Losing sight of the trees for the forest?

Pairs trading and attention shifts

Heiko Jacobs and Martin Weber*

This draft: October 2011

Appendix attached

Abstract

This paper tests asset pricing implications of the investor attention shift hypothesis proposed in recent theoretical work. Our objective is to directly assess how the dynamics of investor inattention affect the relative pricing efficiency of linked assets. We create a novel proxy for investor distraction in the time series and explore its impact in a promising and so far widely neglected setup: Stock pairs trading (Gatev et al. (2006)), a popular proprietary relative arbitrage approach. Relying on almost 50 years of daily data for the US stock market as well as on evidence from eight major international stock markets, we provide broad and robust evidence for substantial distraction effects. For instance, the average one-month return on long-short US stock pairs that open on high distraction days is about twice as high as the return on pairs that open on low distraction days. Several conceptually quite diverse tests confirm the idea of time-varying investor attention affecting pairs trading profitability.

Keywords: Limited attention, investor distraction, attention shifts, pairs trading, relative-value arbitrage

JEL Classification Codes: G12, G14

---

*Heiko Jacobs is from the Lehrstuhl für Bankbetriebslehre, Universität Mannheim, L 5, 2, 68131 Mannheim. E-Mail: jacobs@bank.bwl.uni-mannheim.de. Martin Weber is from the Lehrstuhl für Bankbetriebslehre, Universität Mannheim, L 5, 2, 68131 Mannheim, and CEPR, London. Email: weber@bank.bwl.uni-mannheim.de. We would like to thank seminar participants at the 2011 Helsinki Finance Summit on Investor Behavior, at the 18th Annual Meeting of the German Finance Association as well as at the University of Mannheim for valuable comments. This paper is scheduled for presentation at the 2011 Financial Research Association Meeting.

1 Introduction

Relaxing the strict assumptions of traditional models, recent theoretical work argues that investors have limited information processing capabilities. Consequently, they have to optimally allocate their finite attention across several aggregation levels, which is done depending on priority and urgency. As Peng and Xiong (2006) put it: "In severely constrained cases, the investor allocates all attention to market- and sector-level information and ignores all the firm-specific data" (p. 565). In this paper, we empirically explore asset pricing implications of this attention shift hypothesis. Our objective is to directly test how the dynamics of investor inattention affect the price formation of linked assets.

We create a novel proxy for time-varying investor distraction and explore its role in a natural, promising, and so far widely neglected setting: Pairs trading (Gatev et al. (2006)), a popular relative-value arbitrage approach, which bets on the future performance of two assets with very similar past performance. Relying on close to 50 years of daily data for the US stock market as well as on insights from eight major international markets, we provide broad and robust evidence for distraction effects. For instance, pairs trading is much more profitable than usual when stocks diverge on so called high distraction days, during which turbulent market conditions are assumed to demand investors' full attention. It is much less profitable than usual when stocks diverge on low distraction days, during which we expect sufficient resources to process complex interactions at the firm-level.

Given the vast amount of news available in financial markets, recent models propose and empirical work verifies that investors exhibit category learning behavior[1]: Most effort is typically spent on processing news relevant primarily on some aggregated level, as this information tends to be most important for the valuation of an investor's overall portfolio. The remaining capacities are used to process more disaggregated (e.g. firm-level) news.

[1] Theoretical work includes e.g. Barberis and Shleifer (2003), Peng (2005), or Peng and Xiong (2006). On the empirical side, a vivid example for category thinking is given in Cooper et al. (2001): During the internet bubble, firms that simply changed their name to dot.com names, but not their business model, experienced positive abnormal post-announcement returns. Similarly, mutual funds that change their names for cosmetic reasons to appear more like a current hot return style, have been shown to attract positive abnormal inflows without improving their performance (Cooper et al. (2005)). An implication of category learning is excessive return comovement as recently identified in various settings. See e.g. Barberis et al. (2005), Greenwood (2008), or Boyer (2011) on index effects, Green and Hwang (2009) on price effects, or Pirinsky and Wang (2006) on location effects.

The dynamics of this setup lend support to the idea that the relative inattention to disaggregated information should be particularly high in exceptional market conditions: The need to focus on understanding shocks in the big picture should leave fewer resources available to concentrate on details. Empirically investigating this prediction first requires answers to two questions: How to proxy for unobservable investor attention allocation? And which return anomaly is likely to be particularly affected by attention shifts?

A few studies so far have dealt with these questions and thereby focussed on explaining variations in the post-earnings announcement drift. We contribute to this literature in two ways: We propose a novel proxy for investor distraction, and we apply it to pairs trading, a conceptually different and untested setup. The proxy aims at quantifying the unexpected daily information load market participants need to process in order to timely assess the overall market situation. Building on the premise that information shocks partly manifest themselves in abnormal returns, we do so by condensing the magnitude and dissemination of unanticipated daily return shocks in a broad range of market segments into a single ratio. We perform yearly decile sorts of the proxy in the baseline analysis, and particularly concentrate on "high distraction days" (decile 10) as opposed to "low distraction days" (decile 1). We then test whether the proxy has predictive power for the profitability of short-term pairs trading, whose mechanisms are illustrated in section 2.2. In short, out of the in total 200 million eligible sets of two stocks in our baseline US sample, we find those whose prices have moved together the most historically. Each month, we select the 100 pairs with minimum distance between normalized historical return paths, and then trade them over the adjacent six months. Specifically, whenever the difference in cumulative daily returns of any of these top pairs exceeds a certain threshold, we short the relatively overpriced winner and buy the relatively underpriced loser. If the future resembles the past, prices are likely to finally convergence again, thereby generating positive returns on zero-cost portfolios. We are interested in whether it makes any difference whether stocks diverge on high or low distraction days. Do shocks in limited attention towards firm-level information hinder market participants from keeping relative prices in line?

Please insert figure 1

The essence of our findings is captured in figure 1. It displays, in event-time, the average one-month return on long-short US stock pairs by distraction proxy decile ranks. Findings are based on more than 100,000 round-trip trades between January 1962 and December 2008. The lower (upper) line can be interpreted as lower (upper) bound for the return achieved by the strategy (see section 2). There appears to be a close to monotonic increase in the profitability by distraction deciles. The return on pairs opening on low (high) distraction days is far lower (higher) than returns on average. The difference between decile 1 and 10 is not only highly statistically, but also economically significant: The return on high distraction days is more than twice as high as the return on low distraction days. This time-varying profitability of pairs trading might also be interpreted in the sense of Grossman and Stiglitz (1980). The distraction proxy is likely to identify moments in which gathering and processing firm-specific information is particularly costly, so that the market has to provide higher payoffs as compensation.

These return differences appear robust. For instance, they are not sensitive to the specific design of the distraction proxy. They are also quite persistent over time, including the recent past. An approach which aims at isolating the distraction effect in calendar time exhibits little exposure to well-known risk premia. Firms tend to be large and liquid, and standard pair characteristics are very similar across distraction deciles. To mitigate concerns that unobserved variables might drive our results, we also examine, with similar results, returns to a subset of pairs that happen to diverge on both high and low distraction days.

Moreover, findings are not confined to the US market. The return difference between high and low distraction days is a persistent phenomenon which, with varying degree, is observable in each of the eight major non-US stock markets we additionally study.

To gain deeper insights, we conduct a variety of additional tests whose results appear in line with the idea of the dynamics of investor distraction contributing to pairs trading profitability. Alternative proxies for limited attention, which we derive from the previous literature, often have an incremental effect. Pairs opening immediately before holidays, when investor distraction is likely to be particularly high, tend to be more profitable and to converge more often than pairs on average. The impact of investor distraction appears lower for pairs consisting of firms from the same industry or for pairs consisting of whole value-weighted industries. Finally, pairs particularly neglected (covered) by the media appear more (less) profitable, and exhibit a higher (lower) sensitivity to changes in the level of investor distraction.

Our empirical approach provides a promising setup to gain insights about the impact of time-varying investor inattention on asset pricing for several reasons.

1. Its quantitative nature allows us to identify sets of linked firms for which cross-stock information transfer is likely to be inhibited in moments of high distraction. The monthly top 100 pairs represent only an extremely small fraction of all eligible pairs, and are identified by exhaustive matching in normalized daily price space. In the baseline analysis, we concentrate on firms from different industries. In this way, it provides an intuitive, elegant way of identifying firms, which, despite mainly operating in different segments, are likely to be somehow economically related. Such pairs are interesting candidates for our scenario as industrial boundaries have been shown to along with informational boundaries induced by specialization of important market participants such as analysts or fund managers (e.g. Hong et al. (2007), Menzly and Ozbas (2010)). Moreover, there appears to be substantial information content in residual pairwise stock return comovement, even when an exhaustive list of explaining variables is relied on (e.g. Chen et al. (2010), Chordia et al. (2011)). In sum, the link between two typical firms in our analysis might be thought of as potentially being strong, but simultaneously often also less explicit, obvious, and transparent, and thus prone to being neglected particularly easily.
2. The type of return predictability in pairs trading is different from the type of return predictability that has been linked to limited attention in the literature so far. Previous studies assets have analyzed the lagged price response of stocks to their own past returns (e.g. Hong et al. (2000)), lead-lag effects between portfolios of stocks (e.g. Hong et al. (2007), Hou (2007)) or return predictability along the supply chain (Cohen and Frazzini (2008), Menzly and Ozbas (2010)). Pairs trading, however, is about predicting the relative short-term performance of two individual, typically rather large stocks with an often non-obvious relationship, out of which neither is the systematic leader. Linking this type of cross-predictability of returns to variations in investor distraction, is, to our knowledge, new.
3. The nature of pairs trading profits fits well with the idea of attention constraints

impeding timely information spill-over. It is a short-term strategy, whose profitability tends to almost monotonically decline in event-time (e.g. Engelberg et al. (2009) and section 3). Importantly, the day of divergence appears to be a critical date. In fact, a large fraction of the cumulative return difference upon divergence is attributable to the day of divergence itself. Thus, identifying circumstances in which this behavior is ex ante more likely to be caused by temporary market frictions is a key to the strategy's success.

4. In general, comprehensive empirical studies on pairs trading are still rare. This is surprising given its large seemingly abnormal returns reported in Gatev et al. (2006) as well as its apparent popularity among practitioners. Moreover, very little is known about pairs trading in international markets, even though only very few trading strategies have survived the test of time and independent scrutiny. As a consequence, it is still an open question when, where, and why pairs trading is particularly profitable. We address this gap in the literature with a data set comprising about 14,000 stocks with 25 million firm days from eight major non-US stock markets.

5. Our findings might shed light on other pervasive empirical puzzles. A number of scenarios are related to pairs trading in that there also appear primarily short-term price discrepancies between similar assets, which are often difficult to reconcile with standard theory.[2] In a broader sense, our results might thus help to better understand how limited attention affects the efficiency of related assets in practice.

The remainder of this paper is organized as follows. Section two discusses related research and explains our empirical design. Section three presents baseline findings, both for the US and international markets. Section four provides a number of robustness checks. Section five contains further tests that provide additional evidence of the link between time-varying investor inattention and pairs trading profits. Section six concludes.

[2] For instance, Lee et al. (1991), Pontiff (1995), Chay and Trzcinka (1999), and Cherkes et al. (2009) focus on the relationship of the prices of closed-end fund shares and the per share market value of the assets hold by the funds. Lamont and Thaler (2003) and Mitchell et al. (2002) study situations where a firm's market value is less than the value of its ownership stake in publicly traded subsidiary. Scruggs (2007), Rosenthal and Young (1990), Froot and Dabora (1999), and Jong et al. (2009) study price parity deviations of dual-listed companies ("Siamese Twins"). Gagnon and Karolyi (2010) study discrepancies between the prices of US and home-market shares of companies with cross-listed stocks. Smith and Amoako-Adu (1995), Zingales (1995), and Schultz and Shive (2010) study dual class shares issued by the same company, that differ in voting rights, but have equal cash flow rights.

2 Related Literature and Empirical Design

2.1 Limited Investor Attention

Extensive evidence from psychological research shows that attention is a scarce cognitive resource. Focussing attention on one task necessarily goes along with a substitution of cognitive resources from other tasks (Kahneman (1973)). Building on these insights, a growing body of empirical and theoretical work highlights the importance of attention constraints in finance. This research argues that market participants have to be selective in information processing and thus potentially neglect value-relevant information. In the presence of at least some limits to arbitrage, this can induce temporary mispricings.

As investor attention allocation is not observable, the identification of promising proxies is a challenge for empirical work. A substream of this literature defines simple, intuitive time-series proxies for limited investor attention and employs them to explain variations in the magnitude of return anomalies. So far, these papers almost exclusively focus on the post-earnings announcement drift. Hirshleifer et al. (2009) use the number of same day earnings releases. They find that, on days where many earnings announcements compete for investors' attention, average immediate market reactions are weaker, but post-announcement abnormal returns are higher. Qualitatively similar results are reported by DellaVigna and Pollet (2009) and Peress (2008). DellaVigna and Pollet (2009) rely on Fridays, on which, as they argue, investors are distracted by the upcoming weekend. Peress (2008) additionally employs the daily number of firms covered in the Wall Street Journal. Hou et al. (2009) use down market periods, during which investors are assumed to often "put their heads in the sand" (see also Karlsson et al. (2009)). They study, in addition to the post-earnings announcement drift, variations in momentum profits.

We contribute to this literature by providing consistent evidence based on a novel proxy, which relies on the intuition behind models on the dynamics of attention allocation such as Peng (2005) or Peng and Xiong (2006). It aims at identifying days on which market participants are likely to be forced to spend more (or less) resources than usual on understanding "the big picture". Implementing this idea leaves many degrees of freedom. To assure robustness, we first construct a baseline proxy and then, in section 4.6, extensively test the sensibility of our findings with ten modified proxies.

For the construction of the baseline proxy, a four-step procedure is employed. First, for January 1960 to December 2008, we compute daily value-weighted returns for the 49 Fama and French (1997) industries[3], thereby taking into account all common shares (CRSP share code 10 or 11) traded on NYSE, AMEX or NASDAQ. Second, we decompose industry returns to construct daily industry-specific return shocks. Specifically, the shock $AR_{i,t}$ for industry i at day t is defined as the absolute difference between the actual industry return $R_{i,t}$ and its expected return as given by a simple OLS market model:

$$AR_{i,t} = \mid R_{i,t} - \widehat{\alpha}_{i,t} - \widehat{\beta}_{i,t} R_{m,t} \mid \tag{1}$$

Parameter estimates are obtained from rolling time-series regressions based on daily return data over the previous year. We later augment the model with well-established risk factors. Third, we condense these shocks into a single measure $Distraction_t$. Out of the several plausible weighting schemes, we choose an approach that takes the expected level and frequency of industry-specific shocks into account. Each industry weight $w_{i,t}$ is determined by the inverse of the volatility $\sigma_{i,t}$ of the industry shock variable $AR_{i,t}$ as follows:

$$Distraction_t = \sum_{i=1}^{49} w_{i,t} AR_{i,t} \text{ where } w_{i,t} = \frac{\frac{1}{\sigma_{i,t}}}{\sum_{i=1}^{49} \frac{1}{\sigma_{i,t}}} \tag{2}$$

Shock volatilities $\sigma_{i,t}$ are estimated from $AR_{i,t}$ over the previous year. The weighting approach fits with the intuition that a pronounced return shock in an industry for which large shocks are a common occurrence is less likely to unexpectedly demand extra attention than a pronounced shock in an industry which "usually behaves as expected".[4] Figure 3 displays the time-series of the resulting raw proxy from January 1962 to December 2008.

Please insert figure 3

[3]Specifically, we use the 48 industries defined in Fama and French (1997), and group stocks that are not assigned to any industry in category 49. In the baseline approach, we focus on this industry classification as markets are partially segmented by industrial boundaries (e.g. Menzly and Ozbas (2010)), as the Fama and French (1997) system appears to identify economically similar companies (e.g. Chan et al. (2007)), as it is widely relied on in academic studies, and as it is available for the whole sample period.

[4]In unreported findings, we find that volatility-weighting often appears a compromise between value-weighting, where few large industries substantially drive the aggregate measure, and equal-weighting, where the impact of small industries is much stronger. Nevertheless, all three weighting schemes result in highly pairwise correlated (0.9 or greater) distraction proxies. All weighting schemes are relied on in later robustness tests in section 4.6.

An intuitive way of exploring its ability to identify distracting moments is to explore when it reaches its maximum. Indeed, these days are appealing. The highest values (in descending order) are achieved on October 20, 1987 (the day following the stock market crash), on March 15, 2000 (massive sell-off of technology stocks towards traditional industries right before the burst of the bubble), and on September 17, 2001 (first trading day after 9/11).[5] Figure 3 also reveals that there are several phases (but no general time trend), in which proxy values typically differ substantially from the sample average.[6]

To thoroughly test for distraction effects despite these episodes, we perform yearly sorts of the proxy as the final fourth step. For each year separately, we assign a decile-based rank to each trading day. Pooling the data results in roughly 1,180 days for each decile rank. Figure 3 shows the time-series of the monthly number of high distraction days (decile 10) and low distraction days (decile 1). It also shows Spearman rank order correlation coefficients between distraction proxy decile ranks and the rank order of market-level variables. The proxy is only weakly correlated with standard risk premiums and (moderately) positively with factors each assumed to capture a specific aspect of turbulent markets (squared market return, market turnover, rolling ten day return volatility from t-10 to t-1).

The appendix illustrates the relationship between the distribution of (different types of) abnormal industry returns and distraction proxy deciles ranks. Higher ranks tend to go along both with a generally more pronounced level of abnormal returns and with a higher dispersion of these returns. However, the maximum weight of a single industry-level shock is similar across decile ranks and moreover tends to be moderate. Together, these findings suggest that high distraction days typically identify moments where there is widespread turbulence across markets segments.

2.2 Pairs Trading

There are still only few comprehensive empirical studies on pairs trading so far, possibly due to its proprietary and computationally intensive nature. Gatev et al. (2006) report

[5] Note that the proxy not simply picks up market movements. The value-weighted market return on these days is 0.40%, 0.68% and -5.07%, respectively.

