

NBER WORKING PAPER SERIES

PAIRS TRADING: PERFORMANCE OF A
RELATIVE VALUE ARBITRAGE RULE

Evan G. Gatev
William N. Goetzmann
K. Geert Rouwenhorst

Working Paper 7032
<http://www.nber.org/papers/w7032>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1999

We are grateful to Peter Bossaerts, Jon Ingersoll, and Carl Schecter for helpful discussions on this topic. We thank the International Center for Finance at the Yale School of Management for research assistance. We thank the participants in the Finance workshop at Vanderbilt University for their comments. The views expressed in this paper are those of the authors and do not reflect those of the National Bureau of Economic Research.

© 1999 by Evan G. Gatev, William N. Goetzmann, and K. Geert Rouwenhorst. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Pairs Trading: Performance of a Relative Value Arbitrage Rule
Evan G. Gatev, William N. Goetzmann,
and K. Geert Rouwenhorst
NBER Working Paper No. 7032
March 1999

ABSTRACT

We test a Wall Street investment strategy known as “pairs trading” with daily data over the period 1962 through 1997. Stocks are matched into pairs according to minimum distance in historical normalized price space. We test the profitability of several trading rules with six-month trading periods over the 1962-1997 period, and find average annualized excess returns of up to 12 percent for a number of self-financing portfolios of top pairs. Part of these profits may be due to market microstructure effects. Nevertheless, our historical trading profits exceed a conservative estimate of transaction costs through most of the period. We bootstrap random pairs in order to distinguish pairs trading from pure mean-reversion strategies. The bootstrap results suggest that the ‘pairs’ effect differs from previously documented mean reversion profits.

Evan G. Gatev
Yale School of Management
135 Prospect Street
P.O. Box 208200
New Haven, CT 06511

William N. Goetzmann
Yale School of Management
135 Prospect Street
P.O. Box 208200
New Haven, CT 06511
and NBER
william.goetzmann@yale.edu

K. Geert Rouwenhorst
Yale School of Management
135 Prospect Street
P.O. Box 208200
New Haven, CT 06511

I. Introduction

Wall Street has long been interested in quantitative methods of speculation. One popular short-term speculation strategy is known as "pairs trading." The strategy has at least a fifteen year history on Wall Street and is among the proprietary "statistical arbitrage" tools currently used by hedge funds as well as investment banks. The concept of pairs trading is disarmingly simple. Find two stocks whose prices have moved together historically. When the spread between them widens, short the winner and buy the loser. If history repeats itself, prices will converge and the arbitrageur will profit. It is hard to believe that such a simple strategy, based solely on past price dynamics and simple contrarian principles, could possibly make money. If the U.S. equity market were efficient, risk-adjusted returns from pairs trading should not be positive.

In this paper, we examine the risk and return characteristics of pairs trading with daily data over the period 1962 through 1997. Using a simple algorithm for choosing pairs, we test the profitability of several straightforward, self-financing trading rules. We find average annualized excess return of about 12 percent for top-pairs portfolios. We show that part of this may be due to institutional factors. In particular, Jegadeesh and Titman (1995) find evidence that a significant proportion of profits from short-term contrarian strategies may be due to asynchronous trading and the bid-ask bounce. After the bid-ask bounce and transaction costs are taken into account, pairs trading yields reduced -- but still positive and significant -- returns.

Regardless of the magnitude of the profits from pairs trading, our results reveal something about the mechanism and performance of actual relative-price arbitrage activities. This is useful, because, despite considerable theory about market efficiency, economists have little empirical information about how such efficiency is maintained. Event studies show us that new information

is rapidly impounded into stock prices, presumably through speculative activity, however by their very nature event studies are conditional upon significant news releases that focus attention upon the securities. Studies about active speculators and institutional investors are tantalizing -- they suggest some degree of herding behavior that may be motivated by opportunity for speculative gains, however we do not yet really understand why certain stocks attract institutional attention. In this paper, we investigate the mechanisms and profits of relative pricing in a setting in which speculators are trading in securities that are close economic substitutes for each other. While pairs trading may be only one of a number of "convergence" strategies employed by arbitrageurs, it has the attraction of corresponding quite nicely to a model-free, but highly intuitive "Law of One Price," in a way that will be made clear below.

Our results fall in a gray area of profitability -- attractive enough to sustain low cost market participants whose presence maintains relative efficiency, but not attractive enough to support large amounts of investor capital. Our study explores the statistical characteristics of pairs trader portfolio returns, and thus helps understand one interesting mechanism of market efficiency -- relative pricing of close substitutes.

II. Background

II.1 History

In the mid-1980's, the Wall Street quant Nunzio Tartaglia assembled a team of physicists, mathematicians and computer scientists to uncover arbitrage opportunities in the equities markets. Tartaglia's group of former academics used sophisticated statistical methods to develop high-tech trading programs, executable through automated trading systems, that took the intuition and trader's "skill" out of arbitrage and replaced it with disciplined, consistent filter rules. Among other things,

Tartaglia's programs identified pairs of securities whose prices tended to move together. They traded these pairs with great success in 1987 -- a year when the group reportedly made a \$50 million profit for the firm. Although the Morgan Stanley group disbanded in 1989 after a couple of bad years of performance, pairs trading has since become an increasingly popular "market-neutral" investment strategy used by institutional traders as well as hedge fund managers. The increased popularity of quantitative-based statistical arbitrage strategies has also apparently affected profits. In a recent New York Times interview, David Shaw, head of one of the most successful modern quant shops and himself an early Tartaglia's protégé, suggests that recent pickings for quant-shops have become slim -- he attributes the success of his firm D.E. Shaw to early entry into the business. Tartaglia's own explanation for pairs trading is psychological. He claims, that "...Human beings don't like to trade against human nature, which wants to buy stocks after they go up not down."¹ Could pairs traders be the disciplined investors taking advantage of the undisciplined over-reaction displayed by individual investors? This is at least one possible -- albeit psychological - explanation for our results, which is consistent with Jegadeesh and Titman's (1995) finding that contrarian profits are in part due to over-reaction to company-specific information shocks rather than price reactions to common factors.

II.2 Data Snooping

In our study we have not searched over a strategy space to identify successful trading rules, but rather we have interpreted practitioner description of pairs trading as straightforwardly as possible. Our rules follow the general outline of first "find stocks that move together," and second

¹ Hansell, Saul, 1989, "Inside Morgan Stanley's Black Box," *Institutional Investor*, May, p.204.

"take a long-short position when they diverge." A test requires that both of these steps must be parameterized in some way. How do you identify "stocks that move together?" Need they be in the same industry? Should they only be liquid stocks? How far do they have to diverge before a position is put on? When is a position unwound? We have made some straightforward choices about each of these questions. We put positions on at a two-standard deviation spread, which might not always cover transactions costs even when stock prices converge. Although it is tempting to try potentially more profitable schemes, the danger in data-snooping refinements outweigh the potential insights gained about the higher profits that could result from learning through testing.

As it stands now, data-snooping is a serious concern in our study. Pairs trading is closely related to a widely studied subject in the academic literature -- mean reversion in stock prices.² We consider the possibility that we have simply reformulated a test of the previously documented tendency of stocks to revert towards their mean at certain horizons. To address this issue, we develop a bootstrapping test based upon random pair choice. If pairs-trading profits were simply due to mean-reversion, then we should find that randomly chosen pairs generate profits, i.e. that buying losers and selling winners in general makes money. This simple contrarian strategy is unprofitable over the period that we study, suggesting that mean reversion is not the whole story.

