial (ZINB) regressions of the number of spreads widening events along the path of pairs convergence trades with the maximum holding horizon of 6-month. The zero-inflation equation ("auxiliary equation") is a logistic regression with independent variables including Converge10 (an indicator variable taking value of one if the pair converges within ten days; and zero otherwise), Avg_PESPR_Change and a constant term. Only the regression coefficients and associated *t*-statistics of the main equations are reported. In all regressions, the likelihood ratio tests and Vuong tests reject the Poison regression model and simple negative binominal regression model in favor of the zero-inflated negative binomial regression models at 1% level. All regressions compute the clustered standard errors, where the cluster is defined by year, month and industry. Standard errors are in parentheses. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	1.781***	1.750***	1.802***	1.846***	1.907***	1.997***	1.968***	1.842***
	(0.070)	(0.071)	(0.069)	(0.071)	(0.080)	(0.076)	(0.084)	(0.071)
Avg_PESPR	-4.369***	-3.762***	-4.269***	-4.251***	-4.091***	-3.723***	-3.827***	-3.081**
	(1.378)	(1.359)	(1.369)	(1.373)	(1.386)	(1.379)	(1.397)	(1.515)
Avg_PESPR_Change	0.374	0.433	0.369	0.352	0.353	0.410	0.394	0.386
	(1.109)	(1.110)	(1.111)	(1.113)	(1.111)	(1.105)	(1.106)	(1.114)
Avg_Turn	0.050**	0.035*	0.045**	0.042*	0.037*	0.029	0.038*	0.041*
	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Avg_dTurn_Change	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
News	0.107***	0.105***	0.107***	0.109***	0.110***	0.110***	0.110***	0.109***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Coverage	0.012	0.009	0.011	0.011	0.011	0.013	0.012	0.011
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Avg_Ret_pst1mth	0.033	0.042	0.035	0.035	0.036	0.043	0.036	0.038
	(0.111)	(0.109)	(0.110)	(0.110)	(0.110)	(0.109)	(0.110)	(0.110)
Avg_Ret_pst12mth	-0.109***	-0.107***	-0.109***	-0.110***	-0.109***	-0.107***	-0.109***	-0.110***
	(0.040)	(0.039)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
Avg_Ret_pst36mth	-0.057***	-0.053***	-0.056***	-0.055***	-0.055***	-0.053***	-0.056***	-0.056***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Avg_BM	0.056**	0.049*	0.051**	0.045*	0.038	0.036	0.037	0.044*
	(0.026)	(0.025)	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)	(0.026)
Log(Avg_MktCap)	0.043***	0.037***	0.040***	0.039***	0.033***	0.022***	0.025***	0.038***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Avg_mRetVola	-1.070***	-0.973***	-1.054***	-1.075***	-1.049***	-0.987***	-1.045***	-1.073***
	(0.326)	(0.323)	(0.325)	(0.324)	(0.323)	(0.322)	(0.323)	(0.323)
Common_Holding_Ratio		0.192***						
		(0.045)						
Common_Analyst_Ratio	-0.048							
	(0.036)							
DIF_F12_D3				-0.056***	-0.050***	-0.042**	-0.046***	-0.046***
				(0.016)	(0.016)	(0.016)	(0.016)	(0.017)

DIF_F12_D3 x Liquidity								-0.055*
								(0.030)
DIF_F12_D3 x Size							-0.034**	
							(0.015)	
DIF_F12_D3 x Common_Holding						-0.047***		
						(0.014)		
DIF_F12_D3 x Common_Analyst					-0.018			
					(0.013)			
Number of Clusters	1409	1409	1409	1409	1409	1409	1409	1409
Observations	27703	27703	27703	27703	27703	27703	27703	27703