[6] This appears in line with recent findings on the behavior of idiosyncratic volatility (e.g. Brandt et al. (2010), Fink et al. (2010)).

statistically and economically significant profits between 1962 and 2002, which are not driven by standard risk factors, unrealized bankruptcy risk, or short sales constraints. Focussing on same-industry pairs between 1993 and 2006, Engelberg et al. (2009) further explore cross-sectional characteristics of pairs trading. They find that part of the profits to pairs trading seem to stem from differential immediate response to common information, i.e. news that affects both stocks in the pair. Thus, studying dynamics in the level of investor distraction, which might cause such frictions, seems an intuitive way to gain deeper insights. Do and Faff (2010) and Do and Faff (2011) report a declining trend in standard pairs trading profitability over the recent past, which is partly driven by higher fraction of nonconvergent pairs. Again, identifying and understanding scenarios in which pairs are ex ante more likely to converge appears crucial to understand the price formation process. In the international context, Andrade et al. (2005) document annual excess returns of about 10% for the Taiwanese stock market between 1994 and 2002. They show that uninformed trading shocks are a major driver of the strategy's profitability.

For our empirical analysis in the US stock market, we obtain daily stock price data on all common shares (CRSP share code 10 or 11) traded on NYSE or AMEX on any time between January 1960 and December 2008. We impose several restrictions to assure that only relatively large and liquid stocks enter the analysis. We discard all stocks with at least one missing return or zero trading volume on any day of the 12 months estimation period, during which pairs are matched. Moreover, we only consider stocks whose market capitalization is larger than the median of the NYSE/AMEX stock universe at that time. To mitigate data mining concerns and to facilitate comparison with previous work, we widely follow the methodology developed in Gatev et al. (2006).[7] Specifically, at the first day of the 12 months estimation period, we set the price of each eligible stock to equal unity. We use daily price data to compute stock-specific time-series of cumulative total returns (with reinvested dividends) over the whole estimation period. A simple algorithm is then relied on to determine to what extent two stocks, which we require to belong

[7]Note that their setup slightly differs from our baseline scenario in several dimensions. Their sample period is shorter, their eligible stock universe broader, their maximum holding period longer, and their method to identify the top 100 pairs slightly different. In unreported results, we have replicated their main analysis to the extent possible. We obtained findings very similar to theirs. In section 4, we assess the sensitivity of our results. They are robust with regard to several plausible changes in methodology (i.e. regarding maximum holding period or top pair identification). Moreover, we provide out-of-sample evidence for in total eight major international stock markets.

to different (out of the 49) Fama and French (1997) industries, have moved together historically. The algorithm is intended to provide a parsimonious, intuitive framework to identify pairs. Let $R_{i,t}$ ($R_{j,t}$) be the normalized return series of stock i (j) in estimation period t, comprising of trading days 1 to n. The distance measure is then defined as:

$$\frac{1}{n}\sum_{i=1}^{n}(R_{i,t} - R_{j,t})^2 \tag{3}$$

We compute this value for all possible pair combinations, whose number grows quadratically with the number of eligible stocks. Then, we choose, at the beginning of each month, the top 100 pairs with minimum distance. These top pairs only represent a tiny fraction (on average less than 0.03%) of all pairs, which aims at identifying strongly linked firms. The 100 pairs are then eligible for trading in the immediately following six months evaluation period. At the beginning of this period, prices are again set to equal unity. If the spread between the cumulative return series of two substitutes exceeds a certain threshold, we go long in the relatively underpriced stock and short in the relatively overpriced stock. Following Gatev et al. (2006), we open a pair if prices diverge by more than two historical standard deviations, as estimated from equation 3. The self-financing pair is then hold for up to one month. If prices convergence before this cut-off date, the trade is closed with a gain. If prices do not convergence within a month, positions are offset, which, if prices diverge even further, results in a loss.[8] A pair may trade several times during the six months evaluation period. The amount of money invested in later trades differs depending on whether we report event-time results (the baseline analysis) or calendar-time results. In event-time, we just again go one dollar long (short) into the cheap (expensive) stock. In calendar-time, proceeds from previous trades are reinvested, which implies that pairs in a portfolio are weighted by the cumulative returns of the component pairs. The bottom-line differences between both methods are small though. We initiate the pairs estimation period at the beginning of every month from January 1960 to July 2007, leading to an evaluation period from January 1961 to December 2008.

Figure 2 illustrates the trading process with examples. Pairs therein open several times during the trading period, however not always in the same direction. This is a common

[8]A third reason for closing a pair is delisting of a firm. In this case, we use the delisting return or the last available price. Unreported results suggest that the economic impact of this scenario on our findings is weak as the likelihood of delisting within the month after divergence is low. The qualitative nature of our main findings remains unchanged even if we set the return of the long stock to -100% when it is delisted.

behavior. Out of the large fraction of pairs that open in total at least three times, roughly 85% hold each stock at least once in both a long and a short position. Therefore, the phenomenon is different from the lead-lag relationship studied in earlier work (e.g. Lo and MacKinlay (1990), Hou (2007), Hong et al. (2007)).

Please insert figure 2

Despite the strict screening process as outlined above, microstructural effects might still be an issue. To mitigate these concerns, we report the results from two return computation schemes, that might be considered a lower and upper bound for the magnitude of our findings. For the upper bound, we simply compute returns on zero-cost portfolios between the day of divergence (closing price) and the day of convergence (closing price). For simplicity, we call this scheme "no waiting". For the lower bound, we employ a more conservative return computation approach as discussed in Gatev et al. (2006). Specifically, we skip one day after the divergence and add one day following the crossing of the prices. This method is intended to account for the impact of the bid-ask spread. Moreover, it works strongly against finding effects attributable to investor distraction, as any information overlooked at the day of divergence might be impounded into prices during the next day without entering our return estimates. We call this scheme "one day waiting".

3 Baseline Results

3.1 US evidence

We have to ensure that our findings capture the impact of variations in investor inattention rather than variations in other important variables. Therefore, table 1 compares firm-level and pair-level variables separately for all trades, for trades opening on low distraction days (decile 1) and for trades opening on high distraction days (decile 10). Inferences are very similar when we also include the remaining deciles in the analysis.

Please insert table 1

Characteristics related to liquidity and limits to arbitrage are deemed particularly relevant. Our proxies for liquidity comprise market capitalization (NYSE/AMEX decile rank), the Amihud (2002) illiquidity ratio and average pre-event turnover. The latter two are estimated from daily data over the pairs formation period. Besides market capitalization, we employ idiosyncratic risk as a proxy for limits to arbitrage (e.g. Pontiff (2006)). The role of arbitrage constraints will later be analyzed in more detail. Following previous literature, we compute idiosyncratic risk as the volatility of the residual from time-series regressions of daily stock returns on factors for the market premium, size, value and momentum. Again, the twelve months immediately preceding the pair's trading period serve as estimation period. Table 1 shows four main findings. First, as aimed at with our selection criteria, firms in general tend to be large and liquid. The medium firm belongs to the ninth NYSE/AMEX decile and has an average daily turnover of 0.11%. Relying on IBES data from 1980 on (see e.g. Hong et al. (2000)), the medium firm is covered by nine analysts. Second, there are typically only small differences in firm characteristics within pairs and across distraction deciles. The only statistically significant result is found for idiosyncratic risk, where differences seem small from an economic perspective. Third, with regard to industry structure, both firms and pairs are, in the overall picture, well diversified. For instance, both on high and low distraction days, firms from all 49 industry groups and pairs from well more than 600 industry group combinations are traded. However, utility stocks pose an exception. They make up close to 30% of all sample firms and are part of all top industry group combinations. We address this issue in later tests. Fourth, the day of divergence appears an interesting date. Pairs on average are opened when cumulative standardized returns have diverged by 6.68%. More than 40% of this difference are on average attributable to the day of divergence itself. Thus, understanding what causes prices of related stocks to diverge exactly on these days is critical for the success of any strategy that bets on short-term convergence. On high (low) distraction days, the return spread at the day of divergence is significantly larger (smaller) than on average, which seems consistent with an investor attention story.

We first perform univariate analysis to examine the impact of time-varying investor distraction on pairs trading. To this end, we compare the average event-time return on pairs sorted by distraction proxy decile ranks as observed on the day of pair divergence. Table 2 shows findings based on more than 100,000 round trip trades from January 1962 to

December 2008. Panel A (B) reports findings under the “no waiting” (“one day waiting”) return computation scheme. The first row in both panels reports returns on zero-cost pairs generated within the month following the day of divergence. If pairs converge before this cut-off date, we assume that the proceeds are hold in cash with zero interest rate until the month has passed. As all cash-flows before the cut-off date are positive by construction, this is a conservative approach.

Please insert table 2

In line with previous literature, traditional pairs trading averaged over the whole sample period appears highly profitable. One month returns are estimated between 97 (“one day waiting”) and 138 (“no waiting”) basis points. When analyzing returns by investor distraction deciles, a clear pattern emerges. Pairs opening on low distraction days are far less profitable, both from a statistical and an economic point of view, than pairs on average. The one month return ranges only from 53 (“one day waiting”) to 89 (“no waiting”) basis points. On the other hand, pairs opening on high distraction days are far more profitable than pairs on average. Estimates range here from 130 to 190 basis points. The difference between decile 10 and decile 1 amounts to highly significant and economically meaningful 77 to 101 basis points per month. The effect is not confined to the extreme distraction deciles: Decile 2 to 9 show an almost monotonic increase in profitability. The appendix provides more detailed information about the return distribution by distraction deciles.

To gain more insights, we study the mechanisms behind these return differences. In general, higher returns on pairs opening on specific days can stem from three sources: First and foremost, the probability of convergence can be higher. This is indeed what a limited attention story would predict for high distraction days. As implied by e.g. the model of Peng and Xiong (2006), we expect, all else equal, cross-stock information to diffuse more slowly during these times. Imagine, for instance that common news is released, which clearly and directly affects the first firm in the pair, but only indirectly and less clearly the second firm. Market frictions due to high investor distraction are then likely to prevent the news from being instantaneously and fully impounded into the price of the second stock. This might induce temporary price divergence and thus the opening of the pair. Consistent with this line of reasoning, far more pairs open on high distraction than low distraction days (see table 2). When investors finally become fully aware of the link between both firms, relative prices should adjust gradually and the pair is likely to finally converge again. The second row of panel A and B shows that this prediction is supported by the data. The average fraction of pairs converging within the month after divergence is 36.3%. For pairs opening on low distraction days, however, the convergence rate is only 33.4%. This value almost monotonically increases by distraction deciles, culminating in a convergence rate of 39.7% for decile 10. In other words, simply switching from low distraction to high distraction days increases the likelihood of convergence by almost 20 percent. Figure 4 shows the probability of convergence on a given day in event time. In line with the idea of time-varying investor distraction being an important driver of divergence, the likelihood of convergence within the first event days is considerably higher (lower) for pairs diverging on high (low) distraction days. After about five days, convergence rates begin to approximate each other more closely, until they appear indistinguishable.

Please insert figure 4

Higher returns might also stem from average returns conditioned on convergence being higher. Arguably, this is also what a limited attention explanation of our findings would suggest. To the extent that a slower cross-stock information flow on high distraction days translates ceteris paribus into a higher cumulative return difference at the time of divergence, we would expect to finally gain higher returns upon convergence. Table 2 shows findings supporting this line of reasoning. The difference in returns upon convergence between decile 10 and decile 1 is estimated to range between 56 ("one day waiting") and 106 ("no waiting") basis points. In other words, and as also shown in table 2, simply switching from low distraction to high distraction days appears to increase the average return upon convergence by roughly 10%. Again, figure 4 provides graphical evidence. It also shows that average daily returns almost monotonically decline in event-time, in particular for pairs opening on high distraction days. This again highlights the important role of the day of divergence.

Finally, a third potential source of profit is that the average return conditioned on non-convergence could be less negative for pairs opening on high distraction days. A limited attention story does not imply that this should be the case: Non-convergence is (comparatively more) suggestive of idiosyncratic news affecting only one stock in the pair (e.g. Engelberg et al. (2009)). As this type of information is arguably often easier to grasp and process than common news affecting the potentially complex relationship between both firms in the pair, attention constraints should be less binding. Again, table 2 displays findings consistent with this line of reasoning. The difference in returns upon non-convergence between decile 10 and decile 1 is virtually zero and statistically insignificant. In fact, the returns for this scenario are similar across all distraction deciles.

Taken together, findings are in line with the implications of investor distraction. Pairs opening on high distraction days are more attractive for two reasons: They are more likely to convergence, and, if they do, they generate higher abnormal returns.

To control for other factors that might partially drive our findings so far, we conduct several multivariate tests. Main results are presented in table 3. The dependent variable is the pooled one-month event-time return on long-short pairs. For brevity, we only report the more conservative results from the "one day waiting" return computation scheme. The independent variable of interest is the investor distraction proxy. In different specifications, we employ either the distraction proxy decile rank or a high/low distraction dummy, which is zero for low distraction days (decile 1) and one for high distraction days (decile 10). The remaining independent variables comprise up to three control sets. The first set controls for calendar and industry effects (indicator variables for year, month, day of week as well as pair industry group combinations). The second set controls for market-level conditions on the day of divergence (market return, squared market return, market turnover, 10 day rolling volatility, factors for daily return premia on size, value, momentum and short-term reversal). The third set includes almost all pair and firm characteristics outlined at the beginning of this section (see table 3 for details).

Please insert table 3

Depending on the model specification, the difference in one month abnormal returns between pairs opening on high distraction days and those opening on low distraction days is estimated to range from 39 to 73 basis points in the multivariate case. Coefficients remain all strongly statistically significant at the one percent level.

3.2 International evidence

A powerful way to test the validity of our baseline findings is to evaluate the success of the approach in independent samples. This is particular appealing as, up till now, hardly anything is known about the nature of pairs trading profits in other major international stock markets. Finding similar return patterns across countries would strongly suggest that our findings represent a generalized phenomenon rather than are attributable to US-specific factors or elaborate data mining. Therefore, we study the dynamics of pairs trading profits in eight countries. This number is somewhat arbitrarily set and meant to be a compromise between maximizing the sample size and minimizing the fraction of error-prone daily return and volume data as well as the number of months with too few eligible stocks for a reasonable analysis. Given this trade-off, we rely on Japan, UK, France, Germany, Switzerland, Italy, Netherlands, and Hongkong. These markets represent the eight largest non North-American stock markets based on domestic stock market capitalization at the end of 2002, as reported by Datastream. This date roughly represents the middle of the sample period for most of these countries, for which we gather data from the Compustat Global Daily Stock File. Depending on the availability of reliable trading volume data, the sample period starts at some point in the middle of the 90ies and ends, for all markets, in December 2009. The appendix gives more detailed information about the samples. In total, the analysis is based on an initial data set of about 14,000 stocks accounting for 25 million firm days. The computation of the country-specific distraction proxy relies on the 10 GICS industry sectors. For the country-specific monthly top 100 pairs, we discard all stocks with at least one missing return or at least two zero/missing trading volume days within the 12 months estimation period. Apart from that, the analysis closely mirrors the US approach in table 2. Main findings from in total about 200,000 round-trip trades are displayed in table 4.

Please insert table 4

Findings reveal that traditional pairs trading appears in general highly profitable in all countries. This may seem surprising given the fact that we focus on the recent past in which returns on pairs trading in the US have been far smaller than in earlier periods. However, even under the conservative "one day waiting" scheme, annualized returns range from 6%

(Italy) to more than 13% (Germany, France). While these results are interesting in their own right, we again focus on the role of investor distraction at the day of divergence. We find strong evidence for distraction effects. With the exception of Japan, the return on pairs opening on low (high) distraction days is smaller (larger) than average sample returns in every singly country. The positive return difference between decile 10 and decile 1 is, with the exception of Italy and Japan, persistently and strongly statistically significant. In addition, their size is economically meaningful and sometimes even very large. Moreover, the nature of pairs trading profits in Japan does not seem to be that different. Focussing on distraction quintiles instead of deciles yields findings which become more in line, both statistically and economically, with the ones obtained for the other countries. Digging deeper, panel C of table 4 reveals that the key to the pronounced profits on high distraction days is again a higher likelihood of convergence. From low distraction days to high distraction days, the fraction of converging pairs increases between 12% (from 43.8% to 49.1% in Japan) to 87% (from 15.9% to 29.8% in Hongkong). Remarkably, while self-financing pairs trading generates seemingly abnormal returns in all countries, its nature appears to differ substantially e.g. with regard to the number of pairs traded, the impact of the "one day waiting" scheme, or the (unconditional) fraction of converging pairs. Exploring the sources and consequences of these cross-sectional differences might be an interesting field for further research. In any case, the results in the overall picture strongly confirm the baseline results obtained for the US market.

4 Robustness Checks

In this section, we test the sensitivity of our baseline findings from various perspectives. For the sake of brevity and if not mentioned otherwise, we only report results obtained under the more conservative "one day waiting" scheme.

4.1 Subperiod analysis

To assess whether our findings are robust across time, we repeat the analysis for three consecutive subperiods of (close to) equal length. Panel A of table 5 shows findings for the periods from 1962 to 1977, from 1978 to 1993, and from 1994 to 2008.

Please insert table 5

In line with results from previous work (Gatev et al. (2006), Do and Faff (2010)), returns to traditional pairs trading seem to decline over time, though they remain statistically significant. For the most recent subperiod, the one-month return is only 24 basis points, possibly as a result of the increasing popularity of such strategies as well as decreasing transaction costs. More importantly though, in all cases, returns originating from divergence on high (low) distraction days are higher (lower) than on average. In fact, between 1994 and 2008, returns on pairs opening on low distraction days are even negative. Nevertheless, for high distraction days, the average return in the same period is 78 basis points, which is three times larger than the return obtained from unconditional pairs trading. The difference between decile 10 and decile 1 is highly significant in all subperiods, both economically and statistically. In sum, our results appear robust across time.