Although the effect we document is not merely an extension of previously known anomalies, it is still not immune to the data-snooping argument. Indeed we have explicitly "snooped" the data to the extent that we are testing a strategy we know to have been actively exploited by risk-arbitrageurs. As a consequence we cannot be sure that past trading profits under our simple

² See, for example, Rosenberg and Rudd (1982), DeBondt and Thaler (1985), Bossaerts (1988), Conrad and Kaul (1989), Jegadeesh (1990), Lehman (1990), Lo and MacKinlay (1990) and Lakonishok, Shleifer and Vishny (1994) for evidence on mean reversion in stock prices and potential explanations for the phenomenon.

strategies will continue in the future. This potential critique has another side, however. The fact that pairs trading is already well-known risk-arbitrage strategy means that we can simply test current practice rather than develop our filter-rule *ad hoc*.

II.3 Relative Pricing

Asset pricing can be viewed in absolute and relative terms. Absolute pricing values securities from fundamentals such as discounted future cash flow. This is a notoriously difficult process with wide margin for error. Recent papers by Bakshi and Chen (1997) and Lee, Myers and Swaminathan (1997) are heroic attempts to build quantitative value investing models. Relative pricing is only slightly easier. Relative pricing means that two securities that are close substitutes for each other should sell for the same price -- it does not say what that price will be. Thus, relative pricing allows for bubbles in the economy, but not necessarily arbitrage or profitable speculation. The Law of One Price [LOP] -- and a "near-LOP" is applicable to relative pricing -- even if that price is wrong.

Ingersoll (1987) defines the LOP as the "proposition ... that two investments with the same payoff in every state of nature must have the same current value." In other words, two securities with the same prices in all states of the world should sell for the same thing. Chen and Knez (1995) extend this to argue that "*closely integrated* markets should assign to similar payoffs prices that are *close*." They argue that two securities with similar, but not necessarily matching payoffs across states should have similar prices. This is of course a weaker condition, and subject to bounds on prices for unusual states, however it allows the examination of "near-efficient" economies, or in Chen and Knez' case, near integrated markets. Notice that this theory corresponds to the desire to

find two stocks whose prices "move together" as long as we can define states of nature as the time-series of observed historical trading days.

We use an algorithm to choose pairs based on the criterion that they have had the same or nearly the same state prices historically. We then trade pairs whose prices closely match in historical state-space, since the LOP suggests that in an efficient market, their prices should be nearly identical. In this framework, the current study can be viewed as a test of the LOP and near-LOP in the U.S. equity markets, under certain stationarity conditions. We are effectively testing the integration of very local markets -- the markets for specific individual securities. This is similar in spirit to Bossaerts (1988) test of co-integration of security prices at the portfolio level. We further conjecture that the marginal profits to be had from risk arbitrage of these temporary deviations is crucial to the maintenance of first-order efficiency. We could not have the first effect without the second.

III. Methodology

Our implementation of pairs trading has two stages. We form pairs over a twelve month period (formation period) trade them in the next six-month period (trading period). Both twelve months and six months are chosen arbitrarily and have remained our horizons throughout the study.

III.1 Pairs Formation

In each pairs-formation period, we screen out all stocks from the CRSP daily files that have one or more days with no trade. This serves to identify relatively liquid stocks as well as to facilitate pairs formation. Next, we construct a cumulative total return index for each stock over the

formation period. We then choose a matching partner for each stock by finding the security that minimizes the sum of squared deviations between the two normalized price series. Pairs are thus formed by exhaustive matching in normalized daily "price" space, where price includes re-invested dividends.

We use this approach because it best approximates the description of how traders themselves choose pairs. Interviews with pair traders suggest that they try to find two stocks whose prices "move together." In addition to "unrestricted" pairs, we will also present results by sector, where we restrict both stocks to belong to the same broad industry categories defined by Standard & Poors: Utilities, Transportation, Financial and Industrials. Each stock is assigned to one of these four groups, based on the stock's SIC code. The minimum-distance criterion is then used to match stocks within each of the groups.

III.2 Trading Period

Once we have paired up all liquid stocks in the formation period, we study the top 5 and 20 pairs with the smallest historical distance measure, in addition to the 20 pairs after the top 100 (pairs 101-120). This last set is valuable because most of the top pairs share certain characteristics which will be described in detail below. On the day following the last day of the pairs formation period, we begin to trade according to a pre-specified rule. We chose rules based on the proposition that we open a long-short position of when the pair prices have diverged by a certain amount, and close the position when the prices have reverted. Following practice, we base our rules for opening and closing positions on a standard deviation metric. We open a position in a pair when prices diverge by more than two historical standard deviations, as estimated during the pairs formation period. We

unwind the position at the next crossing of the prices. If prices do not cross before the end of the trading interval, gains or losses are calculated at the end of the last trading day of the trading interval. Since the positions are effectively self-financing portfolios, we report the payoffs by going one dollar short in the higher-priced stock and one dollar long in the lower-priced stock.

Figure 1 provides illustrates the pairs trading strategy using two stocks, Kennecott and Uniroyal in the six month period starting in August of 1962. The top two lines represent the normalized price paths with dividends re-invested, and the bottom line indicates the opening and closing of the strategy on a daily basis. It is clear why these two firms paired with each other. They generally tended to move together over the trading interval. Notice that the position first opens in the seventh trading day of the period and then remains open until day 36. Over that interval, the spread actually first increased significantly before convergence. The pair remains close in price during the period and cross frequently. The pair opens five times during the period, however not always in the same direction. Neither stock is the "leader." In our example, convergence occurs in the final day of the period, although this is not always the case.

III.3 Excess Return Computation

Because pairs may open and close at various points during the six-month trading period, the calculation of the excess return on a portfolio of pairs is a non-trivial issue. Pairs that open and converge during the trading interval will have positive cash flows. Because pairs can re-open after initial convergence, pairs can have multiple positive cash flows during the trading interval. Pairs that open but do not converge will only have cash flows on the last day of the trading interval when all positions are closed out. Therefore, the payoffs to pairs trading strategies are a set of positive cash

flows that are randomly distributed throughout the trading period, and a set of cash flows at the end of the trading interval which can either be positive or negative. For each pair we can have multiple cash flows during the trading interval, or we may have none in the case when prices never diverge by more than two standard deviations during the trading interval. Because the trading gains and losses are computed over long–short positions of one dollar, the payoffs have the interpretation of excess returns. The excess return on a pair during a trading interval is computed as the sum of the payoffs during the trading interval.³

We consider two measures of excess return on a *portfolio* of pairs: the return on committed capital and the return on actual employed capital. The excess return on committed capital takes the sum of the payoffs over all pairs during the trading period, and divides it by the number of pairs in the portfolio. This measure of excess return is clearly conservative – if a pair does not trade for the whole of the trading period, we still include a dollar of committed capital in our calculation of excess return. A hedge fund would presumably be more flexible in its sources and uses of funds. In such case computing excess return relative to the actual capital employed may give a more realistic measure of the trading profits. We calculate the excess return on employed capital as the sum of the pair payoffs divided by the number of pairs that actually open during the trading period. For example, if only one pair trades in the top 5 portfolio, then the excess return on employed capital would be five times the excess return on committed capital.

³ This is a conservative approach to computing the excess return, because it implicitly assumes that all cash is received at the end of the trading period. Because any cash flows during the trading interval is positive by construction, it ignores the fact that these cash flows are received early, and understates the computed excess returns.

III.4 Pairs Portfolios

We initiate the strategy, trading the pairs at the beginning of every month in the sample period, with the exception of the first twelve months, which are needed to estimate pairs for the strategy starting in the first month. The one-month waiting period before rebalancing is an arbitrary choice. One way to motivate the six-month trading period with one-month period between re-matching of pairs is to think of a proprietary trading desk which delegates the management of the six portfolios to six different traders whose evaluation periods are staggered by one month. At the end of a trading period, the corresponding trader starts with a new portfolio whose pairs are formed over the prior year.

IV. Empirical Results

IV.1 Strategy Profits

Table I summarizes the excess returns for the pairs portfolios that are unrestricted in the sense that the matching stocks do not necessarily belong to the same broad industry categories. Panel A shows the results when positions are opened at the end of the day that prices diverge, and closed at the end of the day of price convergence. Panel B reports the returns if we open and close the positions at the end of the day following divergence and convergence. Panel A indicates that a portfolio of the five best pairs earns an average excess return of 5.98 percent on committed capital over a six-month period while the portfolio of all 1870 pairs yields 4.10 percent. The corresponding t -statistics of the mean, computed using Newey-West standard errors, are 9.13 and 10.70 respectively. The corresponding average six-month excess returns on employed capital are slightly higher at 6.01 and 4.23 percent. These excess returns are large in a statistical sense, and seem to

suggest that pairs trading is profitable.

The remainder of Panel A provides information about the return distributions of pairs portfolios. There are diversification benefits from combining multiple pairs in a portfolio. As the number of pairs in a portfolio increases, the portfolio standard deviation falls, as does the range of the realized returns and the frequency of negative portfolio excess return during a six-months period. For example, during the full sample period of 34 years, a portfolio of 20 pairs has only 6 six-month periods with negative payoffs, while a portfolio of 5 pairs returns negative profits in 11 trading periods. The distribution of pairs payoffs is skewed right and peaked relative to the normal distribution.

Since pairs-trading is in essence a contrarian investment strategy, the returns may be biased upward due to the bid-ask bounce (Jegadeesh (1990), Jegadeesh and Titman (1995), Conrad and Kaul (1989)). In particular, the strategy buys sells stocks that have done well relative to their match and buys those that have done poorly. Part of any observed price divergence is potentially due to price movements between bid and ask quotes: conditional on divergence the winner's price is more likely to be an ask quote and the loser's price a bid quote. Since we have used these same prices for the start of trading, our returns may be due to the fact that we are implicitly buying at bid quotes (losers) and selling at ask quotes (winners). The opposite is true at the second crossing (convergence): part of the drop in the winner's price can reflect a bid quote, and part of the rise of the loser's price an ask quote.

To address this issue, we explore the effect of waiting a day before initiating a position. Panel B of Table I gives the excess returns when we initiate positions in each pair on the day following the divergence and liquidate on the day following the crossing. The average excess returns

drop by about 200 basis points. While the excess returns are still significantly positive in a statistical sense, the dramatic drop in the excess returns suggests that a non-trivial portion of the profits in Panel A may be due to bid-ask bounce. It is difficult to quantify which portion of the profit reduction is due to bid-ask bounce and which portion stems from true mean reversion in prices due to rapid market adjustment. None-the-less, this difference raises questions about the economic significance of our results when we include transactions costs. We will return to a detailed discussion of this issue at a later point in the paper.

IV.2 Trading Statistics and Portfolio Composition

Table II provides detailed analysis of the trading statistics and composition of the pairs portfolios. What are the characteristics of the stocks that are matched into pairs? How often does a typical pair trade? Because pairs trading is an active investment strategy, it is important to evaluate the profitability relative to the trading intensity of the portfolios. An important question is whether positions are opened too early, i.e. does the two standard deviation trigger effectively limit the profits to be made from the strategy, and how does the trigger relate to the bid-ask spread of the stocks? In the second line of panel A in Table II the average 2 standard deviation "open position" trigger is reported. For the top five pairs, the position typically opens when prices have diverged by 4% or more. This is a relatively narrow gap in prices. The optimal trigger point in terms of profitability may actually be much higher than 2 standard deviations, although we have not experimented to find out. The trigger spread increases for 20 pairs and all pairs because the standard deviation of the pairs is increasing as the proximity of the securities in price space diverges. Panel A also shows that nearly all pairs open in the six-month trading period and that the average number

of round-trips per pair in the trading interval, including closing out the position at the end of the six months, is 2.4. This means that most of the pairs open during the trading interval, and many pairs open multiple times.

Panel B of Table II describes the composition of the pairs in terms of market capitalization and industry membership. In terms of size, the average stock in the top 5 and top 20 pairs belongs to the third decile from the top. 68% of the pairs' stocks belong to the top three size deciles, and 87% come from the top five size deciles of the CRSP data. Most pairs combine stocks from different size deciles (i.e. size "mixed" pairs), and the average size differs by a single decile.

A breakdown of the pairs by industry composition is instructive. On average, 82% of the stocks in the top 20 pairs are utility stocks, despite the fact that utilities represent a fairly small proportion of the stocks in the whole sample. The "deeper" 20 pairs have a larger proportion of industrials, and for this group, "cross-sector" pairs occur frequently: about 9 times out of 20. These break-downs raise the question whether the profitability of pairs trading is limited to the utility sector, or whether pairs strategies are also profitable in other sectors of the market. We will return to this question at a later point in the paper.

IV.3 Time-Series Characteristics of Performance

Table III reports evidence about the risk of pairs trading. We calculate the average daily dollar change in the position value for the strategy across all pairs, under the assumption that these positions are marked to market at the close of each day and then cumulate these to monthly excess returns and compare them to the S&P 500 and treasury bills. We report the excess returns to the one-day wait trading strategy in annualized form and for comparison we report total returns to

treasury bills and the S&P 500. The top Panel of Table III shows that the pairs trading portfolios are much less volatile than the S&P 500, with annualized standard deviations ranging from 4% to 6% per year. The combination of low volatility and high average return explains that the annual Sharpe ratios for the pairs portfolios are high relative to the S&P500. Also, the excess returns tend to be positively autocorrelated at the monthly horizon. This suggests that the profitability of pairs position accrues gradually over time, and is not simply driven by short-term price reversals.

In order to explore the systematic risk exposure of the pairs portfolios, we regress their monthly excess returns on the three factors of Fama-French (1996). The exposures of pairs portfolios to the market return are generally small, and not significantly different from zero. This is not surprising because pairs formation matches stocks which are highly correlated, and therefore tend to move together under all market conditions. Exposures to the other FF factors is generally positive and significant. Pairs returns are positively correlated with the difference between small and big stocks (SMB), and the difference between value and growth stocks (HML). However, these exposures are not sufficient to explain their average returns: all of the 3-factor regression intercepts are significantly different from zero. Exposure to the FF factors explains only about 100bp of the average annual performance of the pairs returns.

The last panel of Table III shows the results of regressing the pairs returns on the Ibbotson factors: the excess return on the S&P 500, a U.S. small stock premium, a U.S. bond default premium and a U.S. bond horizon premium.⁴ The risk exposure to the S&P 500 is slightly negative, the

⁴These indices were obtained from the Ibbotson Associates EnCorr analyzer. U.S. small stock premium is the geometric difference between the small stock return and the S&P 500 stock return. The default premium is the geometric difference between the total return to long term corporate bonds and the total return to long term government bonds. The horizon premium is the geometric difference between total return to long term government bonds and treasury bills.

exposure to the small stock premium is significantly positive, the exposure to the default premium is positive and the exposure to the horizon premium is positive. Thus, when corporate bonds increase in price relative to government bonds, the pairs portfolios make money. There are a range of possible explanations for this pattern. Relatively cheaper borrowing rates by arbitrageurs may force stock prices closer to equilibrium values, or common factors affecting convergence in both stock and bond markets may be responsible. The portfolios also appear to be sensitive to shifts in the yield curve. I.e. when long-term spreads decrease, pairs trading is more profitable. At first glance, the sensitivity to term-structure measures may be explained by the presence of interest rate sensitive Utility stocks in many of the top pairs. However, interest rate movements also seem to matter for the more broadly diversified pairs portfolios. In sum, the pairs portfolios seem to be exposed to different sources of systematic risk. While the R-squared is low, these portfolios are evidently not factor-neutral, despite being constructed in a way that should essentially match up economic substitutes.