4.2 Variations in the data set

As shown in table 1, firms from the utility sector represent about 30% of all firm-level observations. To analyze whether our findings represent a widespread phenomenon, we control for the impact of utility firms in two ways. In the first scenario, we simply exclude these stocks. We rerun the selection process and the baseline analysis, but only consider pairs that do not include any utility firm. In the second scenario, we identify the monthly top 100 pairs under the constraint that each firm is only considered once at maximum. This approach not only decreases the fraction of utility stocks to roughly 17%, but also changes the composition of the data set considerably.[9] Panel B of table 5 verifies, however, that the baseline findings are robust to such variations in the eligible pairs universe.

[9] Again, we rank pairs by minimum distance of firms' normalized return series. However, we skip a pair if at least one of the firms is already a component of any higher-ranked pair. As a result, the top 100 pairs always consist of 200 different firms. The rank of the last selected pair is always 100 in the baseline scenario, but on average 782 in the new scenario.

4.3 Alternative econometric approach

We also modify our empirical design by running Fama-MacBeth-type regressions. We first estimate yearly pooled cross-sectional regressions of one-month pairs returns on distraction proxy decile ranks and then use the time-series of the resulting coefficient to assess its statistical significance. Panel C of table 5 shows the result. The coefficient is positive in about 75% of all years, highly statistical significant, and economically meaningful.

4.4 Same pairs

Though descriptive statistics and multivariate tests suggest that firm and pair characteristics are widely comparable across distraction deciles, the analysis could potentially still suffer from an omitted variables problem. To mitigate these concerns, we analyze a subsample specifically designed to isolate the impact of variations in investor distraction. We restrict our focus to those pairs which happen to diverge both at least once on a low and at least once on a high distraction day. Implementing this idea requires to impose a restriction on the maximum time span between these events. We here report results for the subset of pairs, for which the time difference between the average date the divergence on low distraction days took place and the average date the divergence on high distraction days took place, is less than a year.[10] In total, this leaves 5,488 trades on high or low distraction days. This procedure controls for all firm and pair-level variables, including unobserved ones, that do not vary within this typically rather short time period. Findings are shown in panel D of table 5. Results verify that inferences from our baseline findings carry over.

4.5 Implementability

The distraction proxy is constructed from yearly decile sorts. This implies that the information it contains is not fully available in real time to market participants, in particular in the first months of a given year. To assure that a trading strategy based on our find-

[10]Results are not sensitive to this specific choice. We have experimented with modifications, such as omitting the one year restriction or such as only considering the first trade in each six months trading period.

ings would actually have been implementable, we have modified the proxy construction by relying on rolling historical values. Specifically, the sorting into deciles for a given day is now based on the raw proxy values over the immediately preceding 250 trading days. Put differently, all required information would have been easily at hand in real time. The appendix provides a transition matrix between the baseline proxy and its modification. In the final pairs trading sample, both variables are highly positively correlated (0.87). And indeed, as the appendix shows, findings broadly carry over. Univariate results are only slightly weaker and multivariate results even slightly stronger than in the baseline case.

4.6 Modified distraction proxies

Our results might partly be driven by the specific design of our distraction proxy. To address potential data mining concerns, we repeat our baseline analysis with ten alternative proxies, which modify the original approach in many dimensions. Specifically, these proxies differ with respect to the type and number of market segments used (49 industries, 100 portfolios sorted on book-to-market and size, 25 portfolios sorted on size and short-term reversal), with respect to the weighting scheme of return shocks (volatility weighting, equal weighting, value weighting, interquartile range), with respect to the model of expected return (market model, four factor model), and with respect to the type of returns used (abnormal returns, raw returns as in Stivers and Sun (2010)). Table 6 gives more information about these proxies and also shows returns by distraction deciles.

Please insert table 6

Table 6 provides a pervasive picture. The difference between decile 10 and decile 1 is statistically highly significant in all cases. Moreover, it keeps its strong economic importance. Finally, results are not confined to the extreme deciles. Instead, in many cases, there seems to be a close to monotone relationship between distraction deciles and returns.

4.7 Limits to arbitrage

In unreported results, we find that only few of the market-level and pair-level control variables in the multivariate analysis are persistently statistically significant. The strongest effect (t-statistic 4.48 in model 4 and 2.60 in model 8 in table 3), however, is found for average idiosyncratic risk, supporting the notion that pairs consisting of difficult to arbitrage firms generate larger returns on average. This cross-sectional finding appears in line with insights from related literature (e.g. Gagnon and Karolyi (2010), Engelberg et al. (2009))). It further raises the question whether similar forces are at work in the time-series and whether these might reduce the importance of our distraction proxy. As a proxy for arbitrage risk in the time-series, we rely on the Chicago Board Options Exchange Market Volatility Index (VIX). The VIX a popular measure of the volatility implied in S&P 500 index options and widely considered a forward-looking measure of overall market uncertainty.[11] Several theories suggest that the amplification of fundamental shocks might impede arbitrageurs from eliminating potential mispricings, or alternatively, that the behavior of constrained arbitrageurs themselves may amplify fundamental shocks (e.g. Long et al. (1990), Shleifer and Vishny (1997), Hong et al. (2011)). The appendix shows results from various regressions of pooled one-month pairs trading returns on the VIX and the distraction proxy over the period January 1990 to December 2008. We both employ raw values of the VIX as well as its yearly decile ranks, computed as for the distraction proxy.

We find that time-varying risk in arbitrage activities does appear to matter, as the VIX often shows up as a significant variable. However, the distraction proxy has a pronounced incremental impact, and remains economically and statistically highly significant.

4.8 Return factor exposure

In the following, we test whether the return difference between high and low distraction days is attributable to loadings on pervasive well-known risk factors. In an attempt to transfer the event-time results of our baseline analysis to calendar time to the extent possible, we extend the maximum holding period from one month (our baseline setup) to six months (as in Gatev et al. (2006)). Doing so works against finding differences across distraction deciles.[12] Separately for pairs opening on high and on low distraction days, we

[11] The VIX is not included in the baseline analysis, as daily data is only available on a daily basis from 1990 on. In the time-series, the correlation of our distraction proxy and the raw VIX (yearly decile ranks of the VIX) is 0.22 (0.31).

[12] Note that this can be inferred from e.g. figure 4: Pairs opening on high (low) distraction days are characterized by a higher (lower) probability of convergence as well as higher (lower) average returns during the first few days after divergence.

construct a time series of daily portfolio returns which are weighted by the cumulative returns of the component pairs. Returns for both time series are then compounded to calculate monthly returns. Finally, we compute the difference between the monthly return for the high distraction portfolio and the monthly return for the low distraction portfolio.

We then regress this time series on a number of well-established risk factors. The first model includes the Fama and French (1993) factors. The second model additionally includes factors designed to capture persistent patterns in return autocorrelation at different time lags. Specifically, we rely on factors for short-term reversal, medium-term momentum, and long-term reversal. The inclusion of these variables is motivated by the contrarian nature of pairs trading, whose success could at least partly be subsumed by these risk premiums. The third model is augmented by the traded liquidity factor constructed in Pástor and Stambaugh (2003). It is intended to control for the strategy's exposure to the aggregate (market-wide) liquidity risk (see e.g. Avramov et al. (2006), Engelberg et al. (2009)). For data availability reasons, this model starts in January 1968, which is six years later than the other models.

Table 7 verifies that, in contrast to standard pairs trading[13], the distraction strategy appears market-neutral. It hardly loads notably on any risk premium. Alphas, however, are persistently statistically significant. In sum, the return difference between pairs opening on high and low distraction days does not seem attributable to standard risk factors.

Please insert table 7

After roughly two weeks, pairs behave similarly. Moreover, pairs not converging within the first days are increasingly unlikely to converge at all (Engelberg et al. (2009) and figure 4). We extend the maximum holding period to six months to obtain smooth time-series of returns on pairs opening on high or low distraction days. As high respectively low distraction days only comprise a tenth of all trading days, simply computing time series returns with a maximum holding period of one month yields a substantial fraction of non-trading days with missing returns. However, relying on distraction quintiles instead of deciles, or simply sticking to the original approach does not change the qualitative nature of our findings.

[13]In line with findings in e.g. Gatev et al. (2006), we find that the monthly return series on traditional pairs trading loads significantly negative on momentum and positive on short-term reversal.

5 Further insights

In this section, we conduct additional tests to further establish the link between time-varying investor distraction and pairs trading profitability.

5.1 Time-series evidence

As a first test, we analyze the role of other investor distraction proxies, which are inspired by previous work. Specifically, we construct four simple alternative dummy variables for limited attention in the time series. Following DellaVigna and Pollet (2009) and Peress (2008), we employ a Friday dummy. Following the idea developed in Hirshleifer et al. (2009) and Peress (2008), we construct a variable based on the number of same-day events competing for investors' attention. To this end, we compute the number of pairs that start trading on a given day. There is considerable variation in each year and no general time trend. The latter is not surprising, as the number of pairs eligible for trading remains constant after the first six months of the sample period, which we exclude. However, the number of opening pairs is significantly positively correlated (0.25) with distraction proxy decile ranks. Therefore, we rely on the residuals from a regression of the logarithmized number of diverging pairs on distraction proxy decile ranks. Finally, we condense this information into a dummy variable. We use the top and bottom quintile to identify days with an unexpectedly large number (dummy=1) or small number (dummy=0) of newly opening pairs. With regard to the third and fourth proxy, we follow Hou et al. (2009) and Karlsson et al. (2009) who provide evidence that investors tend to be less attentive during down market periods. We rely on NBER recession dates and create a dummy variable that takes the value of one if NBER classifies a month as recession. We also create an alternative dummy that is one if the cumulative three year value-weighted market return is negative and zero otherwise.

We then imitate our baseline approach of section 3 by regressing pair returns separately on each of these limited attention dummies (specification 1) as well as additionally on the full set of control variables used in model 8 of table 3, including the distraction proxy decile rank (specification 2). As we have four alternative proxies, two regression specifications, and two return computation schemes, we run 16 regressions in total. Each of the attention dummies is constructed in a way that a positive coefficient is expected. The main findings are presented in panel A of table 8.

Please insert table 8

As predicted, the coefficient is positive in all 16 cases.[14] The persistent positive sign of the coefficients is broadly consistent with the idea of limited attention affecting the relative efficiency of linked assets, although most proxies lack significance once one controls for calendar, industry, market and pair characteristics. In all multivariate regressions, however, the distraction proxy decile rank remains highly statistically and economically significant, suggesting that its explanatory power tends to be greatest.

Given these insights, we explore whether the sensitivity of pairs trading returns to distraction proxy decile ranks becomes even higher once one considers the proxy's possible interaction with the attention proxies inspired by previous work. Specifically, consider the following two extreme scenarios. The first sample consists of all days with distraction decile rank 10 for which, at the same time, the alternative distraction proxy also identifies a distracting situation (e.g. a Friday or a day with a higher than expected number of diverging pairs). In these cases, attention constraints should become particularly binding. The second sample consists of all days with distraction decile rank 1 for which, at the same time, the alternative distraction proxy also identifies a situation in which attention constraints should be less binding. We expect the difference in returns of pairs opening in one of these two extreme situations to be larger than the return difference between decile 10 and decile 1 in our baseline scenario (see tables 2 and 3).

To explore this possibility, we regress one-month pairs returns on the distraction proxy decile rank, the alternative limited attention dummy variable and the interaction effect (as well as a large set of control variables). Again, we have 16 regression specifications in total. Panel B of table 8 shows coefficients obtained for each interaction effect, which are persistently positive as expected. Panel C shows the implied percentage change in return difference between high and low distraction scenarios, as outlined above, when

[14] By far the strongest effect, both statistically and economically, is found for the attention proxy based on NBER recession periods. Note, however, that in contrast to e.g. decile ranks of our novel proxy, recession months are far from being uniformly distributed across sample years.

benchmarked against our baseline findings. As the return difference increases in each case, findings lend further support to the notion that time-varying limited attention is an important explanatory factor for pairs trading profitability.

Next, we study whether pairs trading is particularly profitable immediately before those seven federal holidays for which NYSE has been closed over the whole sample period.[15] We expect investor distraction to be particularly high in these times. This line of reasoning is backed up by DellaVigna and Pollet (2009) who provide evidence that, even on "ordinary Fridays", investors are distracted by the upcoming weekend. It is also motivated by work on holiday effects (e.g. Frieder and Subrahmanyam (2004), Hong and Yu (2009)). We compare returns on pairs that diverge on the last trading day before the holiday with returns on pairs that open on any other day. Specifically, for each year and each holiday separately, we determine whether the mean (median) pre-holiday pairs trading return is larger or smaller than the return over the rest of the year. Panel D of table 8 shows the fraction of years in which pre-holiday pairs trading is more profitable. The fraction is larger than 50% in 11 out of 14 cases, and often also statistically significant. Before Christmas and New Year's Day, pairs trading seems particularly successful. For instance, in about 70% of all sample years, mean and median returns from pairs opening on the last trading day of the year are higher than corresponding returns over the rest of the year. This is substantially driven by a considerably higher than usual fraction of converging pairs. While on average less than 37% of pairs converge within a month, more than 50% do if they diverge immediately before New Year's Day.

5.2 Cross-sectional evidence

We finish our analysis by exploring some cross-sectional implications of our setup. Specifically, we are interested in determining which pairs react most sensitively to changes in investor distraction on the date of pair divergence. Results for the "one day waiting" ("no waiting") scheme are reported in table 9 (the appendix).

Please insert table 9

[15] In chronological order of occurrence, these holidays are: New Year's Day, Washington's Birthday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, Christmas Day.

5.2.1 Alternative Assets

We expect the sensitivity to be positively related to the degree of informational frictions between the constituents of the pair. So far, we have focused on pairs with stocks from different industries. As market participants often specialize along industrial boundaries (e.g. Hong et al. (2007), Menzly and Ozbas (2010)), we expect the sensitivity to be lower for pairs consisting of stocks from the same industry. Therefore, we modify our baseline approach by again identifying the monthly top 100 pairs, but now only considering same-industry pairs. We expect the sensitivity to be even lower if we concentrate on trading pairs of whole industries instead of pairs of stocks. To verify this prediction, we identify the monthly top 20 pairs out of all possible combinations of the 49 value-weighted industries as constituents.[16]

We compute two measures of return sensitivity to time-varying investor distraction. The first measure is the return difference between pairs opening on high distraction days (decile 10) and those opening on low distraction days (decile 1). The second is the coefficient on the distraction proxy decile rank, as obtained from regressions of one-month event-time returns on the proxy such as in table 3. Panel A of table 9 shows findings supporting our line of reasoning. Compared to the baseline pairs universe, the sensitivity to the level of investor distraction is slightly lower for pairs with stocks from the same industry. It is considerably lower, though still significant.

5.2.2 Common Industry Segments

Panel B shows findings from a test similar in spirit. While pairs are based on firms from different industries, it is reasonable to expect that at least some of them operate, at a more disaggregated intra-firm industry level, in some common business segments. The economic link for these pairs will arguably be more visible, which renders it comparatively less likely that shocks in limited attention will cause prices to diverge. To test this hypothesis, we exploit the fact that, starting from 1977, firms have to disclose detailed financial information of any industry segment comprising more than 10% of total consolidated

[16]The lower number of top pairs is chosen as the maximum number of eligible pairs is only 49*48/2=1,176, and thus extremely low when compared to the baseline approach.

yearly sales. To gather this information, we rely on sales data reported in the Compustat fundamentals annual files as well as Compustat segment files, which are then merged with the CRSP data. Several screening procedures are intended to assure data quality.[17] For each pair traded at least once between January 1977 and December 2008, we determine whether firms have at least one business segment in common. We find that about 18% of pairs that satisfy all data requirements share at least one segment, where segments are again defined by the 49 Fama/French industries. We then conduct an analysis analogous to the one described in the previous paragraph. Findings reported in Panel B suggest that the return sensitivity to time-varying investor distraction appears indeed lower for pairs with same industry segments, though the difference is not always significant.

5.2.3 Press coverage

Finally, we explore the role of press coverage. Previous work focussing on other setups has shown that the extent of a firm's media coverage appears directly linked to the speed of information diffusion and thus to price efficiency (e.g. Peress (2008), Huberman (2001)). As press articles catch many investors' attention (e.g. Barber and Odean (2008)), disseminate information to a broad audience, and increase firm visibility (e.g. Fang and Peress (2009)), coverage should help to keep relative prices in line, also and in particular in turbulent moments. Thus, first, a highly covered pair should be less profitable than a pair that is widely neglected by the press. Second, the highly covered pair should also react less sensitive to distracting overall market situations.

To explore these predictions, we rely on the Dow Jones News Service (DJNS) database as accessible via factiva. Due to its comprehensive coverage, this database has widely been used in previous studies, and argued to be "the best approximation of public news for traders" (Chan (2003), p. 230). For each firm that meets our data requirements on pairs trading (see section 2.2) at some point after 1990, we collect the yearly number of news

[17] See e.g. Berger and Ofek (1995) or Cohen and Lou (2011) for more detailed information about the relevant segment reporting regulations. We loosely follow these studies in imposing various criteria firms have to meet in order to enter this test. First, for a given year, firms are required to have data both in the Compustat fundamentals annual as well as segment file. Second, the sum of reported segment sales must be within 1% of total sales. Third, we exclude segment eliminations or segments with missing SIC codes.The Compustat segment file provides four-digit SIC codes for industry segments, which we transform in the 49 Fama/French industries.

articles between 1991 and 2008.[18] As this number is strongly positively related to firm size (e.g. Fang and Peress (2009)), we perform yearly regressions of ln (1+number of news) on ln(average market capitalization). We use the yearly top and bottom quintile of the resulting yearly residuals to identify firms with particularly high or low DJNS coverage. Finally, we define a pair as being highly (lowly) covered, if both of its components are firms with high (low) coverage in the year the divergence occurs.

Panel C of table 9 compares return characteristics for both types of pairs. Our predictions largely prove true. The one-month return difference between pairs receiving disproportionately much coverage and those widely neglected reaches at least 80 basis points. In fact, trading highly covered pairs turns out to be completely unprofitable, whereas trading lowly covered pairs is considerably more profitable than trading the average pair in 1991 to 2008. Moreover, as predicted, the pair's sensitivity to changes in the level of investor distraction is statistically and economically significantly higher for lowly covered pairs.