Value-at-risk is a potentially useful framework for evaluating pairs trading risk. Table IV reports both weekly and monthly VAR numbers by summarizing the quantiles of the empirical distributions. The worst week out of the entire period from 1962 through 1996 was an 11.18 % loss below the riskless rate for the top five pair portfolio and an 8.85 % loss below the riskless rate for the top 20 portfolio. Only one per hundred weeks did these portfolios lose more than 2.55 % and 1.32 % respectively. The worst month out of the entire period from 1962 through 1996 was a 15.93 % loss below the riskless rate for the top five pair portfolio and a 13.81 % loss below the riskless rate for the top 20 portfolio. Only one per hundred months did these portfolios lose more than 4.65 % and 2.28 % respectively.

The VAR is useful because it provides a gauge to the potential leverage that could be applied to these strategies. A five-to-one leverage ratio applied to the top 5 pairs would appear to have been adequate to cover the worst monthly loss in the 35 year period. Although the lessons of recent history have taught us not to rely too heavily on historical VAR measures for gauging capital needs for exploiting convergence strategies, the pairs portfolios seem to be exposed to relatively little risk.

Figure 2 shows the monthly performance of the top 20 pairs, based upon the same-day trading rule. Pairs trading has declined in profitability dramatically from the 1970's and 1980's to a low point at the end of our sample when the returns were sometimes negative. Figure 3 compares the cumulative excess returns of the top 20 one-day-waiting strategy the cumulative excess returns of investment in the S&P 500 index. The smooth index of the pairs trading portfolio contrasts dramatically with the volatility of the stock market. Pairs trading performed well over difficult times for U.S. stocks. When the U.S. stock market suffered a dramatic real decline from 1969 through 1980, the pairs strategy had some of its best performance. By contrast, for the last four years in the sample, the market has performed exceptionally well, but pairs trading profits have been flat. Perhaps after its discovery in the early 1980's by Tartaglia and others, competition has decreased opportunity. On the other hand, pair trading might simply be more profitable in times when the stock market performs poorly.

IV.4 Pairs trading by industry group

The pairs formation process thus far has been entirely mechanical. A computer stock has the opportunity to match with a steel firm, and a utility with a bank. This does not mean that these matches are likely. As shown in Table II, the fraction of mixed pairs is typically well below 50

percent. Common factor exposures of stocks in the same industry will make it more likely to find a match within the same sector. Also, firms that are in industries where cross-sectional differences in factor exposures are small or return variances are low are more likely to end up among the top ranking of pairs. For this reason it is perhaps not surprising that many of the top pairs thus far contain utilities. Are the profits to pairs trading consistent across sectors? We also examine the returns on pairs trading where stocks are matched only within the four large sector groupings used by Standard and Poors: Utilities, Transportation, Financials, and Industrials. The results are given in Table V. As in Table I, the pairs are traded with a one-day delay before opening and closing a position in order to minimize the effect of trading on the bid-ask bounce. The six-month excess returns are the largest in the Utilities sector, with 5.04 percent ($t=10.70$). The profits for the other industry groups are somewhat lower, but all statistically significant, with the average Transportation, Financials, and Industrials top 5 pairs earning 3.21 ($t = 4.46$), 3.74 ($t = 5.67$) respectively and 2.43 ($t = 4.57$) percent over a six-month period.

Table VI gives a more detailed picture of the return distributions and trading characteristics of the pairs trading strategies by sector. It shows that the excess return distributions of the Transportation and Industrial pairs are skewed left, while pairs formed from Utilities and Financials are skewed right. With the exception of the top 5 pairs portfolio of Transportation all distributions have negative excess kurtosis relative to a normal distribution. Table VI shows that the relatively large excess returns for Utilities come at a cost: it is also the most trading intensive sector with around 2.5 roundtrips per pair. Transportation pairs experience the least amount of trading with 1.6 roundtrips per pair. The conclusion from these Tables is that pairs trading is profitable in every broad sector category, and not limited to a particular sector, although effects are most pronounced for

Utilities.

IV.5 Pairs trading and contrarian investment

Because pairs trading bets on return reversals it is an example of a contrarian investment strategy. Positive returns to contrarian investing have been documented in the at long horizons (DeBondt and Thaler (1985,1987), Lakonishok, Shleifer and Vishny (1994)). For our study, which examines convergence within a six-month period, the relevant studies are those which have explored return reversals at short-term horizons between one week and one month: (Lehman (1990), Jegadeesh (1990)). Given the average length of time that our pairs positions stay open (1.5 months), this raises the question whether our pairs trading strategies are merely a disguised way of exploiting these previously negative autocorrelations. In an attempt to explore this possibility, we conduct a bootstrap where we construct random pairs within sectors, and trade these pairs following the same rules as our actual pairs. Based on the findings of Lehmann (1990) and Jegadeesh (1990), we expect these random pairs to generate positive profits on average. To evaluate whether our strategies are profitable beyond the documented short-term reversals, we compute the p -values of our pairs trading profits under the bootstrapped distribution. The results are in the last lines of each of the panels in Table VI. Surprisingly, we find that the mean of the bootstrapped distribution is positive in only two of four sectors. However, in each instance the documented profits have a very low p -value under the bootstrapped distribution.

IV. 6 Transactions costs

Table I shows that the average excess return of unrestricted pairs strategies falls from 5.98 per cent to 3.99 per cent per month if we postpone the trades to the day following the crossing. This drop in the excess returns implies an estimate of the lower bound of the average bid-ask spread and hence the transactions costs of trading in the sample. While actual transactions costs may exceed this lower bound, it is informative to know whether the trading profits are large enough to survive this conservative estimate of transactions costs.

Suppose the extreme case where the prices of the winner at the first crossing (divergence) are ask prices and the loser are bid prices. If the next day prices are equally likely to be at bid or ask the delaying trades by one day will reduce the excess returns on average by half the sum of the spreads of the winner and the loser. If at the second crossing (convergence) of the pairs the winners is trading at the bid, and the loser at the ask, waiting one day will reduce the excess returns on average again by one half of the sum of the bid-ask spreads of both stocks. In this extreme case, waiting a day before trading reduces the return on each pair by the roundtrip transactions costs in that pair. Because we trade each pair on average 2.4 times during the trading interval, the drop in the excess returns of 200 basis points by waiting one day reflects the cost of 2.4 roundtrips, which implies a transactions costs of 83 bp per pair per roundtrip. This may be interpreted as an estimated effective spread of 42 bp. This is a conservative estimate (lower bound) of the transactions costs because it assumes that postponing trading by one day leads to a price movement within the spread for every trade.⁵ It is also consistent with transactions costs estimated by Peterson and Fialkowski

⁵ To the extent that not all trades are affected, the excess return reduction reflects the transactions costs on fewer than 2.4 roundtrips, which leads to a larger estimate of the roundtrip spread.

(1994), who find that the average effective spread for stocks in the CRSP database in 1991 was 37 basis points. Indeed, since 86% of the stocks in the top 20 pairs were in the top 5 deciles of CRSP stocks, we would expect the effective spread to be even lower than 37 bp.

Do our trading strategies survive these transactions costs? The profits on our trading strategies in the top 5 – 20 pairs in Table I range from 368 to 288 basis points over a six-month period. If the prices used to compute these excess returns are equally likely to be at bid or ask, which seems a reasonable assumption, we have to correct these excess returns to reflect that in practice we buy at ask and sell at bid prices. In other words, we have to subtract the roundtrip trading costs to get an estimate of the profits after transactions costs. Our conservative estimate of transactions costs of 83 bp times 2.4 rounds trips per pair results in an estimate of 200 bp transactions cost per pair per six month period. This gives average net profits ranging from 168 to 88 bp over each six-month period. Comparing these profits to the reported standard errors, we conclude that they are both economically and statistically significant.

Further analysis is required to get more precise estimates of influence of transactions costs of pairs trading strategies. An important question in this context is whether the trading rule that we have used to open and close pairs can be expected to generate economically significant profits *even if pairs trading works perfectly*. Because we use a measure of historical standard deviation to trigger the opening of pairs, and since this estimated standard deviation is the smallest among all pairs it is likely to underestimate the true standard deviation of a pair. As a consequence, we may simply be opening pairs "too soon" and at a point that we cannot expect it to compensate for transactions costs

even if the pair subsequently converges. Preliminary and unreported results suggest that this is indeed the case for some of our pairs.

There is a second reason why our trading strategies require "too much" trading. We open pairs at any point during the trading period when the normalized prices diverge by two standard deviations. This is not a sensible rule towards the end of a trading interval. For example, suppose a divergence occurs at the next to last day of the trading interval. The convergence has to be substantial in order to overcome the transactions cost that will be incurred when we close out the position on the next day (the last day of the trading interval). Preliminary results suggest that this is also an important source of excess trading.

V. Conclusion

We examine contrarian strategies based on the notion of cointegrated prices in a reasonably efficient market, known on Wall Street as Pairs Trading. We form pairs of stocks, which are close substitutes according to a minimum distance criterion using a metric in price space. We find that trading suitably formed pairs of stocks exhibits profits, with are robust to conservative estimates of transaction costs. These profits are uncorrelated to the S&P 500, however they do exhibit some sensitivity to the spreads between small and large stocks and between value and growth stocks in addition to the spread between high grade and intermediate grade corporate bonds and shifts in the yield curve. Because the strategies are trading intensive, the profitability of the strategy clearly depends upon the price and the impact of execution.

Our findings are consistent with the hypothesis of positive cross-autocorrelation in security returns at the daily level. And they appear not to be explained by simple mean reversion

documented in the literature. Whether or not pairs trading is consistently profitable, it may be an important mechanism of relative price equilibration by market participants. It is not surprising that few individual investors -- even day traders -- have heard about pairs trading. Larger players such as institutions are likely to have a relative advantage in their ability to command leverage to take positions and their ability to execute trades cheaply. On the other hand competition in the industry and the price impact of large trades may be important factors limiting the scale of pairs trading.

References:

- Bakshi, Gurdip and Zhiwu Chen, 1997, "Stock Valuation in Dynamic Economies," working paper, Ohio State University.
- Daniel, Kent, David Hirshleifer, and A. Subrahmanyam, 1996, A theory of overconfidence, self-attribution, and security market under- and over-reaction, working paper, Northwestern University and University of Michigan and UCLA.
- Bossaerts, Peter, 1988, Common nonstationary components of asset prices, *Journal of Economic Dynamics and Control* 12, 347-364.
- Brock, William, Joseph Lakonishok and Blake LeBaron, 1992, Simple technical trading rules and the stochastic properties of stock returns, *Journal of Finance* 47, 1731-1764.
- Chen, Zhiwu and Gurdip Bakshi, 1995, "Measurement of Market Integration and Arbitrage," *Review of Financial Studies* 8(2), pp. 287-325.
- DeBondt, Werner and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-805.
- DeBondt, Werner and Richard Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.
- DeBondt, W. F. M., and R. H. Thaler, 1995, Financial Decision-Making in Markets and Firms: A Behavioral Perspective, in *Finance*, ed. by R. Jarrow, V. Maksimovic, and W. Ziemba. Elsevier/North Holland, Amsterdam, vol. 9 of *Handbook in Operations Research and Management Science*, chap. 13, pp. 385-410.
- Fama, Eugene F. and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 131-155.
- Ingersoll, Jonathan Jr., 1987, *Theory of Financial Decision-Making*, Rowman and Littlefield, New Jersey.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and Selling Losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1995, Overreaction, Delayed Reaction, and Contrarian Profits, *The Review of Financial Studies* Vol.8 No. 4, 973-93.

Lee, Charles, M.C., James N. Meyers and Bhaskaran Swaminathan, 1997, "What is the Intrinsic Value of the Dow?" Working paper, Cornell University, University of Washington and Cornell University.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lo, Andrew W. and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

Peterson, Mitchell and David Fialkowski, 1994, "Posted versus effective spreads: Good prices or bad quotes?" *Journal of Financial Economics*, 35, 269-292.

Ready, Mark, 1997, Profits from technical trading rules, working paper, University of Wisconsin – Madison.

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

Table I: Excess Returns on Unrestricted Pairs Trading Strategies

Summary statistics of the 6-month excess returns on equally-weighted portfolios of pairs. We trade according to the rule that opens a position in a pair when the prices of the stocks in the pair diverge by two historical standard deviations. The "top n" portfolios include the n pairs with least distance measures, and the portfolios "20 after top 100" has the pairs after the top 100 pairs. The average number of pairs in the all-pair portfolio is 1870. There are 408 observations, from 2/1963 until 12/1997

Portfolio	top 5	top 20	20 after top 100	All
Panel A: Excess Return Distribution (No Waiting)				
Mean (Committed capital)	0.0598	0.0601	0.0451	0.0410
Standard error (Newey-West)	0.0065	0.0055	0.0038	0.0038
t-statistic	9.1363	10.8866	11.8900	10.7022
Excess return distribution				
Median	0.0549	0.0543	0.0457	0.0325
Standard deviation	0.0704	0.0526	0.0482	0.0354
Kurtosis	0.7323	2.5906	0.2319	1.5074
Skewness	0.4352	0.9781	0.1500	1.1730
Minimum	-0.1423	-0.0792	-0.0793	-0.0248
Maximum	0.2972	0.3124	0.1946	0.1756
Observations with excess return < 0	16%	9%	16%	7%
Mean excess return on employed capital	0.0601	0.0606	0.0456	0.0423
Panel B: Excess Return Distribution (1 day Waiting)				
Mean (Committed capital)	0.0368	0.0396	0.0354	0.0288
Standard Error (Newey-West)	0.0051	0.0042	0.0034	0.0033
t-statistic	7.2635	9.3286	10.4686	8.6709
Excess return distribution				
Median	0.0297	0.0360	0.0357	0.0231
Standard deviation	0.0581	0.0426	0.0453	0.0312
Kurtosis	1.2998	3.2452	0.0950	0.9549
Skewness	0.3509	0.9127	0.0426	0.9333
Minimum	-0.1402	-0.0976	-0.0827	-0.0335
Maximum	0.2860	0.2478	0.1781	0.1479
Observations with excess return < 0	24%	16%	22%	15%
Mean excess return on employed capital	0.0370	0.0400	0.0358	0.0297

Table II: Trading Statistics and Portfolio Composition

Trading statistics and portfolio composition for portfolios of pairs formed over a 12 month period according to a minimum distance criterion, and then traded over the subsequent 6 month period. We trade according to the rule that opens a position in a pair when the prices of the stocks in the pair diverge by two historical standard deviations. The "top n" portfolios include the n pairs with least distance measures, and the portfolios "20 after top 100" has the pairs after the top 100 pairs. The average number of pairs in the all-pair portfolio is 1870. There are 408 observations, from 2/1963 until 12/1997