6 Conclusion

Understanding how markets impound information into stock prices is one of the central concerns of financial economics. We provide new insights by analyzing how the price formation of linked stocks, as identified by pairs trading techniques, is affected by investor distraction, as quantified with a novel proxy. Our results lend support to the notion that the relative efficiency of linked assets might not be stable over time, but be affected by short-term investor attention shifts. Pairs opening on high distraction days, on which exceptional market circumstances force investors to concentrate on understanding the big picture, are much more profitable than pairs opening on low distraction days. This key finding does not only hold in the US market, but also in all eight major international stock markets we study. It is economically meaningful, statistically significant, and survives a number of robustness checks. Several further tests are also consistent with the idea of investor attention constraints being an important source of friction in financial markets.

[18] Tetlock (2010) argues that DJNS articles before November 1996 might suffer from some measurement error and survivorship bias towards larger firms. Therefore, we have replicated the following analysis also for the subperiod 1997 to 2008. The qualitative nature of our findings does not change.

References

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

Andrade, S. C., V. di Pietro, and M. S. Seasholes, 2005, "Understanding the profitability of pairs trading," *Unpublished working paper, U.C. Berkeley, Northwestern University.*

Avramov, D., T. Chorida, and A. Goyal, 2006, "Liquidity and autocorrelations in individual stock returns," *Journal of Finance*, 61, 2365–2394.

Barber, B. M., and T. Odean, 2008, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *Review of Financial Studies*, 21, 785–818.

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

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

Berger, P. G., and E. Ofek, 1995, "Diversification's effect on firm value," *Journal of Financial Economics*, 37, 39–65.

Boyer, B. H., 2011, "Style related comovement: Fundamentals or labels?," *Journal of Finance*, 66, 307–332.

Brandt, M., A. Brav, J. R. Graham, and A. Kumar, 2010, "Idiosyncratic volatility puzzle: Time trend or speculative episodes?," *Review of Financial Studies*, 23, 863–899.

Chan, L. K., J. Lakonishok, and B. Swaminathan, 2007, "Industry classifications and return comovement," *Financial Analysts Journal*, 63, 56–70.

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

Chay, J., and C. A. Trzcinka, 1999, "Managerial performance and the cross-sectional pricing of closed-end funds," *Journal of Financial Economics*, 52, 379–408.

Chen, H., S. Chen, and F. Li, 2010, "Empirical investigation of an equity pairs trading strategy," *Unpublished working paper, University of British Columbia, University of Michigan.*

Cherkes, M., J. Sagi, and R. Stanton, 2009, "A liquidity-based theory of closed-end funds," *Review of Financial Studies*, 22, 257–297.

Chordia, T., A. Goyal, and Q. Tong, 2011, "Pairwise correlations," *Unpublished working paper, Emory University, University of Lausanne, Singapore Management University.*

Cohen, L., and A. Frazzini, 2008, "Economic links and predictable returns," *Journal of Finance*, 63, 1977–2011.

Cohen, L., and D. Lou, 2011, "Complicated Firms," *Journal of Financial Economincs, forthcoming.*

Cooper, M. J., O. Dimitrov, and R. Rau, 2001, "A rose.com by any name," *Journal of Finance*, 56, 2371–2388.

Cooper, M. J., H. Gulen, and P. R. Rau, 2005, "Changing names with style: Mutual fund name changes and their effect on fund flows," *Journal of Finance*, 60, 2825–2858.

DellaVigna, S., and J. M. Pollet, 2009, "Investor inattention and Friday earnings announcements," *Journal of Finance*, 64, 709–749.

Do, B., and R. Faff, 2010, "Does simple pairs trading still work?," *Financial Analysts Journal*, 66, 83–95.

Do, B. H., and R. W. Faff, 2011, "Are pairs trading profits robust to trading costs?," *Working paper, Monash University.*

Engelberg, J., P. Gao, and R. Jagannathan, 2009, "An anatomy of pairs trading: The role of idiosyncratic news, common information and liquidity," *Unpublished working paper, University of North Carolina, University of Notre Dame, Northwestern University.*

Fama, E. F., and K. R. French, 1993, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— , 1997, "Industry costs of equity," *Journal of Financial Economics*, 43, 153–193.

Fang, L., and J. Peress, 2009, "Media coverage and the cross-section of stock returns," *Journal of Finance*, 64, 2023–2052.

Fink, J., K. E. Fink, G. Grullon, and J. P. Weston, 2010, "What drove the increase in idiosyncratic volatility during the internet boom?," *Journal of Financial and Quantitative Analysis*, 45, 1253–1278.

Frieder, L., and A. Subrahmanyam, 2004, "Non-secular regularities in returns and volume," *Financial Analysts Journal*, 60, 29–34.

Froot, K., and E. Dabora, 1999, "How are stock prices affected by the location of trade?," *Journal of Financial Economics*, 53, 189–216.

Gagnon, L., and G. A. Karolyi, 2010, "Multi-market trading and arbitrage," *Journal of Financial Economics*, 97, 53–80.

Gatev, E. G., W. N. Goetzmann, and K. G. Rouwenhorst, 2006, "Pairs trading: Performance of a relative-value arbitrage rule," *Review of Financial Studies*, 19, 797–827.

Green, T. C., and B.-H. Hwang, 2009, "Price-based return comovement," *Journal of Financial Economics*, 93, 37–50.

Greenwood, R., 2008, "Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights," *Review of Financial Studies*, 21, 1153–1186.

Grossman, S. J., and J. E. Stiglitz, 1980, "On the impossibility of informationally efficient markets," *American Economic Review*, 70, 393–408.

Hirshleifer, D., S. S. Lim, and S. H. Teoh, 2009, "Driven to distraction: Extraneous events and underreaction to earnings news," *Journal of Finance*, 64, 2289–2325.

Hong, H. G., J. D. Kubik, and T. Fishman, 2011, "Do Arbitrageurs Amplify Economic Shocks?," *forthcoming Journal of Financial Economics.*

Hong, H. G., T. Lim, and J. C. Stein, 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies," *Journal of Finance*, 55, 265–295.

Hong, H. G., W. Torous, and R. Valkanov, 2007, "Do industries lead stock markets?," *Journal of Financial Economics*, 83, 367–396.

Hong, H. G., and J. Yu, 2009, "Gone fishin': Seasonality in trading activity and asset prices," *Journal of Financial Markets*, 12, 672–702.

Hou, K., 2007, "Industry information diffusion and the lead-lag effect in stock returns," *Review of Financial Studies*, 20, 1113–1138.

Hou, K., L. Peng, and W. Xiong, 2009, "A tale of two anomalies: The implications of investor attention for price and earnings momentum," *Unpublished working paper, Ohio State University, City University of New York, Princeton University.*

Huberman, G., 2001, "Familiarity breeds investment," *Review of Financial Studies*, 14, 659–680.

Jong, A. D., L. Rosenthal, and M. A. V. Dijk, 2009, "The risk and return of arbitrage in dual-listed companies," *Review of Finance*, 13, 495–520.

Kahneman, D., 1973, "Attention and Effort," *Prentice Hall, New Jersey.*

Karlsson, N., G. Loewenstein, and D. Seppi, 2009, "The ostrich effect: Selective attention to information," *Journal of Risk and Uncertainty*, 38, 95–115.

Lamont, O. A., and R. H. Thaler, 2003, "Can the market add and subtract? Mispricing in tech stock carve-outs," *Journal of Political Economy*, 111, 227–268.

Lee, C. M., A. Shleifer, and R. H. Thaler, 1991, "Investor sentiment and the closed-end fund puzzle," *Journal of Finance*, 46, 75–109.

Lo, A. W., and C. A. MacKinlay, 1990, "When are Contrarian Profits due to Stock Market Overreaction?," *Review of Financial Studies*, 3, 175–205.

Long, J. B. D., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1990, "Noise trader risk in financial markets," *Journal of Political Economy*, 98, 703738.

Menzly, L., and O. Ozbas, 2010, "Market segmentation and cross-predictability of returns," *Journal of Finance*, 65, 1555–1580.

Mitchell, M., T. Pulvino, and E. Stafford, 2002, "Limited arbitrage in equity markets," *Journal of Finance*, 57, 551–584.

Pástor, L., and R. F. Stambaugh, 2003, "Liquidity risk and expected returns," *Journal of Political Economy*, 111, 642–685.

Peng, L., 2005, "Learning with information capacity constraints," *Journal of Financial and Quantitative Analysis*, 40, 307–329.

Peng, L., and W. Xiong, 2006, "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, 80, 563–602.

Peress, J., 2008, "Media coverage and investors' attention to earnings announcements," *Unpublished working paper, Insead.*

Pirinsky, C., and Q. Wang, 2006, "Does corporate headquarters location matter for stock returns?," *Journal of Finance*, 61, 1991–2015.

Pontiff, J., 1995, "Closed-end fund premia and returns implications for financial market equilibrium," *Journal of Financial Economics*, 37, 341–370.

——— , 2006, "Costly arbitrage and the myth of idiosyncratic risk," *Journal of Accounting and Economics*, 42, 35–52.

Rosenthal, L., and C. Young, 1990, "The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever N.V./PLC," *Journal of Financial Economics*, 26, 123–141.

Schultz, P., and S. Shive, 2010, "Mispricing of dual-class shares: Profit opportunities, arbitrage, and trading," *Journal of Financial Economics*, 98, 524–549.

Scruggs, J. T., 2007, "Noise trader risk: Evidence from the Siamese twins," *Journal of Financial Markets*, 10, 76–105.

Shleifer, A., and R. W. Vishny, 1997, "The limits of arbitrage," *Journal of Finance*, 52, 35–55.

Smith, B. F., and B. Amoako-Adu, 1995, "Relative prices of dual class shares," *Journal of Financial and Quantitative Analysis*, 30, 223–239.

Stivers, C., and L. Sun, 2010, "Cross-sectional return dispersion and time variation in value and momentum premiums," *Journal of Financial and Quantitative Analysis*, 45, 987–1014.

Tetlock, P. C., 2010, "Does public financial news resolve asymmetric information?," *Review of Financial Studies*, 23, 3520–3557.

Zingales, L., 1995, "What determines the value of corporate votes?," *Quarterly Journal of Economics*, 110, 1047–1073.

Figure 1: One Month Return on Pairs by Distraction Proxy Decile Ranks in Event Time

This figure shows the average one month event-time return on US long-short stock pairs sorted by the distraction proxy decile rank on the day of pair divergence. Findings are based on more than 100,000 round-trip trades between January 1962 and December 2008 (see sections 2 and 3 for details). After divergence, pairs are hold for up to one month. If they do not converge again before this cut-off date, positions are offset. If pairs converge before this cut-off date, the proceeds are hold in cash until the full month has passed. As any cash-flow before the cut-off date is positive by construction, this is a conservative approach. For the upper bound of the return to the strategy, we compute returns on zero-cost portfolios between the day of divergence (closing price) and the day of convergence (closing price). We refer to this scheme as "no waiting". For the lower bound, we employ a more conservative return computation approach as discussed in Gatev et al. (2006). Specifically, we skip one day after the divergence and add one day following the crossing of the prices. We refer to this scheme as "one day waiting".

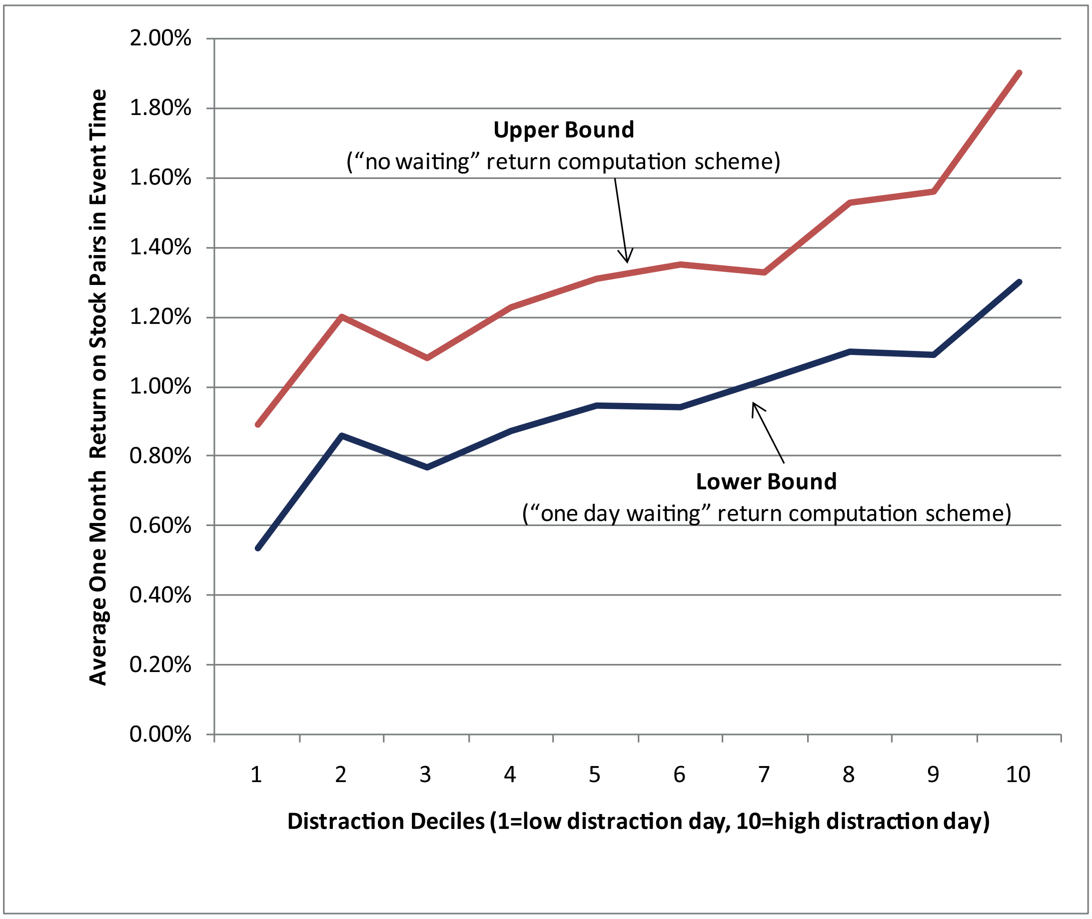
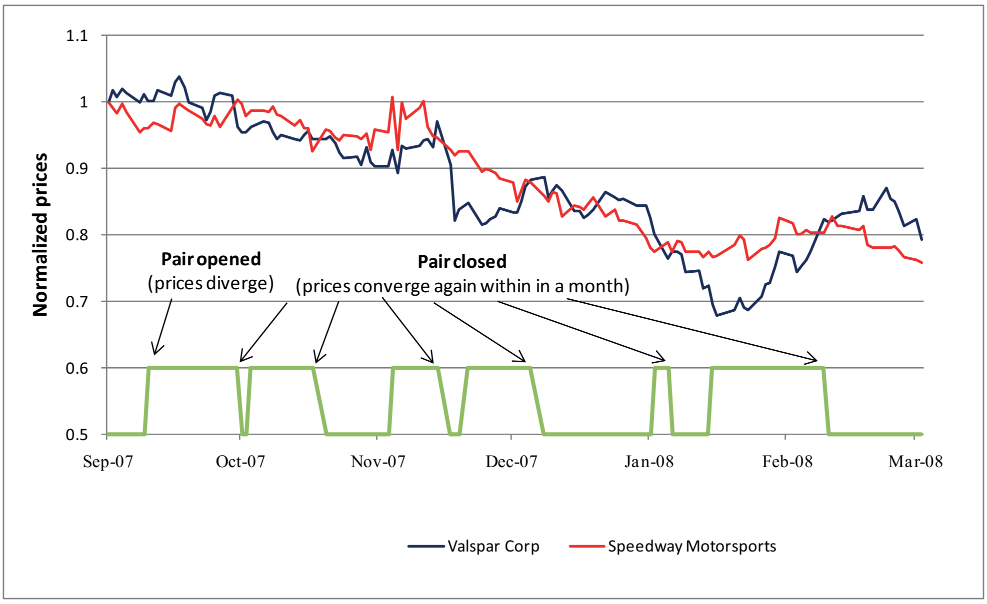


Figure 2: Illustration of Pairs Trading Process

This figure illustrates the trading process with two examples. Each month, the 100 pairs with minimum distance between normalized 12 month daily return indices are selected. They are then eligible for trading in the immediately following six months evaluation period. At the beginning of this period, prices are set to equal unity. If the spread between the cumulative return series of the two stocks exceeds two historical standard deviations (as estimated in the estimation period), we go one dollar long in the relatively underpriced stock, which is financed by short-selling the relatively overpriced stock. The self-financing pair is then hold for up to one month. If prices convergence before this cut-off date, the trade is closed with a gain. If prices do not convergence within a month, positions are offset, which, if prices diverge even further, results in a loss. A pair may open several more times during the trading period. In this case, the trading process is repeated as outlined above.