Portfolio	top 5	top 20	20 after top 100	All
Panel A: Trading Statistics				
Mean (Committed capital)	0.0598	0.0601	0.0451	0.0410
Average 2 std dev 'open position' trigger	0.0399	0.0449	0.0683	0.1518
Average number of pairs traded	4.97	19.81	19.77	1814.46
Mean round-trip trades per pair	2.4	2.2	1.9	1.7
Standard deviation of round-trips, per pair	0.8	0.5	0.3	0.2
Average time open, per pair, months	3.8	3.8	4.0	3.9
Std dev of open position, per pair, months	0.7	0.5	0.4	0.2
Mean excess return on employed capital	0.0601	0.0606	0.0456	0.0423
Panel B: Pair Portfolio Composition				
Average decile of a pair's stock	3.1	3.2	4.0	5.8
Pairs' stocks in top 3 size deciles	68%	63%	52%	29%
Pairs' stocks in top 5 size deciles	87%	86%	72%	46%
Mixed size pairs	3.7	15.4	15.8	1494.9
Average decile difference for mixed pairs	1.0	1.0	1.0	1.0
Utility stocks	81%	82%	33%	9%
Transportation stocks	1%	1%	2%	3%
Financial stocks	4%	4%	16%	12%
Industrial stocks	13%	13%	48%	76%
Mixed sector pairs	0.8	3.4	8.6	549.2

Table III: Systematic Risk

This table reports the monthly risk profile for portfolios of pairs formed and traded according to the "Wait One Day" rule discussed in the text, over the period 8/1963 until 12/1997. The "top n" portfolios include the n pairs with least distance measures, and the portfolios "20 after top 100" has the pairs after the top 100 pairs. The average number of pairs in the all-pair portfolio is 1870. The U.S. Small stock premium is the monthly geometric difference between small company stock total returns and large company stock total returns. The Ibbotson factors are from the Ibbotson EnCorr analyzer. U.S. bond default premium is the monthly geometric difference between total return to long-term corporate bonds and long term government bonds. The U.S. bond horizon premium is the monthly geometric difference between investing in long term government bonds and U.S. treasury bills. The Fama and French Factors include HM: the high book-to-market stock returns minus low book to market stock returns and SMB: small stock returns minus large stock returns. Returns for the portfolios are in excess of the riskless rate. S&P 500 returns are calculated in excess of treasury bill returns.

	top 5	top 20	20 aft. 100	All	Eq. Pre.
Wait One Day Portfolio Performance					
Geometric Mean Excess Return, annualized	0.0743	0.0802	0.0711	0.0580	0.0513
Standard deviation, annualized	0.0497	0.0380	0.0384	0.0359	0.1535
Annual Sharpe Ratio	1.5196	2.1260	1.8700	1.6314	0.4062
Monthly Serial Correlation	0.2252	0.3363	0.0803	0.1020	0.0028
Regression on Fama-French Factors					
Intercept	0.0056	0.0061	0.0056	0.0044	0.0070
<i>t</i> -statistic	8.2853	11.849	10.709	9.0694	3.6298
U.S. Equity Risk Premium	0.0102	0.0080	-0.0067	-0.0167	
<i>t</i> -statistic	0.5924	0.6149	-0.5057	-1.3612	
SMB: Small minus Big	0.0404	0.0511	0.0465	0.0688	0.2117
<i>t</i> -statistic	1.7433	2.9192	-0.5057	4.1520	3.2089
HML: High minus Low Book to Market	0.0791	0.0210	0.0180	0.0689	-0.5847
<i>t</i> -statistic	2.8423	3.1582	0.8302	3.4620	-7.8073
R-square	0.0254	0.0412	0.0175	0.0728	0.1610
Durbin-Watson	1.546	1.314	1.805	1.999	2.086
Regression on Ibbotson Factors					
Intercept	0.0059	0.0064	0.0057	0.0047	0.0035
<i>t</i> -statistic	9.0363	12.816	11.163	9.9775	1.8688
U.S. Equity Risk Premium	-0.0296	-0.0259	-0.0206	-0.0408	
<i>t</i> -statistic	-1.6975	-1.9708	-1.5199	-3.2619	
US. Small Stock Premium	0.0477	0.0494	0.0349	0.0496	0.1771
<i>t</i> -statistic	2.7348	3.7614	2.5812	3.9670	3.6334
U.S. Bond Default Premium	0.1288	0.1266	0.1344	0.1680	0.7103
<i>t</i> -statistic	2.0496	2.6748	2.7578	3.7270	4.0631
U.S. Bond Horizon Premium	0.0815	0.0743	0.0440	0.0429	0.6805
<i>t</i> -statistic	2.9237	2.6748	2.0349	2.1458	9.5143
R-square	0.3622	0.0600	0.0358	0.0796	0.1975
Durbin-Watson	1.566	1.351	1.815	1.797	2.100

Table IV. Value at Risk

This table reports the daily and monthly Value-at-Risk percentiles based upon the strategies followed from 8/63 through 12/97

	top 5	top 20	20 / 100	All	Eq. Pre.	T-bills
Monthly Value at Risk						
1%	-0.0227	-0.0151	-0.0198	-0.0133	-0.0961	0.0023
5%	-0.0121	-0.0059	-0.0092	-0.0066	-0.0612	0.0026
10%	-0.0078	-0.0034	-0.0063	-0.0043	-0.0445	0.0029
15%	-0.0049	-0.0015	-0.0040	-0.0033	-0.0346	0.0031
20%	-0.0031	-0.0004	-0.0022	-0.0021	-0.0244	0.0035
Probability below 0	28.70%	22.10%	25.60%	32.30%	43.10%	0.00%
Min. historical observation	-0.1020	-0.0618	-0.0340	-0.0238	-0.2199	0.0021

Table V: Industry Sector Pairs Trading

The Table gives summary statistics for the excess return distributions for pairs trading portfolios by sector. We trade according to the rule which opens a position in a pair the day following the day that prices of the stocks in the pair diverge by two historical standard deviations and closes the position the day following the next crossing of prices.

Portfolio	top 5	top 10	top 20	All
Panel A: Utilities				
Mean Excess Return	0.0504	0.0489	0.0518	0.0478
Median	0.0471	0.0456	0.0484	0.0451
Standard Deviation	0.0545	0.0441	0.0378	0.0284
Kurtosis	0.8676	1.1015	1.3451	1.7417
Skewness	0.4187	0.6080	0.6548	0.7846
Minimum	-0.1009	-0.0806	-0.0362	-0.0221
Maximum	0.2826	0.2322	0.2154	0.1816
Observations with excess return < 0	16 %	11 %	6 %	3 %
Panel B: Transportation				
Mean Excess Return	0.0321	0.0318	0.0243	0.0223
Median	0.0335	0.0335	0.0225	0.0220
Standard Deviation	0.1029	0.0768	0.0600	0.0550
Kurtosis	4.2868	1.2467	0.8636	0.6279
Skewness	-0.3311	-0.1195	-0.1952	0.1560
Minimum	-0.5748	-0.2432	-0.2065	-0.1243
Maximum	0.4654	0.3158	0.2092	0.2081
Observations with excess return < 0	35 %	31 %	30 %	32 %
Panel C: Financials				
Mean Excess Return	0.0374	0.0371	0.0371	0.0352
Median	0.0371	0.0360	0.0348	0.0312
Standard Deviation	0.0825	0.0643	0.0489	0.0401
Kurtosis	1.5977	1.0806	0.8480	0.7545
Skewness	0.0403	0.2103	0.1249	0.4996
Minimum	-0.2737	-0.1527	-0.1391	-0.1015
Maximum	0.3729	0.3051	0.2316	0.1691
Observations with excess return < 0	32 %	27 %	21 %	18 %
Panel D: Industrial				
Mean Excess Return	0.0243	0.0272	0.0270	0.0336
Median	0.0291	0.0295	0.0283	0.0268
Standard Deviation	0.0741	0.0566	0.0465	0.0341
Kurtosis	0.6015	0.0969	0.1107	0.5594
Skewness	-0.3664	-0.0950	0.0058	0.8490
Minimum	-0.2332	-0.1521	-0.1192	-0.0333
Maximum	0.2366	0.1735	0.1582	0.1396
Observations with excess return < 0	34 %	31 %	27 %	14 %

Table VI. Sector Pairs Trading and Random Pairs Bootstrap

The top half of each Panel gives summary statistics of the 6-month excess returns on equally-weighted portfolios of the top n pairs formed by sector. We trade according to the rule which opens a position in a pair the day following the day that prices of the stocks in the pair diverge by two historical standard deviations and closes the position the day following the next crossing of prices. The Newey-West correction used to compute the standard errors uses six lags. The bottom half of each Panel summarizes the distribution of the bootstrapped distribution of the mean excess return of random pairs formation by sector. The p -value corresponds to the probability that bootstrapped profits exceed actual profits.

Portfolio	top 5	top 10	top 20	All
Panel A: Utilities				
Mean Excess Return	0.0504	0.0489	0.0518	0.0478
Standard Error (Newey-West)	0.0047	0.0041	0.0038	0.0031
t-statistic	10.70	11.93	13.70	15.43
Bootstrapped distribution of random pairs:				
mean	0.0009	0.0011	0.0014	
standard deviation	0.0071	0.0051	0.0035	
<p>p-value actual profits</p>	0.0000	0.0000	0.0000	
Panel B: Transportation				
Mean Excess Return	0.0321	0.0318	0.0243	0.0223
Standard Error (Newey-West)	0.0072	0.0059	0.0052	0.0050
t-statistic	4.4577	5.4085	4.6816	4.4797
Bootstrapped distribution of random pairs:				
mean	0.0008	0.0010	0.0011	
standard deviation	0.0075	0.0053	0.0036	
<p>p-value actual profits</p>	0.0000	0.0000	0.0000	
Panel C: Financials				
Mean Excess Return	0.0374	0.0371	0.0371	0.0352
Standard Error (Newey-West)	0.0066	0.0057	0.0046	0.0041
t-statistic	5.6737	6.4955	8.0168	8.6153
Bootstrapped distribution of random pairs:				
mean	-0.0010	-0.0008	-0.0005	
standard deviation	0.0071	0.0050	0.0036	
<p>p-value actual profits</p>	0.0000	0.0000	0.0000	
Panel D: Industrial				
Mean Excess Return	0.0243	0.0272	0.0270	0.0336
Standard Error (Newey-West)	0.0053	0.0043	0.0037	0.0037
t-statistic	4.5681	6.3123	7.2620	8.9602
Bootstrapped distribution of random pairs:				
mean	-0.0015	-0.0009	-0.0011	
standard deviation	0.0074	0.0052	0.0036	
<p>p-value actual profits</p>	0.0000	0.0000	0.0000	