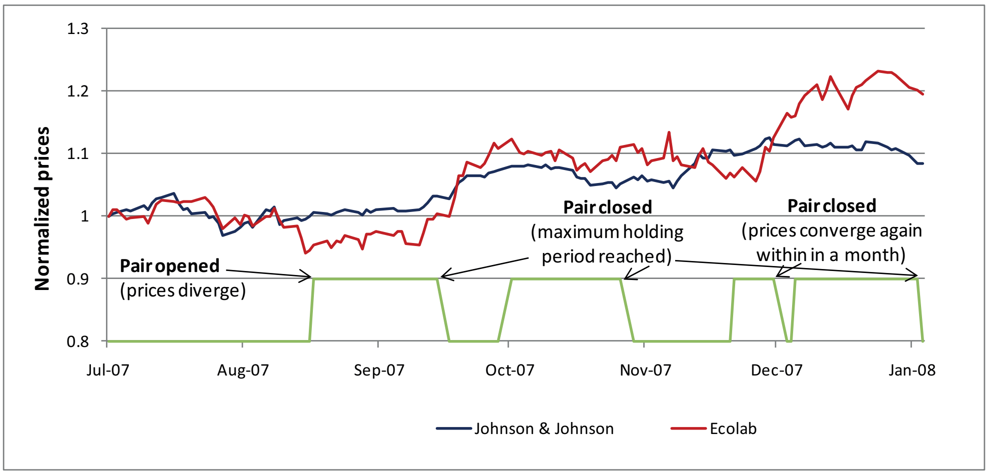


Figure 3: Time Series Characteristics of Investor Distraction Proxy

The upper graph shows the daily development of the raw distraction proxy from January 1962 to December 2008. For its construction, a three-step procedure is employed. First, for January 1960 to December 2008, we compute daily value-weighted returns for the 49 Fama and French (1997) industries. Second, we construct daily return shocks defined as the absolute difference between the actual industry return and its expected return as implied by an OLS market model. Parameter estimates are obtained from rolling time-series regressions based on daily return data over the previous year. Third, shocks are condensed into a single ratio. To this end, industry shocks are weighted by the inverse of the volatility of their shock variable over the previous year. The lower graph shows the monthly number of high and low distraction days based on yearly sorts of the raw distraction proxy in decile ranks. Days with decile rank 10 (1) are referred to as high (low) distraction days. Spearman rank order correlation coefficients between these decile ranks and the rank order of market-level variables based on daily data are as follows. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Market excess return	Small firm factor	Value/growth factor	Momentum factor	Short-term reversal factor	Equal-weighted market return
0.0544***	-0.074***	-0.0231**	-0.1062***	-0.0004	0.0399***
Squared market return	Value-weighted turnover	Equal-weighted turnover	Rolling 10 day volatility	Turnover dispersion	Return dispersion
0.2109***	0.1350***	0.1030***	0.2217***	-0.0848***	-0.0352***

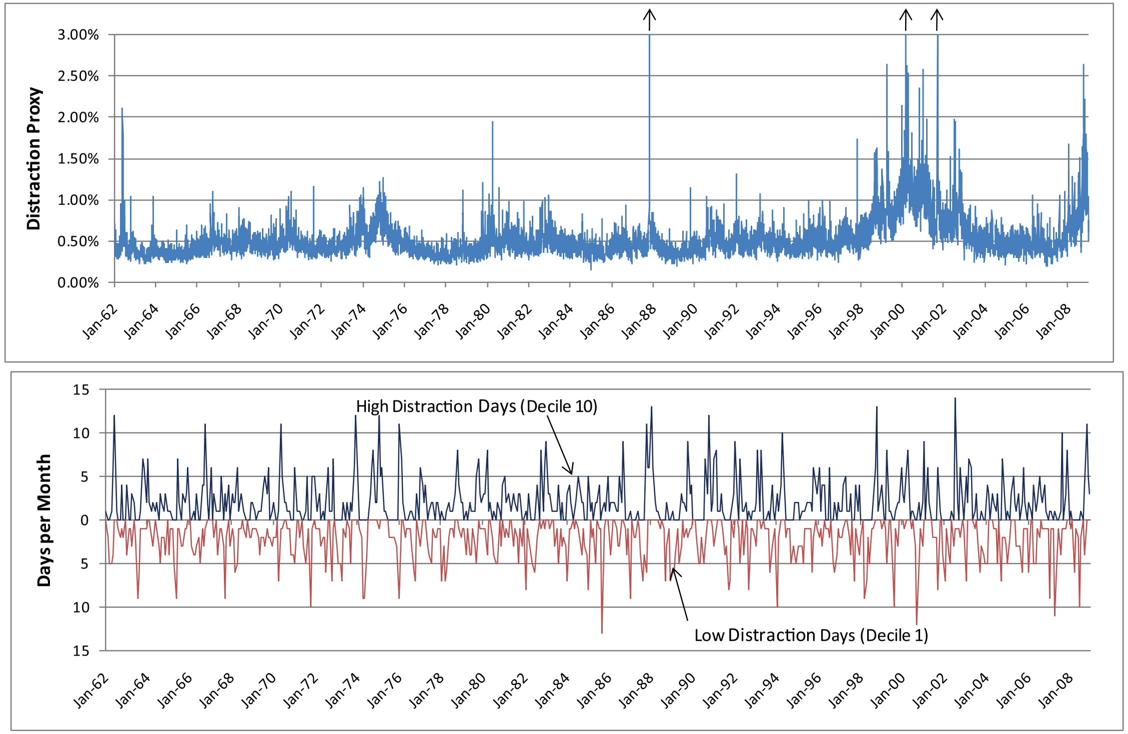
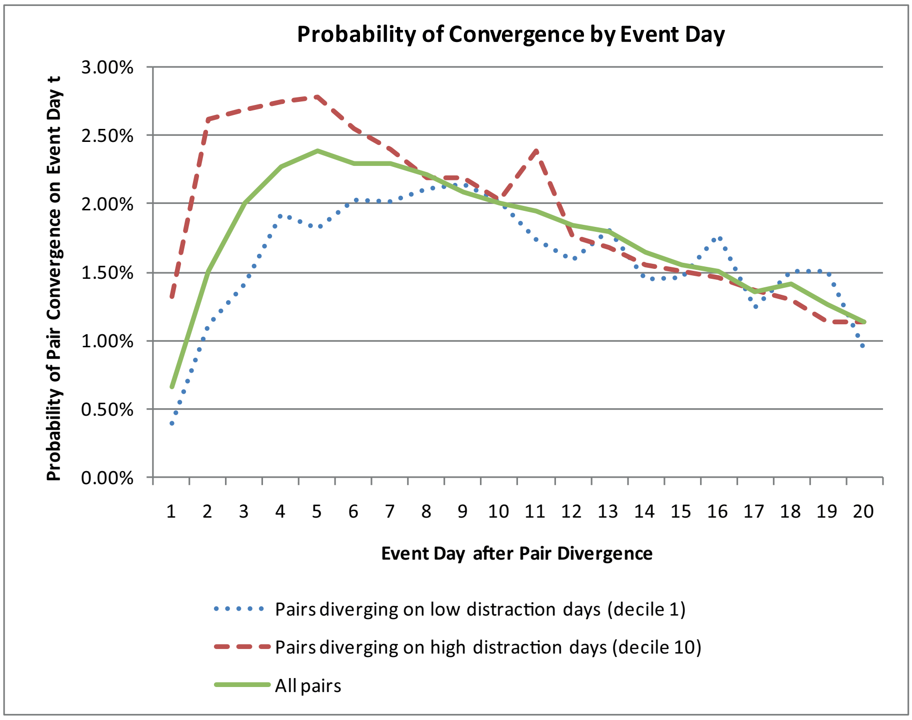


Figure 4: Probability of Convergence and Average Daily Return by Event Day

The upper graph shows the empirical probability of US stock pairs converging on a given event day after divergence. See section 2.2 for a definition of divergence and convergence. The lower graph shows the average daily return of open pairs in event-time. Both figures are based on more than 100,000 round-trip trades between January 1962 and December 2008 (see sections 2 and 3 for details).



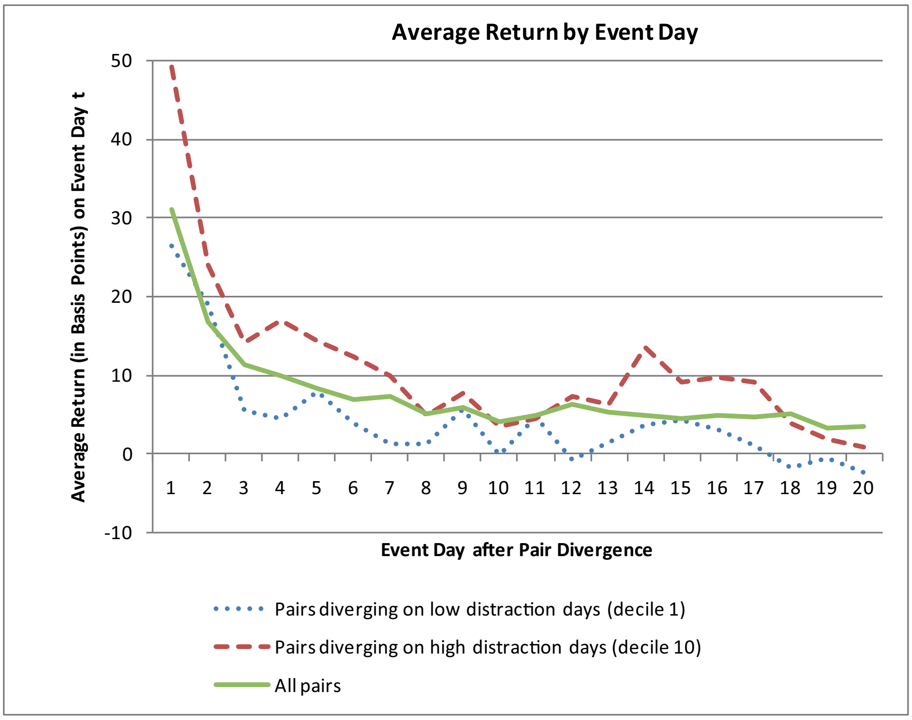


Table 1: Descriptive Statistics for Stock and Pair Characteristics by Distraction Deciles

In panel A, *NYSE/AMEX macap decile* refers to the firm's market capitalization decile rank computed at the beginning of the pair's six month trading period. *Amihud illiquidity ratio* is computed as the average of a stock's absolute daily return divided by its total daily trading volume in million dollars. The estimation period for the illiquidity ratio, for average daily turnover as well as for idiosyncratic risk is the 12 month period ending at the beginning of a pair's trading period. *Idiosyncratic risk* is computed as the standard deviation of the residual obtained from time series regressions of a stock's daily return on factors for the market premium, size, value and momentum. *Maximum industry weight* denotes the largest fraction of sample firms belonging to a specific industry group (out of the 49 Fama/French industries). *Industry concentration* is computed as the sum of squared industry weights. In panel B, the first four rows report within-pair differences of stock characteristics, which are computed as in panel A. The last column reports differences in mean characteristics between decile 10 and decile 1. Standard errors are adjusted for heteroskedasticity and clustered by day of pair divergence. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Variable		All	Distraction Decile 1	Distraction Decile 10	10-1 (Mean)
Panel A: Firm Characteristics by Distraction Deciles					
NYSE/AMEX macap decile	Mean	8.86	8.85	8.86	0.0189
	Median	9	9	9	
Amihud illiquidity ratio	Mean	0.0578	0.0567	0.0566	-0.0002
	Median	0.01	0.0107	0.0094	
Average daily turnover	Mean	0.16%	0.16%	0.17%	0.008%
	Median	0.11%	0.11%	0.11%	
Idiosyncratic risk	Mean	1.12%	1.10%	1.13%	0.032%**
	Median	1.06%	1.05%	1.07%	
No. Analysts (since 1980)	Mean	10.41	10.56	10.21	-0.34
	Median	9	9	9	
No. industry groups		49	49	49	
Maximum industry weight	Fraction	29.14%	29.20%	29.53%	
	Industry	Utilities (across all deciles)			
Industry concentration		0.114	0.115	0.115	
Panel B: Pair Characteristics by Distraction Deciles					
Macap decile difference	Mean	1.25	1.25	1.24	-0.0052
	Median	1	1	1	
Average daily turnover difference	Mean	0.072%	0.072%	0.073%	-0.00%
	Median	0.040%	0.040%	0.039%	
Amihud illiquidity ratio difference	Mean	0.063	0.061	0.061	-0.000
	Median	0.013	0.0130	0.0130	
Idiosyncratic risk difference	Mean	0.248%	0.237%	0.252%	0.015%***
	Median	0.20%	0.19%	0.20%	
Cumulative price difference upon divergence	Mean	6.68%	6.44%	7.04%	0.60%***
	Median	6.28%	6.05%	6.58%	
Return difference at day of divergence	Mean	2.84%	2.38%	3.59%	1.21%***
	Median	2.29%	1.93%	2.89%	
No. industry group combinations		931	623	697	
Maximum industry group weight	Fraction	15.89%	16.04%	15.13%	
	Industries	Utilities/Communication (across all deciles)			
Industry group concentration		0.039	0.041	0.037	
No. round-trip trades		104,125	8,222	14,199	5,977

Table 2: One-month Pairs Trading Returns by Distraction Deciles

This table reports event-time one-month returns on zero-cost portfolios of US stock pairs sorted by distraction proxy decile ranks as observed on the day of divergence. Breakpoints for the deciles are determined separately for each year. Calculations are based on daily data from January 1962 to December 2008. In Panel A (B), trading positions in each pair are initiated on the day of divergence (on the day following the convergence) and liquidated on the day of convergence (on the day following the convergence). *Fraction of convergence* refers to the percentage of pairs that converge within the month following the divergence. *Return if convergence* (*Return if no convergence*) refers to the average return generated, thereby conditioning on pairs that do (do not) converge within the month following divergence. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Distraction Decile	All	1	2	3	4	5	6	7	8	9	10	10-1
Panel A: One Day Waiting, Full Sample Period (1962-2008)												
Return on pairs	0.00973***	0.00534***	0.00859***	0.00770***	0.00873***	0.00946***	0.00943***	0.0102***	0.0110***	0.0109***	0.0130***	**0.0077*****
	(0.000319)	(0.00103)	(0.000956)	(0.000976)	(0.000988)	(0.00100)	(0.00104)	(0.00106)	(0.000917)	(0.000971)	(0.000990)	**(0.00143)**
% of convergence	0.363	0.334	0.343	0.343	0.347	0.362	0.361	0.358	0.376	0.381	0.397	**0.0630*****
Return if	0.0642***	0.0614***	0.0629***	0.0633***	0.0624***	0.0633***	0.0637***	0.0653***	0.0639***	0.0651***	0.0671***	**0.0056*****
convergence	(0.000296)	(0.00101)	(0.000874)	(0.000869)	(0.000920)	(0.00103)	(0.000908)	(0.000910)	(0.000851)	(0.000893)	(0.000909)	**(0.00136)**
Return if	-0.0208***	-0.0224***	-0.0194***	-0.0211***	-0.0195***	-0.0208***	-0.0209***	-0.0201***	-0.0202***	-0.0217***	-0.0213***	**0.00112**
no convergence	(0.000343)	(0.000984)	(0.00102)	(0.00106)	(0.00108)	(0.00102)	(0.00115)	(0.00113)	(0.000978)	(0.00108)	(0.00114)	**(0.0015)**
No. of trades	103,386	8,187	8,679	9,146	9,398	10,048	10,079	10,595	11,019	12,224	14,011	**5,824**
Panel B: No Waiting, Full Sample Period (1962-2008)												
Return on pairs	0.0138***	0.00889***	0.0120***	0.0108***	0.0123***	0.0131***	0.0135***	0.0133***	0.0153***	0.0156***	0.0190***	**0.0101*****
	(0.000341)	(0.00108)	(0.000986)	(0.00103)	(0.00105)	(0.00103)	(0.00106)	(0.00117)	(0.00103)	(0.00104)	(0.00106)	**(0.00151)**
% of convergence	0.363***	0.334	0.343	0.343	0.347	0.362	0.361	0.358	0.376	0.381	0.397	**0.0630*****
Return if	0.0756***	0.0713***	0.0725***	0.0729***	0.0739***	0.0742***	0.0744***	0.0757***	0.0755***	0.0774***	0.0819***	**0.0106*****
convergence	(0.000253)	(0.000735)	(0.000636)	(0.000678)	(0.000723)	(0.000775)	(0.000669)	(0.000707)	(0.000664)	(0.000726)	(0.000937)	**(0.00119)**
Return if	-0.0214***	-0.0223***	-0.0195***	-0.0215***	-0.0204***	-0.0216***	-0.0209***	-0.0216***	-0.0210***	-0.0225***	-0.0223***	**0.0001**
no convergence	(0.000351)	(0.00103)	(0.00105)	(0.00109)	(0.00112)	(0.00105)	(0.00118)	(0.00124)	(0.000964)	(0.00110)	(0.00112)	**(0.00152)**
No. of trades	104,125	8,222	8,738	9,179	9,436	10,094	10,122	10,657	11,146	12,332	14,199	**5,977**

Table 3: Multivariate Analysis: Investor Distraction and Returns on Pairs Trading

This table displays findings from pooled multivariate regressions of the one-month return on zero-cost US stock pairs on a proxy for investor distraction and up to three sets of control variables. The proxy for investor distraction is the *Distraction Proxy Decile Rank* (specifications 1-4) or a *High/Low Distraction Dummy* (specifications 5-8), which is zero for low distraction days (decile 1) and one for high distraction days (decile 10). Pairs trading returns are computed under the conservative "one day waiting" return scheme. The first set of explaining variables controls for calendar and industry effects (indicator variables for year, month, day of week, and pair industry group combinations). The second set controls for market-level conditions on the day of divergence (market return, squared market return, market turnover, 10 day rolling volatility, factors for daily return premia on size, value, momentum and short-term reversal). The third set controls for a number of pair characteristics computed as outlined in table 1 (average firm market capitalization decile rank, ln (average pre-event turnover), ln (average pre-event Amihud illiquidity ratio), average idiosyncratic risk, within-pair differences in these variables, return difference attributable to the day of divergence, ln (average turnover on day of divergence) and ln (difference in turnover on day of divergence)). Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Model specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample period	1/1962-12/2008	1/1962-12/2008	7/1963-12/2008	7/1963-12/2008	1/1962-12/2008	1/1962-12/2008	7/1963-12/2008	7/1963-12/2008
Observations	103,386	103,386	100,426	99,673	22,198	22,198	21,579	21,281
Adjusted R^2	0.08%	2.62%	2.74%	2.83%	0.30%	5.33%	5.58%	5.74%
Distraction Proxy Decile Rank	0.00065***	0.00065***	0.00040***	0.00039***				
	(0.0001099)	(0.0001049)	(0.0001172)	(0.0001183)				
High/Low Distraction Dummy					0.00770***	0.00734***	0.00542***	0.00557***
					(0.0014263)	(0.0014141)	(0.0017108)	(0.0017519)
Controls for calendar and industry effects	no	yes	yes	yes	no	yes	yes	yes
Controls for market-level conditions	no	no	yes	yes	no	no	yes	yes
Controls for pair characteristics	no	no	no	yes	no	no	no	yes