Pair 4, Kennecott and Uniroyal
Starting Month 19620801

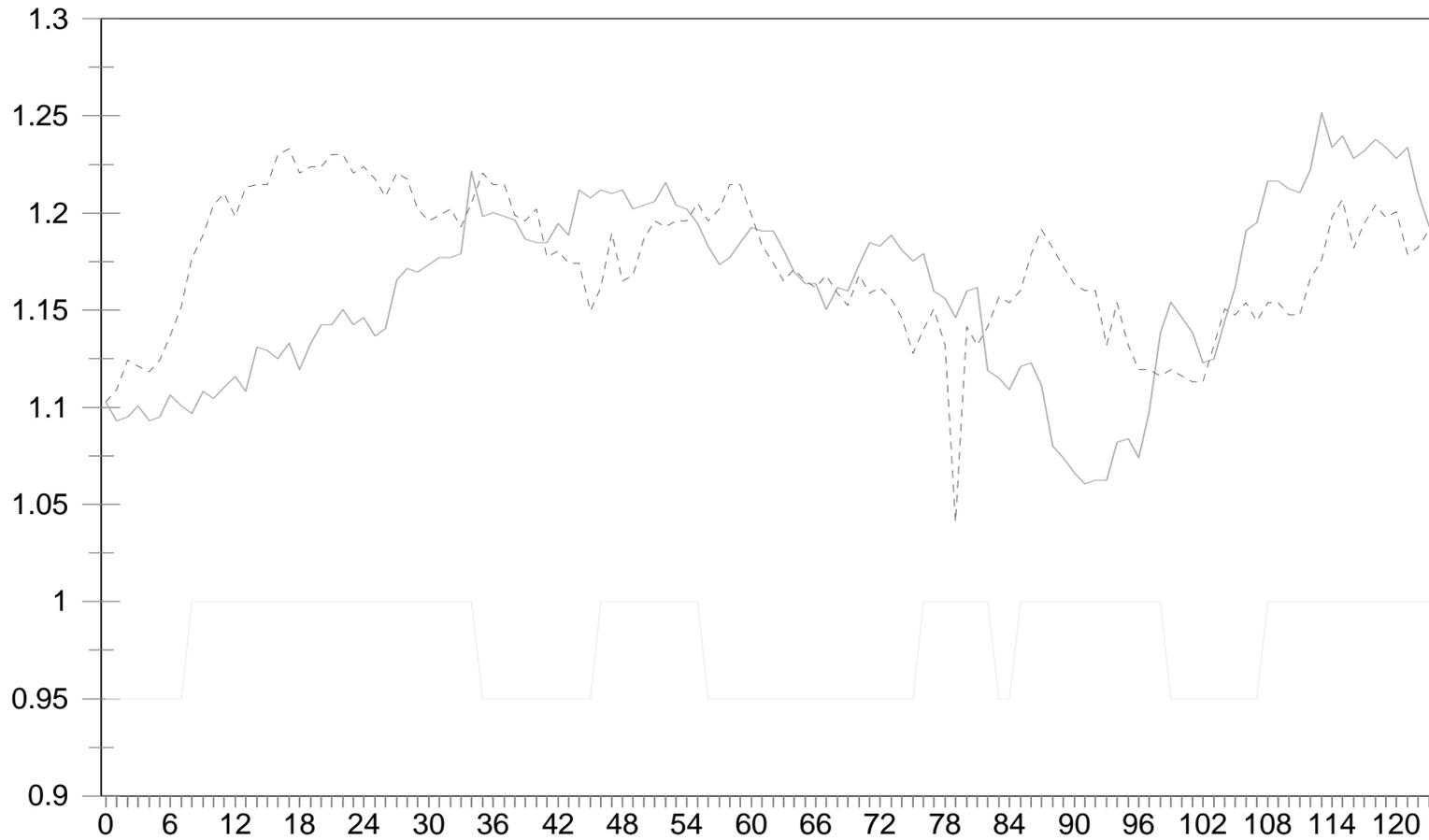


Figure 1

Top 20 Monthly Returns

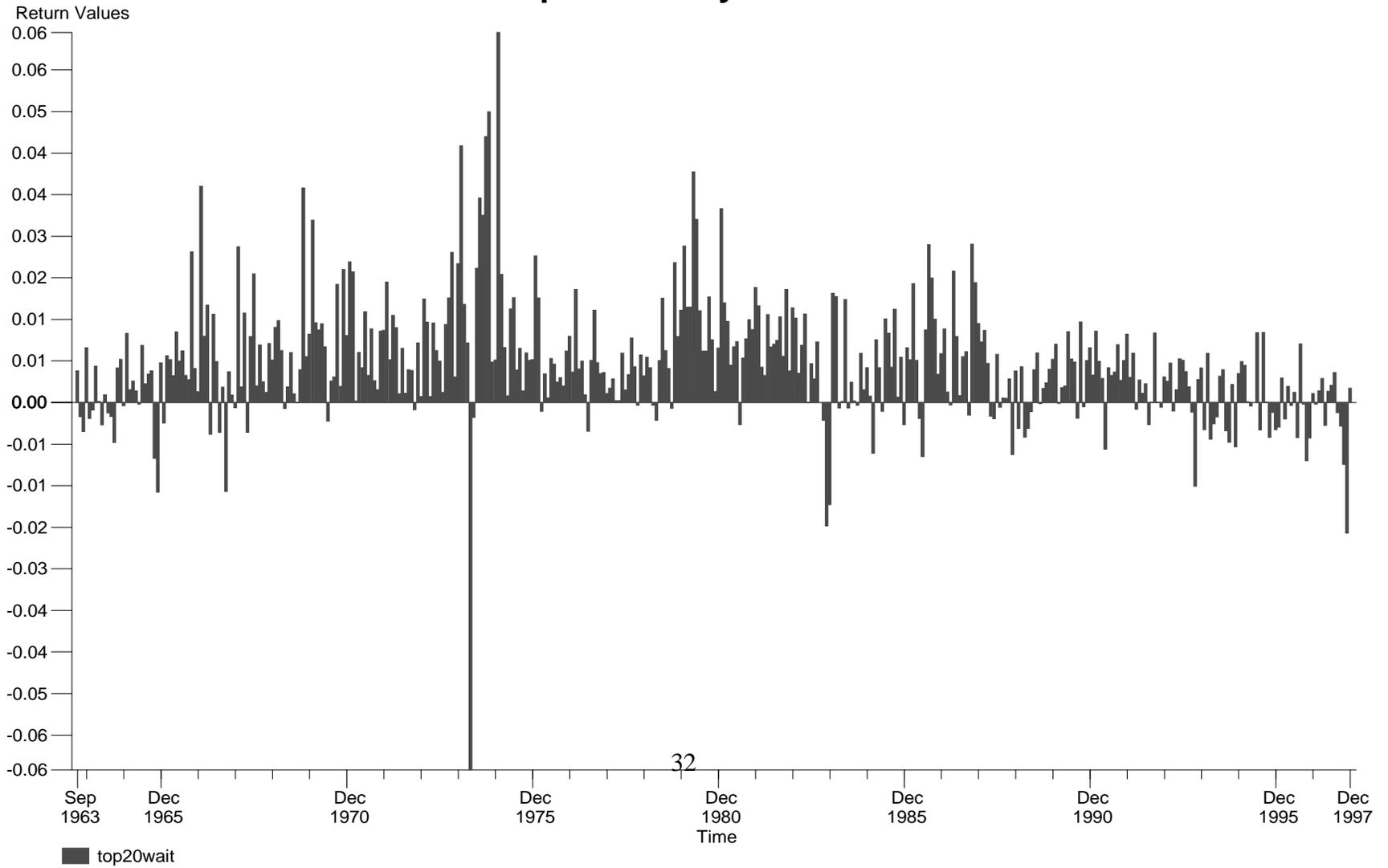


Figure 2

Top 20 & S&P Excess Return

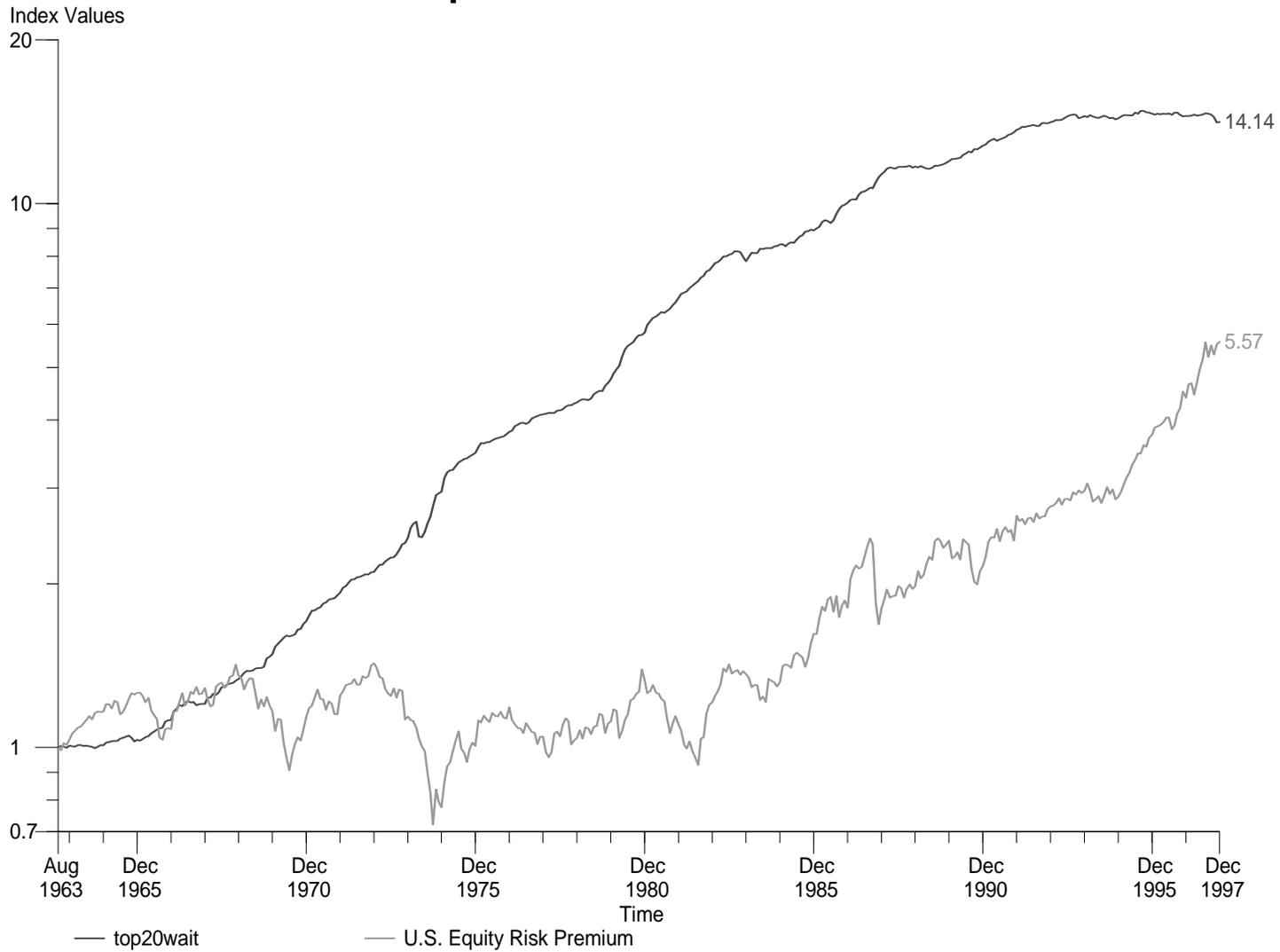


Figure 3