Table 4: Pairs Trading by Distraction Deciles: International Evidence

This table reports profits from pairs trading in international stock markets. The methodology is the same as in the baseline analysis for the US market (see table 2). Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

	Japan	UK	France	Germany	Switzerland	Italy	Netherlands	Hongkong
Panel A: One Day Waiting, Full Sample Period (1962-2008)								
Time Period	1/1995-12/2009	1/1995-12/2009	1/1996-12/2009	1/1996-12/2009	6/1997-12/2009	6/1995-12/2009	1/1995-12/2009	1/1995-12/2009
Observations	36,992	27,185	25,833	26,006	16,841	25,596	21,320	20,894
Panel A: One-month Pairs Trading Abnormal Returns, One Day Waiting								
Pairs Trading Returns: All	0.0103***	0.0073***	0.0111***	0.0109***	0.0054***	0.0048***	0.0082***	0.0100***
	(0.000605)	(0.000762)	(0.000852)	(0.00105)	(0.00130)	(0.00117)	(0.00111)	(0.00223)
Pairs Trading Returns: Decile 1	0.0119***	0.00278	0.0079***	0.0061*	-0.0065	0.0032	0.0037	-0.0099
	(0.00171)	(0.00200)	(0.00270)	(0.00331)	(0.00427)	(0.00350)	(0.00313)	(0.00679)
Pairs Trading Returns: Decile 10	0.0138***	0.0142***	0.0173***	0.0204***	0.0171***	0.0066	0.0203***	0.0263***
	(0.00197)	(0.00245)	(0.00273)	(0.00324)	(0.00510)	(0.00459)	(0.00342)	(0.00601)
Diff 10-1	**0.0019**	**0.0114*** **	**0.0094** **	**0.0143*** **	**0.0237*** **	**0.0034**	**0.0166*** **	**0.0362*** **
	(0.00260)	**(0.00316)**	**(0.00383)**	**(0.00463)**	**(0.00664)**	**(0.00577)**	**(0.00463)**	**(0.00906)**
Panel B: One-month Pairs Trading Abnormal Returns, No Waiting								
Pairs Trading Returns: All	0.0207***	0.0095***	0.0160***	0.0168***	0.0087***	0.0077***	0.0154***	0.0145***
	(0.000685)	(0.000839)	(0.000888)	(0.00111)	(0.00135)	(0.00124)	(0.00125)	(0.00214)
Pairs Trading Returns: Decile 1	0.0216***	0.0033	0.0111***	0.0096***	-0.0035	0.0056	0.0103***	-0.0076
	(0.00189)	(0.00220)	(0.000852)	(0.00331)	(0.00434)	(0.00359)	(0.00373)	(0.00676)
Pairs Trading Returns: Decile 10	0.0270***	0.0197***	0.0227***	0.0301***	0.0219***	0.0117**	0.0315***	0.0334***
	(0.00209)	(0.00259)	(0.00279)	(0.00345)	(0.00520)	(0.00491)	(0.00386)	(0.00561)
Diff 10-1	**0.0055***	**0.0164*** **	**0.0105*** **	**0.0204*** **	**0.0253*** **	**0.0062**	**0.0212*** **	**0.0409*** **
	(0.00282)	**(0.00340)**	**(0.00398)**	**(0.00477)**	**(0.00677)**	**(0.00608)**	**(0.00536)**	**(0.00878)**
Panel C: Fraction of Convergence								
% of Convergence: All	0.448***	0.289***	0.296***	0.296***	0.195***	0.267***	0.232***	0.231***
	(0.00423)	(0.00455)	(0.00411)	(0.00430)	(0.00473)	(0.00417)	(0.00536)	(0.00549)
% of Convergence: Decile 1	0.438***	0.256***	0.269***	0.270***	0.148***	0.250***	0.209***	0.159***
	(0.0127)	(0.0139)	(0.0144)	(0.0132)	(0.0137)	(0.0140)	(0.0178)	(0.0128)
% of Convergence: Decile 10	0.491***	0.341***	0.328***	0.322***	0.244***	0.300***	0.301***	0.298***
	(0.0127)	(0.0147)	(0.0116)	(0.0147)	(0.0160)	(0.0132)	(0.0178)	(0.0160)
Diff 10-1	**0.0537*** **	**0.0850*** **	**0.0589*** **	**0.0514*** **	**0.0965*** **	**0.0500*** **	**0.0920*** **	**0.139*** **
	(0.0179)	**(0.0202)**	**(0.0185)**	**(0.0197)**	**(0.0211)**	**(0.0192)**	**(0.0252)**	**(0.0205)**

Table 5: Robustness Checks

This table presents results from various robustness checks. For brevity, we only report results obtained under the more conservative "one day waiting" return computation scheme. Panel A displays subperiod results from the baseline approach. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence. In Panel B, the way the monthly top 100 pairs are identified is modified. *Excluding utility firms* means we do not consider any pair with at least one firm belonging to Fama/French (1997) industry group 31. *Only different firms* means we do not select a pair if at least one the firms is already a component of any higher-ranked pair. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence. Panel C reports results from Fama-MacBeth-type regressions. We first estimate yearly pooled cross-sectional regressions of one-month pairs returns on distraction decile ranks and then use the time-series of resulting coefficients to assess the statistical significance of the distraction proxy. Newey-West-adjusted standard errors are reported in parentheses. Panel D shows results for *same pairs*, i.e. a subsample of pairs that diverge both at least once on a high distraction day and on a low distraction days. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Subperiod Analysis				
	Distraction Decile: All	Distraction Decile: 1	Distraction Decile: 10	Difference: 10-1
Subperiod: 1962-1977	0.0149***	0.0104***	0.0177***	**0.0073*** **
	(0.000511)	(0.00173)	(0.00153)	**(0.00231)**
Subperiod: 1978-1993	0.0109***	0.00706***	0.0129***	**0.0059** **
	(0.000528)	(0.00175)	(0.00171)	**(0.00245)**
Subperiod: 1994-2008	0.0024***	-0.0025	0.0079***	**0.0104*** **
	(0.000598)	(0.00174)	(0.00187)	**(0.00255)**
Panel B: Variations in the Data Set (Monthly Top 100 Pairs)				
	Distraction Decile: All	Distraction Decile: 1	Distraction Decile: 10	Difference: 10-1
Excluding utility firms	0.01019***	0.00588***	0.01347***	**0.0076*** **
	(0.00031)	(0.00102)	(0.00097)	**(0.00102)**
Only different firms	0.01080***	0.00785***	0.01435***	**0.0065*** **
	(0.00029)	(0.00091)	(0.00087)	**(0.0013)**
Panel C: Alternative Regression Approach				
Coefficient on distraction decile rank		0.00057***		
		(0.0001509)		
Panel D: Limitation to Firms that Diverge both at Least Once on High and Low Distraction Days				
	Distraction Decile: All	Distraction Decile: 1	Distraction Decile: 10	Difference: 10-1
Same pairs	0.0202***	0.0140***	0.0213***	**0.0073*** **
	(0.000611)	(0.00178)	(0.00160)	**(0.00239)**

Table 6: Alternative Distraction Proxies: One-month Pairs Trading Returns by Distraction Deciles

This table reports one-month event-time returns on zero-cost portfolios of stock pairs sorted by distraction proxy deciles as observed on the day of the divergence. Breakpoints for the distraction deciles of each proxy are determined separately for each year. In Panel A, factor loadings with respect to the model of expected returns are estimated from time-series regressions based on daily data over the previous year. In the value-weighted case, abnormal industry returns are weighted by the relative market capitalization of industries. The interquartile range is computed as the 75th percentile of the cross-section of daily abnormal industry returns minus the 25th percentile. Returns on the 100 portfolios formed on book-to-market and size as well as on the 25 portfolios formed on size and short-term reversal are taken from Kenneth French's data library. In Panel B, standard errors (in parentheses) are adjusted for heteroskedasticity and clustered by day of pair divergence. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Description of Alternative Distraction Proxies

Proxy Label	Proxy availability	Computation Scheme	Expected returns	Market Segments
Baseline Distraction Proxy	1/1962-12/2008	Volatility-weighting of absolute abnormal returns	OLS Market Model	49 Fama French Industries
Alternative Distraction Proxy 1	1/1961-12/2008	Equal-weighting of absolute abnormal returns	OLS Market Model	49 Fama French Industries
Alternative Distraction Proxy 2	1/1961-12/2008	Value-weighting of absolute abnormal returns	OLS Market Model	49 Fama French Industries
Alternative Distraction Proxy 3	7/1964-12/2008	Equal-weighting of absolute abnormal returns	Four-Factor Model	49 Fama French Industries
Alternative Distraction Proxy 4	7/1964-12/2008	Value-weighting of absolute abnormal returns	Four-Factor Model	49 Fama French Industries
Alternative Distraction Proxy 5	7/1965-12/2008	Volatility-weighting of absolute abnormal returns	Four-Factor Model	49 Fama French Industries
Alternative Distraction Proxy 6	1/1961-12/2008	Interquartile range of raw abnormal returns (Q3 - Q1)	OLS Market Model	49 Fama French Industries
Alternative Distraction Proxy 7	7/1964-12/2008	Interquartile range of raw abnormal returns (Q3-Q1)	Four-Factor Model	49 Fama French Industries
Alternative Distraction Proxy 8	1/1960-12/2008	Regression approach with raw returns as in Stivers/Sun (2010)	None	49 Fama French Industries
Alternative Distraction Proxy 9	7/1965-12/2008	Volatility-weighting of absolute abnormal returns	OLS Market Model	100 book-to-market x size
Alternative Distraction Proxy 10	7/1964-12/2008	Volatility-weighting of absolute abnormal returns	OLS Market Model	25 size x short-term reversal

Panel B: One Day Waiting, Full Sample Period

	All	1	2	3	4	5	6	7	8	9	10	**10-1**
Attention Proxy 1	0.0097***	0.0056***	0.0067***	0.0092***	0.0082***	0.0091***	0.0097***	0.0108***	0.0116***	0.0100***	0.0133***	**0.0077*** **
	(0.000315)	(0.00100)	(0.000921)	(0.000999)	(0.000991)	(0.00102)	(0.00101)	(0.00100)	(0.000917)	(0.000983)	(0.000981)	**(0.00140)**
Attention Proxy 2	0.0097***	0.0055***	0.0098***	0.0082***	0.0088***	0.0093***	0.0094***	0.0106***	0.0112***	0.0105***	0.0119***	**0.0064*** **
	(0.000315)	(0.000967)	(0.00102)	(0.000970)	(0.00114)	(0.000919)	(0.00101)	(0.000947)	(0.000948)	(0.000950)	(0.000990)	**(0.00138)**
Attention Proxy 3	0.0098***	0.0065***	0.0078***	0.0071***	0.0085***	0.0085***	0.0097***	0.0121***	0.0110***	0.0110***	0.0132***	**0.0067*** **
	(0.000332)	(0.00103)	(0.000967)	(0.00102)	(0.00109)	(0.00103)	(0.000989)	(0.00114)	(0.00105)	(0.000996)	(0.00101)	**(0.00145)**
Attention Proxy 4	0.0098***	0.0065***	0.0087***	0.0093***	0.0091***	0.0100***	0.0092***	0.0098***	0.0101***	0.0116***	0.0119***	**0.0053*** **
	(0.000332)	(0.00104)	(0.00102)	(0.00107)	(0.00113)	(0.00102)	(0.00107)	(0.00102)	(0.00106)	(0.000983)	(0.000997)	**(0.00144)**
Attention Proxy 5	0.0099***	0.0067***	0.0084***	0.0052***	0.0110***	0.0078***	0.0096***	0.0126***	0.0099***	0.0113***	0.0135***	**0.0068*** **
	(0.000336)	(0.00104)	(0.00101)	(0.00102)	(0.00102)	(0.00108)	(0.00112)	(0.00107)	(0.00111)	(0.000978)	(0.00102)	**(0.00146)**
Attention Proxy 6	0.0097***	0.0062***	0.0075***	0.0076***	0.0091***	0.0094***	0.0090***	0.0113***	0.0103***	0.0118***	0.0127***	**0.0065*** **
	(0.000319)	(0.00101)	(0.000983)	(0.000958)	(0.000958)	(0.00101)	(0.00101)	(0.000983)	(0.00101)	(0.00102)	(0.000997)	**(0.00142)**
Attention Proxy 7	0.0098***	0.0056***	0.0092***	0.0093***	0.0081***	0.0064***	0.0097***	0.0104***	0.0114***	0.0111***	0.0141***	**0.0085*** **
	(0.000332)	(0.000979)	(0.00104)	(0.00101)	(0.00102)	(0.00118)	(0.00100)	(0.00108)	(0.000945)	(0.00101)	(0.00104)	**(0.00143)**
Attention Proxy 8	0.0097***	0.0079***	0.0083***	0.0095***	0.0082***	0.0093***	0.0097***	0.0108***	0.0102***	0.0103***	0.0120***	**0.0041*** **
	(0.000315)	(0.00106)	(0.000965)	(0.00101)	(0.00102)	(0.00104)	(0.000910)	(0.000937)	(0.00105)	(0.000998)	(0.000938)	**(0.00142)**
Attention Proxy 9	0.0099***	0.0069***	0.0074***	0.0094***	0.0090***	0.0077***	0.0097***	0.0094***	0.0107***	0.0129***	0.0135***	**0.0066*** **
	(0.000336)	(0.000970)	(0.00107)	(0.00114)	(0.000987)	(0.00103)	(0.000981)	(0.00113)	(0.00103)	(0.00108)	(0.00103)	**(0.00141)**
Attention Proxy 10	0.0098***	0.0061***	0.0093***	0.0080***	0.0087***	0.0095***	0.0105***	0.0081***	0.0108***	0.0116***	0.0135***	**0.00742*** **
	-0.000332	(0.000929)	(0.00102)	(0.00103)	(0.00109)	(0.00104)	(0.00106)	(0.00106)	(0.00101)	(0.00104)	(0.00105)	**(0.00140)**

Table 7: Risk Factor Exposure of High Distraction Minus Low Distraction Portfolio Returns

This table shows results from monthly calendar time-series regressions which aim at quantifying the impact of well-known risk factors on our findings. To this end, we first construct monthly returns on zero-cost portfolios separately for pairs opened on high distraction (decile 10) and low distraction (decile 1) days. To mitigate the problem of days without trading, the maximum holding period is extended to six months (as opposed to one month in the baseline approach). We distinguish between a "no waiting" and a "one day waiting" return computation scheme as outlined in the text. We then compute the difference between the return on the high distraction portfolio and the return on the low distraction portfolio. The resulting time series of monthly long-short returns is regressed on well-known risk factors. The traded liquidity factor is taken from Lubos Pástor's homepage. The remaining factor returns are obtained from Kenneth French's data library. T-statistics (in parentheses) are computed using Newey-West standard errors with six lags.

Model Specification	"no waiting"			"one day waiting"		
	(1)	(2)	(3)	(4)	(5)	(6)
Sample Start	Jan 1962	Jan 1962	Jan 1968	Jan 1962	Jan 1962	Jan 1968
Sample End	Dec 2008	Dec 2008	Dec 2008	Dec 2008	Dec 2008	Dec 2008
Observations	564	564	492	564	564	492
Market Factor	-0.0633	-0.0581	-0.0324	-0.0189	-0.0122	0.0058
t-statistic	(-1.23)	(-1.10)	(-0.59)	(-0.44)	(-0.28)	(0.13)
Size Factor	0.0480	0.0586	0.0957*	0.0638	0.0682	0.0961*
t-statistic	(1.07)	(1.21)	(1.93)	(1.44)	(1.42)	(1.92)
Value Factor	-0.0780	-0.0560	-0.0300	-0.0556	-0.0362	-0.0192
t-statistic	(-1.52)	(-0.90)	(-0.49)	(-1.08)	(-0.63)	(-0.33)
Momentum Factor		0.0270	0.0179		0.0482	0.0371
t-statistic		(0.59)	(0.38)		(1.12)	(0.84)
Short Term Reversal Factor		-0.0094	-0.0263		0.0017	-0.0149
t-statistic		(-0.18)	(-0.48)		(0.032)	(-0.27)
Long Term Reversal Factor		-0.0322	-0.0519		-0.0152	-.02308
t-statistic		(-0.42)	(-0.67)		(-0.23)	(-0.34)
Liquidity Factor			0.0416			0.0588
t-statistic			(0.81)			(1.30)
Intercept	0.0038***	0.0035**	0.0037**	0.0030**	0.0025*	0.0029**
	(2.79)	(2.48)	(2.41)	(2.39)	(1.88)	(2.04)

Table 8: Further Tests (I): Investor Distraction and Pairs Trading Profitability

Panel A reports the impact of alternative limited attention dummies on the profitability of pairs trading. Details about the construction of each proxy are given in the text. Specification 1 displays findings from univariate tests. In specification 2, control variables correspond to those used in model 8 of table 3. Standard errors are adjusted for heteroscedasticity and clustered by day of pair divergence. T-statistics are reported in parentheses. Panel B shows interaction effects and corresponding t-statistics. In specification 1 (2), we regress the one-month event-time pairs return on the distraction proxy decile rank, an alternative limited attention dummy as in panel A, and on the interaction effect (as well as on the full set of controls). Panel C shows the implied percentage increase in return difference between high and low distraction scenarios when benchmarked against the baseline findings in tables 2 and 3. High (low) distraction scenarios are defined as days satisfying both distraction proxy decile rank 10 and alternative limited attention dummy=1 (distraction decile 1 + alternative dummy=0). In Panel D, we compare mean and median returns of pairs opening on the last trading day before federal holidays with mean and median returns of pairs opening on any other day of the year. The table shows the fraction of years in which returns on pre-holiday pairs trading are higher. P-values (in parentheses) are computed from one-sided binominal probability tests with an assumed yearly success rate of 50%. *Excess probability of convergence* is computed as the difference between the fraction of converging pairs that diverged immediately before the federal holiday and the fraction of converging pairs that diverged on any other day of the year. T-statistics are reported in parentheses. In all panels, statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Impact of Limited Attention Proxies Derived from Previous Work					
Model specification	Return Computation	Fridays	No. pairs opening	NBER recessions	3 year market return
Specification 1 (no further controls)	No waiting	0.0006	0.0046***	0.0113***	0.0066***
		(0.74)	(4.20)	(10.50)	(6.36)
	One day waiting	0.0002	0.0033***	0.0090***	0.0055***
		(0.21)	(3.18)	(8.99)	(5.51)
Specification 2 (full set of controls)	No waiting	0.0010	0.0002	0.0050**	0.0006
		(0.42)	(0.18)	(2.55)	(0.37)
	One day waiting	0.0006	0.0007	0.0041**	0.0004
		(0.27)	(0.68)	(2.37)	(0.22)
Panel B: Interaction Effects from Combining the Distraction Proxy with Limited Attention Dummies Derived from Previous Work					
Model specification	Return Computation	Fridays	No. pairs opening	NBER recessions	3 year market return
Specification 1 (no further controls)	No waiting	0.0004	0.0006	0.0006*	0.0004
		(1.54)	(1.50)	(1.78)	(1.21)
	One day waiting	0.0005*	0.0003	0.0004	0.0003
		(1.71)	(0.81)	(1.25)	(1.03)
Specification 2 (full set of controls)	No waiting	0.0004	0.0002	0.0003	0.0002
		(1.34)	(0.62)	(0.79)	(0.79)
	One day waiting	0.0004	0.0002	0.0002	0.0004
		(1.50)	(0.43)	(0.69)	(1.34)
Panel C: Implied Percentage Increase in the Return Difference between High Distraction and Low Distraction Scenarios					
Model specification	Return Computation	Fridays	No. pairs opening	NBER recessions	3 year market return
Specification 1 (no further controls)	No waiting	28.50%	51.47%	139.28%	80.13%
	One day waiting	29.68%	50.86%	150.22%	94.45%
Specification 2 (full set of controls)	No waiting	2.93%	4.59%	135.09%	30.17%
	One day waiting	2.25%	9.73%	154.04%	35.36%
Panel D: Relative Success of Pre-Holiday Pairs Trading					
Pairs Trading Returns		Thanksgiving Day	Christmas Day	New Year's Day	Independence Day
Mean		61%*	70%***	72%***	51%
Median		55%	63%*	70%***	53%
Excess probability of convergence		0.42%	7.39%*	15.77%***	5.61%
(baseline probability: 36.15%-36.3%)		(0.10)	(1.68)	(3.67)	(1.35)
Pairs Trading Returns		Washington's Birthday	Labor Day	Memorial Day	
Mean		61%*	60%	47%	
Median		46%	60%	45%	
Excess probability of convergence		-0.48%	3.87%	3.87%	
(baseline probability: 36.15%-36.3%)		(-0.14)	(1.14)	(0.94)	

Table 9: Further Tests (II): Investor Distraction and Pairs Trading Profitability

Panel A shows the sensitivity of pairs trading returns to distraction proxy decile ranks (as observed on the day of divergence) for several samples of top pairs: The monthly top 100 pairs each consisting of firms from different industries, the monthly top 100 pairs each consisting of firms from the same industries, and the monthly top 20 pairs each consisting of two value-weighted industries. In all cases, we use the Fama/French (1997) classification with 49 industries. The first column shows the return difference between pairs diverging on high distraction days (decile 10) and pairs diverging on low distraction days (decile 1). The approach resembles the methodology used in table 2. The second column shows findings from regressing pairs returns on the distraction proxy decile rank. Panel B reports results from a test similar in spirit. It compares the return sensitivity for pairs whose constituent firms share (do not share) at least one business segment, as described in detail in the text. Panel C compares returns from pairs consisting only of firms with high residual media coverage and pairs consisting only of firms with low residual media coverage, as described in detail in the text. The first column compares average event-time one-month pairs trading returns. The second column shows findings from regressing pairs returns on the distraction proxy decile rank. In all panels, statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Panel A: Impact of Investor Distraction on Pairs Consisting of Alternative Assets			
	Return computation	Diff.: Decile 10-Decile 1	Distraction Decile Rank
Top 100 pairs with stocks	one day waiting	0.00770***	0.00065***
from different industries (N=103,386)		(0.00143)	(0.0000708)
Top 100 pairs with stocks		0.0071***	0.00057***
from the same industry (N=100,726)		(0.00120)	(0.0000925)
Top 20 industry-level pairs		0.0039**	0.00033**
(N=14,180)		(0.00181)	(0.0001394)

Panel B: Pairs With and Without Common Industry Segments (since 1977)			
	Return computation	Diff.: Decile 10-Decile 1	Distraction Decile Rank
No shared industry segment	one day waiting	0.0097***	0.00073***
		(0.0023)	(0.00018)
Shared industry segment		0.0084**	0.00019
		(0.0039)	(0.00030)
Difference		0.0013	0.00054*
		(0.0038)	(0.00029)

Panel C: Pairs with High and Low Residual Media Coverage			
	Return computation	Return: All Deciles	Distraction Decile Rank
Low residual media coverage	one day waiting	0.0077***	0.0010*
		(0.0016)	(0.0006)
High residual media coverage		-0.0001	-0.0012
		(0.0030)	(0.0010)
Difference		0.0082**	0.0022*
		(0.0034)	(0.0012)

Appendix for "Losing sight of the trees for the forest? Pairs trading and attention shifts"

October 2011

Abstract

This appendix contains figures and tables that supplement the analysis in the paper. Figures 1 to 3 illustrate the relationship between specific characteristics of abnormal industry-level returns and resulting distraction proxy decile ranks. Figure 4 shows the empirical cumulative distribution function of the return on US stock pairs sorted by distraction proxy decile ranks as observed on the day of divergence. Table 1 gives an overview of the data samples used in the study. Table 2 reports distribution details of the return on US stock pairs sorted by distraction proxy decile ranks as observed on the day of divergence. Table 3 provides descriptive statistics for pairs trading samples in international stock markets. Table 4 displays a transition matrix of baseline distraction proxy decile ranks and modified ranks available in real time. Table 5 shows findings from multivariate regression when relying on this real-time distraction proxy. Table 6 explores the impact of time-varying arbitrage risk on pairs trading profitability. Table 7 reports cross-sectional tests (no waiting return computation scheme) on the link between investor distraction and pairs trading profitability.

Figure 1: Mean of Abnormal Industry Returns by Distraction Proxy Decile Ranks

The following figures are intended to illustrate the relationship between specific characteristics of abnormal industry-level returns and resulting distraction proxy decile ranks. To this end, we compute, at each day, the mean of different types of abnormal industry returns, where industries are represented by the 49 Fama/French (1997) segments. The upper graph uses (raw) abnormal returns, the middle graph uses absolute abnormal returns, and the lower graph uses weighted absolute abnormal returns, where industry weights are determined by the inverse of the volatility of their shock variables over the previous year. See section 2 of the paper for a detailed description of how abnormal returns and weights are computed. By distraction proxy decile ranks, the boxes illustrate the 25th, 50th, and 75th percentile of the time-series of the mean of these abnormal returns. The adjacent values in the box plot are the most extreme values within 1.5*interquartile range (75th percentile-25th percentile) of the nearer quartile.

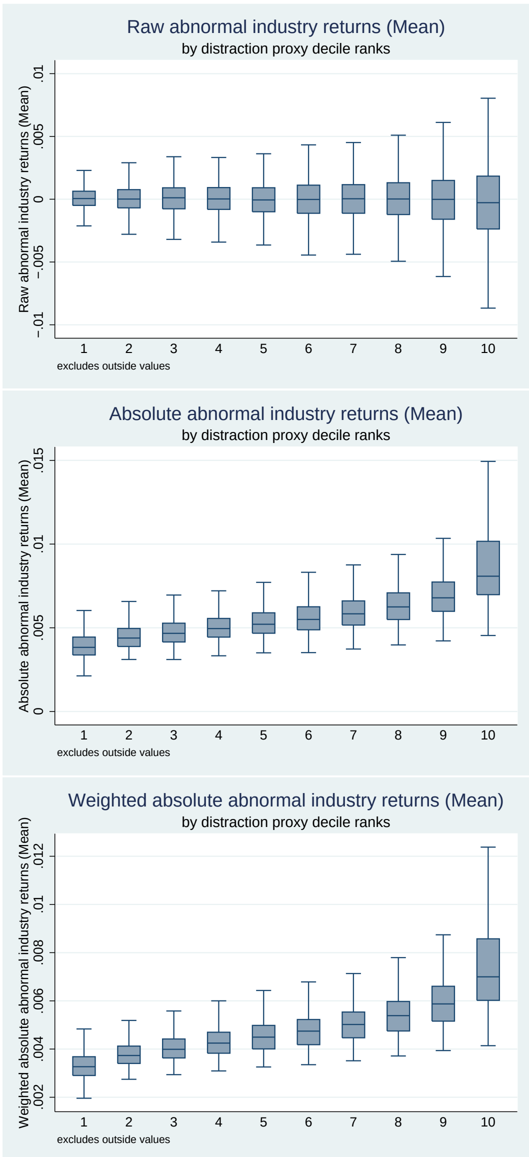


Figure 2: Standard Deviation of Abnormal Industry Returns by Distraction Proxy Decile Rank

The following figures are intended to illustrate the relationship between specific characteristics of abnormal industry-level returns and resulting distraction proxy decile ranks. To this end, we compute, at each day, the standard deviation of different types of abnormal industry returns, where industries are represented by the 49 Fama/French (1997) segments. The upper graph uses (raw) abnormal returns, the middle graph uses absolute abnormal returns, and the lower graph uses weighted absolute abnormal returns, where industry weights are determined by the inverse of the volatility of their shock variables over the previous year. See section 2 of the paper for a detailed description of how abnormal returns and weights are computed. By distraction proxy decile ranks, the boxes illustrate the 25th, 50th, and 75th percentile of the time-series of the standard deviation of these abnormal returns. The adjacent values in the box plot are the most extreme values within 1.5*interquartile range (75th percentile-25th percentile) of the nearer quartile.

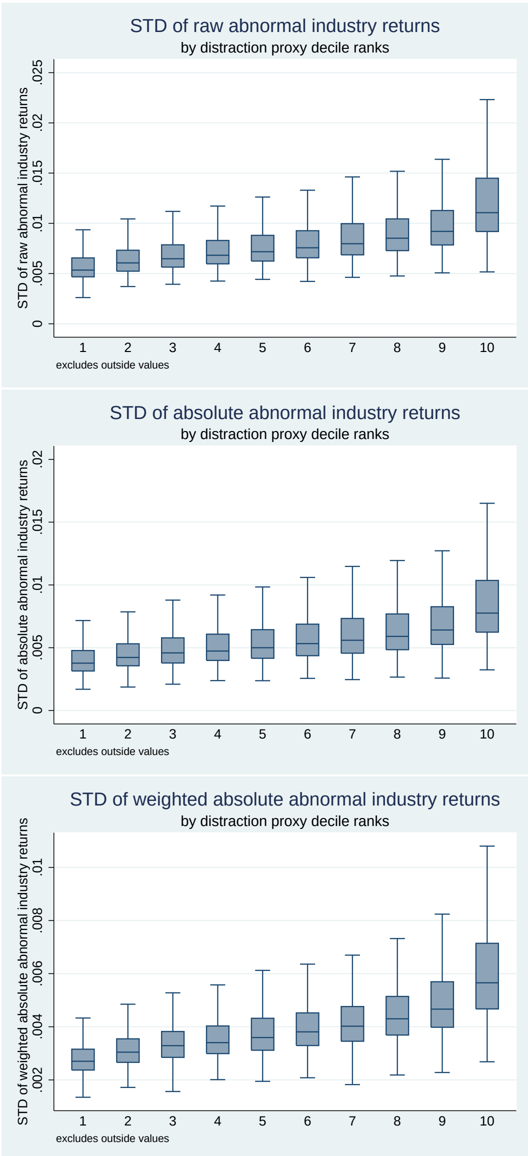


Figure 3: Maximum Fraction of a Single Industry Return Shock by Distraction Proxy Decile Rank

The following figure is intended to illustrate the relationship between specific characteristics of abnormal industry-level returns and resulting distraction proxy decile ranks. To this end, we first compute the baseline distraction proxy as outlined in section 2 of the paper. In short, the proxy is constructed as the sum of weighted absolute abnormal industry returns, where industries are represented by the 49 Fama/French (1997) segments. For each day, we then identify the maximum fraction a single industry return shock accounts for. By distraction proxy decile ranks, the boxes illustrate the 25th, 50th, and 75th percentile of the time-series of this maximum weight. The adjacent values in the box plot are the most extreme values within 1.5*interquartile range (75th percentile-25th percentile) of the nearer quartile.

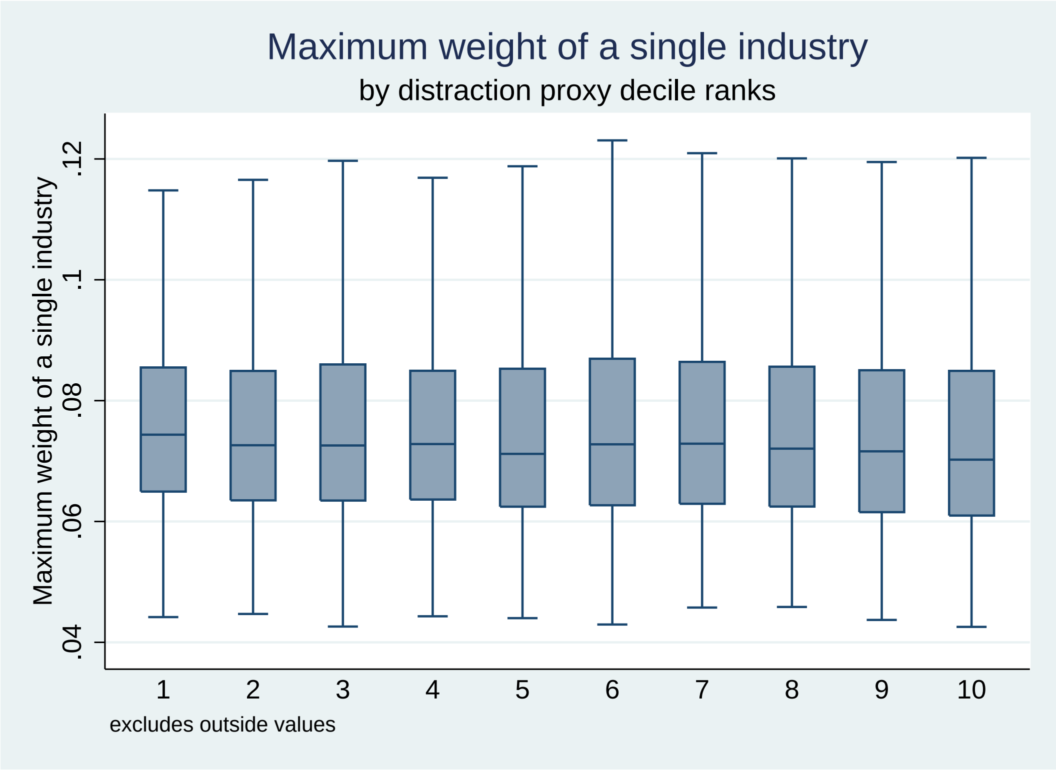


Figure 4: Cumulative Distribution Function of Pairs Trading Profits by Distraction Deciles

This figure shows the empirical cumulative distribution function of event-time one-month returns on zero-cost portfolios of US stock pairs sorted by distraction proxy decile ranks as observed on the day of divergence. Breakpoints for the deciles are determined separately for each year. We only consider low distraction days (=decile 1) and high distraction days (=decile 10). In Panel A (B), trading positions in each pair are initiated on the day of divergence (on the day following the convergence) and liquidated on the day of convergence (on the day following the convergence). For better readability, extreme returns (larger than 20% or smaller than -20%) are not shown. Extreme returns account for roughly 1% of all sample observations.

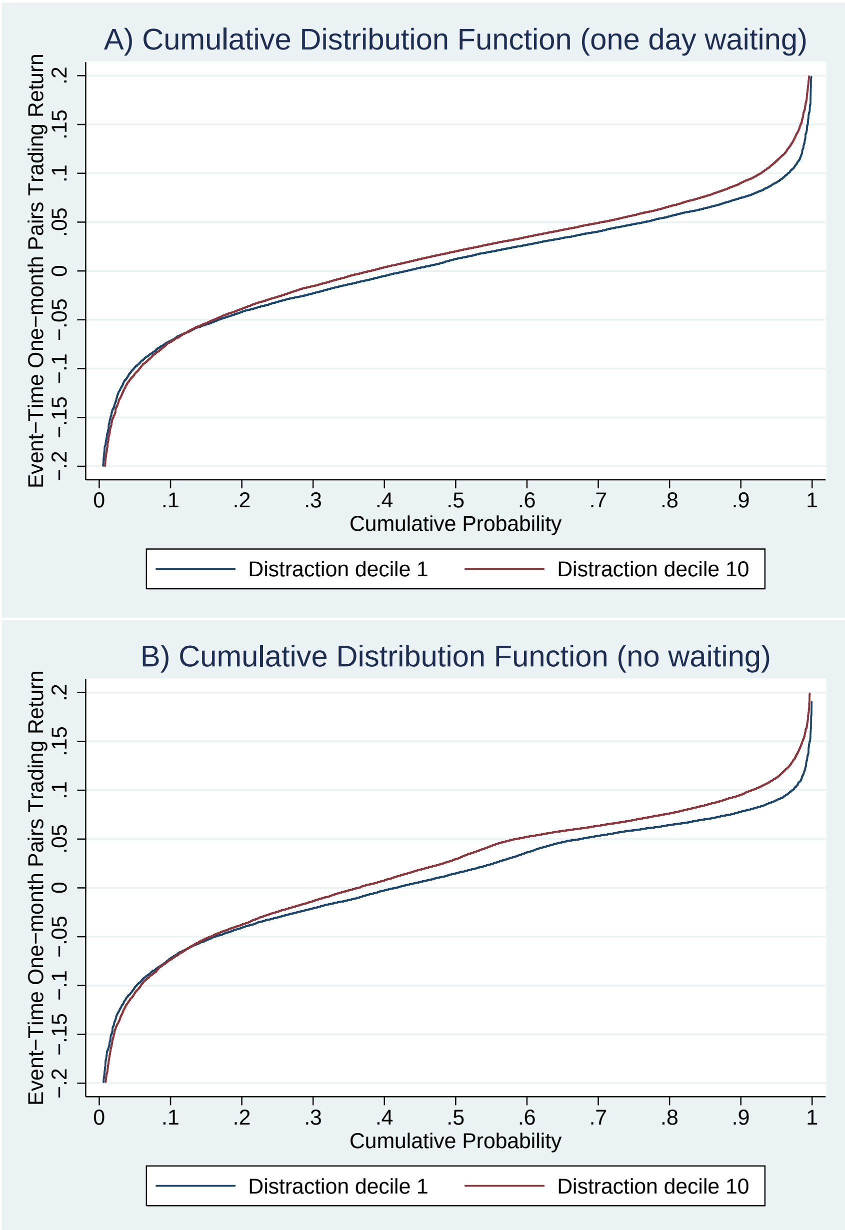


Table 1: Overview of Data Sets Used in the Study

Data Source	Description	Sample Period
CRSP	Daily stock market data for all US firms trading on NYSE, AMEX or NASDAQ	January 1960 - December 2008
Compustat Global Daily Stock File	Daily stock market data for firms of eight large non-North American stock markets (Japan, UK, France, Germany, Switzerland, Italy, Netherlands, Hongkong)	Middle of the 1990ies - December 2009
Compustat Fundamentals Annual and Segment Files	Financial information about firm segments based on yearly data	1977-2008
Factiva	Yearly number of Dow Jones News Services articles about each firm that meets data requirements on pairs trading in some period after 1990	1991-2008
IBES	Number of IBES analysts for sample firms	January 1980 - December 2008
Kenneth R. French's Data Library	Several portfolio returns and risk factors	Middle of the 1960ies - December 2008
Lubos Pástor's homepage	Traded liquidity factor	January 1968 - December 2008
NBER	US business cycle expansions and contractions based on monthly data	January 1960 - December 2008

Table 2: Distribution of One-month Pairs Trading Abnormal Returns by Distraction Deciles

This table reports distribution details of event-time one-month returns on zero-cost portfolios of US stock pairs sorted by distraction proxy decile ranks as observed on the day of divergence. Breakpoints for the deciles are determined separately for each year. Calculations are based on daily data from January 1962 to December 2008. In Panel A (B), trading positions in each pair are initiated on the day of divergence (on the day following the convergence) and liquidated on the day of convergence (on the day following the convergence).

Distraction Decile	All	1	2	3	4	5	6	7	8	9	10
Panel A: One Day Waiting, Full Sample Period (1962-2008)											
N	103,386	8,187	8,679	9,146	9,398	10,048	10,079	10,595	11,019	12,224	14,011
mean	0.97%	0.53%	0.86%	0.77%	0.87%	0.95%	0.94%	1.02%	1.10%	1.09%	1.30%
sd	6.57%	6.20%	6.30%	6.53%	6.32%	6.58%	6.56%	6.54%	6.52%	6.73%	7.05%
p1	-17.83%	-17.23%	-17.08%	-18.08%	-17.37%	-17.04%	-18.70%	-17.72%	-17.73%	-18.06%	-18.59%
p10	-7.04%	-7.16%	-6.60%	-7.02%	-6.95%	-7.02%	-6.95%	-6.98%	-6.94%	-7.28%	-7.28%
p25	-2.73%	-3.16%	-2.73%	-2.77%	-2.70%	-2.76%	-2.61%	-2.70%	-2.58%	-2.68%	-2.64%
p50	1.59%	1.24%	1.41%	1.36%	1.52%	1.57%	1.49%	1.62%	1.71%	1.72%	2.01%
p75	5.23%	4.80%	4.97%	5.05%	5.02%	5.17%	5.24%	5.27%	5.33%	5.47%	5.71%
p90	8.21%	7.49%	7.79%	7.85%	7.90%	.0809082	8.16%	8.28%	8.31%	8.54%	9.00%
p99	15.19%	13.49%	14.62%	14.41%	13.81%	15.04%	15.25%	15.46%	14.92%	16.07%	16.77%
Panel B: No Waiting, Full Sample Period (1962-2008)											
N	104,125	8,222	8,738	9,179	9,436	10,094	10,122	10,657	11,146	12,332	14,199
mean	1.38%	0.89%	1.20%	1.08%	1.23%	1.31%	1.35%	1.33%	1.53%	1.56%	1.90%
sd	6.77%	6.40%	6.41%	6.70%	6.57%	6.75%	6.71%	6.76%	6.69%	6.95%	7.30%
p1	-18.55%	-17.21%	-17.65%	-18.74%	-18.33%	-17.98%	-19.22%	-19.17%	-17.93%	-18.61%	-19.29%
p10	-7.11%	-7.20%	-6.65%	-7.13%	-7.06%	-7.05%	-7.00%	-7.10%	-7.13%	-7.31%	-7.36%
p25	-2.62%	-3.03%	-2.58%	-2.74%	-2.61%	-2.63%	-2.44%	-2.75%	-2.49%	-2.63%	-2.45%
p50	2.14%	1.48%	1.80%	1.68%	1.96%	2.03%	2.09%	2.16%	2.35%	2.54%	2.94%
p75	6.33%	5.90%	5.96%	6.03%	6.11%	6.27%	6.29%	6.33%	6.45%	6.62%	6.95%
p90	8.47%	7.79%	7.99%	8.04%	8.18%	8.33%	8.32%	8.45%	8.57%	8.82%	9.54%
p99	13.80%	12.12%	12.40%	13.12%	12.83%	13.12%	12.99%	13.50%	13.63%	14.30%	15.85%

Table 3: Descriptive Statistics for Pairs Trading Samples in International Stock Markets

In panel A, *Total market capitalization* refers to the value reported by Thomson Financial Datastream for the total domestic market capitalization at the year-end 2002, which roughly marks the middle of the sample period for most countries . In panel B, *Number of industries* states how many of the 10 Global Industry Classification Standard (GICS) industry sectors are represented in the pairs trading sample. *Maximum industry weight* denotes the largest fraction of sample firms belonging to a specific industry group. *Industry concentration* is computed as the sum of squared industry weights. In panel C, values are computed analogously for within-pair industry group combinations.

		Japan	UK	France	Germany	Switzerland	Italy	Netherlands	Hongkong
Panel A: Overall Market Characteristics									
Sample period		1/1995-12/2009	1/1995-12/2009	1/1996-12/2009	1/1996-12/2009	6/1997-12/2009	6/1995-12/2009	1/1995-12/2009	1/1995-12/2009
Total market cap. (in billion USD)		2100.19	1819.29	877.87	651.57	540.76	448.22	429.80	417.61
Number of sample firms		4,873	4,867	1,424	1,302	387	500	354	521
Number of firm days (in million)		12.00	6.35	2.11	2.54	0.32	0.47	0.68	0.87
Panel B: Firm Characteristics									
Number of industries		10	10	10	10	10	10	9	10
Max. industry weight	Fraction	26.50%	20.88%	25.88%	21.29%	30.13%	31.65%	26.99%	38.11%
	Industry	Industrials	Cons. Discretionary	Financials	Industrials	Financials	Financials	Financials	Financials
Industry concentration		16.16%	14.81%	16.46%	21.29%	18.07%	18.79%	17.46%	22.16%
Panel C: Pair Characteristics									
Total no. of pairs traded		36,992	27,185	25,833	26,006	16,841	25,596	21,320	20,894
Average turnover	Mean	0.16%	0.44%	0.21%	0.35%	0.43%	0.41%	0.50%	0.56%
at day of divergence	Median	0.09%	0.28%	0.12%	0.20%	0.25%	0.26%	0.28%	0.29%
Cumul. return difference	Mean	6.53%	8.17%	9.86%	11.70%	12.41%	10.90%	11.13%	16.50%
upon divergence	Median	5.99%	7.37%	9.45%	11.20%	11.47%	9.96%	10.69%	14.96%
Return difference	Mean	3.70%	3.21%	4.07%	4.73%	4.24%	4.30%	4.18%	5.91%
at day of divergence	Median	2.88%	2.40%	3.30%	3.69%	3.00%	3.34%	3.07%	4.46%
No. industry group comb.		41	45	44	42	42	44	36	45
Max. industry group weight	Fraction	11.55%	13.98%	13.39%	12.40%	14.96%	21.29%	19.20%	24.82%
	Industries	Industrials and Financials	Cons. Discretionary and Financials	Industrials and and Financials	Industrials Cons. Discretionary	Health Care and Financials	Con. Discretionary and Financials	Industrials and Financials	Industrials and Financials
Industry group concentration		6.61%	6.00%	6.71%	6.29%	8.31%	9.26%	7.83%	13.05%

Table 4: Transition Matrix: Baseline Distraction Proxy Decile Ranks and Modified Ranks Available in Real Time (1/1963-12/2008, N=102,259)

This table shows the transition matrix for baseline distraction proxy decile ranks and modified decile ranks, which are based on rolling historical raw proxy data. Specifically, for a given day, we use the preceding 250 trading days to compute the distraction decile the day belongs to. This implies that the beginning of the sample period is one year later (January 1963) than in the baseline case. Computations are based on all 102,259 sample pair trades, which are opened between January 1963 and December 2008. This implies that distraction proxy ranks at days where more (less) pairs diverge have a stronger (weaker) impact on the results displayed.

		Modified distraction proxy based on rolling historical data										
	Decile rank	1	2	3	4	5	6	7	8	9	10	Total
	1	64.55%	17.04%	4.82%	1.67%	0.97%	0.49%	0.21%	0.00%	0.00%	0.00%	7.91%
	2	21.02%	42.55%	18.31%	8.68%	5.55%	2.36%	0.81%	0.44%	0.13%	0.00%	8.38%
	3	8.32%	19.41%	36.20%	18.91%	9.46%	8.04%	3.06%	0.95%	0.42%	0.00%	8.83%
	4	4.49%	8.52%	18.25%	33.55%	15.57%	11.03%	7.14%	4.14%	1.28%	0.11%	9.07%
Baseline	5	1.41%	6.03%	8.13%	17.98%	35.40%	17.63%	12.46%	7.23%	2.39%	0.35%	9.70%
	6	0.20%	4.88%	5.87%	6.71%	17.58%	31.48%	17.18%	12.82%	4.99%	1.48%	9.73%
	7	0.00%	1.33%	5.62%	6.60%	5.89%	16.81%	31.62%	21.47%	10.93%	2.46%	10.26%
	8	0.00%	0.23%	1.92%	4.30%	5.64%	7.35%	17.31%	31.14%	24.39%	6.93%	10.70%
	9	0.00%	0.00%	0.89%	1.40%	3.46%	3.52%	7.34%	16.73%	42.47%	21.69%	11.82%
	10	0.00%	0.00%	0.00%	0.20%	0.49%	1.29%	2.88%	5.08%	12.99%	66.97%	13.60%
	Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 5: Multivariate Analysis: Investor Distraction (Real-Time Availability) and Returns on Pairs Trading

This table displays findings from pooled multivariate regressions of the one-month return on zero-cost US stock pairs on a proxy for investor distraction and up to three sets of control variables. The proxy for investor distraction is the *Distraction Proxy Decile Rank* (specifications 1-4) or a *High/Low Distraction Dummy* (specifications 5-8), which is zero for low distraction days (decile 1) and one for high distraction days (decile 10). In contrast to the baseline analysis in the paper, decile ranks are now constructed on the basis of rolling historical data to assure availability in real time. Specifically, for a given day, we use the preceding 250 trading days to compute the distraction decile the day belongs to. Pairs trading returns are computed under the conservative "one day waiting" return scheme. The first set of explaining variables controls for calendar and industry effects (indicator variables for year, month, day of week, and pair industry group combinations). The second set controls for market-level conditions on the day of divergence (market return, squared market return, market turnover, 10 day rolling volatility, factors for daily return premia on size, value, momentum and short-term reversal). The third set controls for a number of pair characteristics computed as outlined in table 1 (average firm market capitalization decile rank, ln (average pre-event turnover), ln (average pre-event Amihud illiquidity ratio), average idiosyncratic risk, within-pair differences in these variables, return difference attributable to the day of divergence, ln (average turnover on day of divergence) and ln (difference in turnover on day of divergence)). Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Model specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample period	1/1963-12/2008	1/1963-12/2008	7/1963-12/2008	7/1963-12/2008	1/1963-12/2008	1/1963-12/2008	7/1963-12/2008	7/1963-12/2008
Observations	101,539	101,539	100,426	99,673	25,963	25,963	25,705	25,538
Adjusted R^2	0.05%	2.63%	2.74%	2.84%	0.30%	5.02%	5.30%	5.36%
Distraction Proxy Decile Rank	0.00052***	0.00067***	0.00045***	0.00050***				
	(0.000109)	(0.000110)	(0.000119)	(0.000120)				
High/Low Distraction Dummy					0.00559***	0.00839***	0.00660***	0.00729***
					(0.001312)	(0.001486)	(0.001615)	(0.001650)
Controls for calender and industry effects	no	yes	yes	yes	no	yes	yes	yes
Controls for market-level conditions	no	no	yes	yes	no	no	yes	yes
Controls for pair characteristics	no	no	no	yes	no	no	no	yes

Table 6: The Impact of Time-Varying Arbitrage Risk (as Proxied by the VIX) and Investor Distraction on Pairs Trading Profitability

This table displays findings from pooled regressions of the one-month return on zero-cost US stock pairs on the proxy for investor distraction and/or the Chicago Board Options Exchange Market Volatility Index (VIX) over the period from January 1990 to December 2008. The proxy for investor distraction is the *Distraction Proxy Decile Rank* (specifications 3,5,6) or a *High/Low Distraction Dummy* (specifications 4,7,8), which is zero for low distraction days (decile 1) and one for high distraction days (decile 10). With regard to the VIX, we either use its raw value (specifications 1,5,7) or its decile ranks, computed separately for each year (specifications 2,6,8). Pairs trading returns are computed under the conservative "one day waiting" return scheme. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Model specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample period	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008	1/1990-12/2008
observations	39,666	39,666	39,666	8,372	39,666	39,666	8,372	8,372
Adjusted R^2	0.12%	0.15%	0.10%	0.30%	0.17%	0.18%	0.48%	0.52%
VIX (Raw Value)	0.00028***				0.00023***		0.00011	
	(0.000079)				(0.000081)		(0.000124)	
VIX (Yearly Decile Ranks)		0.00095***				0.00078***		0.000708*
		(0.000180)				(0.000187)		(0.000403)
Distraction Proxy Decile Rank			0.00079***		0.00060***	0.00053***		
			(0.000177)		(0.000177)	(0.000184)		
High/Low Distraction Dummy				0.01070***			0.009790***	0.00832***
				(0.002210)			(0.002221)	(0.002430)

Table 7: Further Tests (II): Investor Distraction and Pairs Trading Profitability

Panel A shows the sensitivity of pairs trading returns to distraction proxy decile ranks (as observed on the day of divergence) for several samples of top pairs: The monthly top 100 pairs each consisting of firms from different industries, the monthly top 100 pairs each consisting of firms from the same industries, and the monthly top 20 pairs each consisting of two value-weighted industries. In all cases, we use the Fama/French (1997) classification with 49 industries. The first column shows the return difference between pairs diverging on high distraction days (decile 10) and pairs diverging on low distraction days (decile 1). The approach resembles the methodology used in table 2 of the paper. The second column shows findings from regressing pairs returns on the distraction proxy decile rank. Panel B reports results from a test similar in spirit. It compares the return sensitivity for pairs whose constituent firms share (do not share) at least one business segment, as described in detail in the text. Panel C compares returns from pairs consisting only of firms with high residual media coverage and pairs consisting only of firms with low residual media coverage, as described in detail in the text. The first column compares average event-time one-month pairs trading returns. The second column shows findings from regressing pairs returns on the distraction proxy decile rank. In all panels, statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Standard errors (in parentheses) are adjusted for heteroscedasticity and clustered by day of pair divergence.

Panel A: Impact of Investor Distraction on Pairs Consisting of Alternative Assets			
	Return computation	Diff.: Decile 10-Decile 1	Distraction Decile Rank
Top 100 pairs with stocks	no waiting	0.0101***	0.00089***
from different industries (N=103,386)		(0.00151)	(0.0000726)
Top 100 pairs with stocks		0.0094***	0.00083***
from the same industry (N=100,726)		(0.00156)	(0.0001154)
Top 20 industry-level pairs		0.0047**	0.00036**
(N=14,180)		(0.00190)	(0.0001451)

Panel B: Pairs With and Without Common Industry Segments (since 1977)			
	Return computation	Diff.: Decile 10-Decile 1	Distraction Decile Rank
No shared industry segment	no waiting	0.0115***	0.00096***
		(0.0024)	(0.00019)
Shared industry segment		0.0112***	0.00043
		(0.0039)	(0.00030)
Difference		0.00022	0.00053*
		(0.0039)	(0.00030)

Panel C: Pairs with High and Low Residual Media Coverage			
	Return computation	Return: All Deciles	Distraction Decile Rank
Low residual media coverage	no waiting	0.0108***	0.0011**
		(0.0017)	(0.0005)
High residual media coverage		0.0003	-0.0010
		(0.0030)	(0.0011)
Difference		0.0106***	0.0021*
		(0.0035)	(0.0012)