Lecture 23: Pairs Trading

Steven Skiena

Department of Computer Science
State University of New York
Stony Brook, NY 11794–4400

http://www.cs.sunysb.edu/~skiena
Pairs Trading

This strategy was pioneered by Nunzio Tartaglia’s quant group at Morgan Stanley in the 1980’s. It remains an important statistical arbitrage technique used by hedge funds. They found that certain securities were correlated in their day-to-day price movements.

When a well established price correlation between $A$ and $B$ broke down, i.e. stock $A$ traded up while $B$ traded down, they would sell $A$ and buy $B$, betting that the spread would eventually converge.

This divergence between pairs may be caused by temporary supply/demand changes, when a single large investor changes position in a single security.
Potentially Correlated Pairs

- Coca-Cola (KO) and Pepsi (PEP)
- Wal-Mart (WMT) and Target Corporation (TGT)
- Dell (DELL) and Hewlett-Packard (HPQ)
- Ford (F) and General Motors (GM)

Highly-correlated pairs often (but not always) come from the same sector because they face similar systematic risks.
Advantages of Pair Trading: Market Neutrality

The pairs trade helps to hedge sector- and market-risk. If the market or sector crashes, you should experience a gain on the short position and a negating loss on the long position, leaving your profit close to zero in spite of the large move. In a pairs trade, you bet on the direction of the stocks relative to each other.

Trading strategies which are independent of market movements are said to be market neutral. Pairs trading is a mean-reverting strategy, assuming that prices will revert to historical trends. Pairs trading is largely self-funding, since the short sale returns can be used to buy the long position.
Dangers of Pair Trading: News Reaction

The increased divergence between two stocks can be a rational response to news about one of the companies. Transaction costs (e.g. commissions, bid-ask spreads) can eat up the theoretical returns in such an active strategy. Pairs trading became less advantageous as more groups started doing it.

Hedge funds use algorithmic trading strategies that monitor for deviations in price, automatically buying and selling to capitalize on market inefficiencies. Such program trading requires fast reaction/execution time to take advantage of tighter spreads before other groups can.
Identifying Pairs

Correlation analysis can be used to identify interesting potential pairs, but beware of spurious correlation. Competing companies in the same sector make natural potential pairs. Certain companies have several classes of shares trading simultaneously (e.g. common and preferred) which should move largely in unison. Certain companies may be simultaneously traded on multiple exchanges or have international subsidiaries (e.g. Dutch and Royal Shell). A general or sector index portfolio (ETF) might be used as one component of a pair.
Example: Intel vs. SMH

The price ratio is flat long term but fluctuates day to day. The trailing dip would suggest buying Intel and shorting the semiconductor index (SMH). The ratio reverting to its historical mean suggests a logical time to exit both positions.
Achieving Market Neutrality

Stock A: $20 per share, Stock B: $10 per share
**Pairs trade:** Long 100 shares of Stock A: $2,000 Short 200 shares of Stock B: $2,000 Net long/short: $0
This is a true market-neutral position:
*Scenario 1: Both stocks rise 50 percent:* Stock A: $30 Stock B: $15
*Scenario 2: Both stocks fall 50 percent:* Stock A: $10 Stock B: $5
Example: Wells Fargo vs. Washington Mutual

**TABLE 1  WFC/WM STATS**

This pair meets the statistical requirements for a trade. It's highly correlated, and its price is two standard deviations below its one-year average. To trade the pair, buy Wells Fargo and sell Washington Mutual.

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Standard deviation (1 year):</th>
<th>0.044</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 days:</td>
<td>Average historical ratio (1 year):</td>
<td>1.48</td>
</tr>
<tr>
<td>100 days:</td>
<td>Current ratio:</td>
<td>1.39</td>
</tr>
<tr>
<td>1 year:</td>
<td>Divergence:</td>
<td>2 standard deviations</td>
</tr>
</tbody>
</table>

**TABLE 2  SPREAD FUNDAMENTALS**

Both companies’ fundamentals are similar, which is what you’re looking for in a pairs trade.

<table>
<thead>
<tr>
<th></th>
<th>Wells Fargo</th>
<th>Washington Mutual</th>
</tr>
</thead>
<tbody>
<tr>
<td>P/E</td>
<td>14.12</td>
<td>11.51</td>
</tr>
<tr>
<td>Dividend Yld:</td>
<td>3.4%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Market cap:</td>
<td>$106.2 billion</td>
<td>$42.1 billion</td>
</tr>
</tbody>
</table>
A Successful Pairs Trade

That said, Washington Mutual has since collapsed, so observed correlations do break down…
Why Might Pairs Trading Work?

Reversion to the mean requires a driving mechanism; pairs trading would not work were prices truly random walks. The Law of One Price (LOP) is the proposition that two investments with the same payoff in every state of nature must have the same current value. A strong historical correlation suggests they behave the same in a large number of states, and so should be priced the same. Another possible explanation is that changing risks of (rarely occurring) bankruptcies would cause reversion after the non-realization of the event plays out.
Academic Studies: Gatev, Goetzmann, Rouwenhorst


They tested a simple pairs trading strategy on daily data from 1962-2002.
A simple trading rule yields average annualized excess returns of up to 11% for self-financing portfolios of pairs.
We will look at their results as a model for a careful backtesting analysis of a trading strategy.
Strategy Specification

Pairs were formed based on price correlations over a 12-month period, starting every month. Each pair was then traded (possibly multiple times) over the next six months. The pair was “bought” when its ratio spread was outside two standard deviations of its 12-month spread. If normally distributed, this should happen about 5% of the time.
The position is closed when the ratio returned to its historical parity, or when the given six-month period ended. Pair trades were made for $1 in each direction.
Pair Formation

Stocks which did not trade on at least one day in the formation period were eliminated to ensure liquidity. A matching partner for each security was found by minimizing the squared deviations between the two normalized daily price series, where dividends were reinvested. Separate experiments were done to restrict pairs to the same industrial sector. The pairs were ranked by distance, and those within given rank traded.
Return Computation

That pairs open and close at various points during the six-month trading period complicates evaluation of return. Pairs that open and converge will have positive cash flows; Pairs that open but do not converge will have positive or negative cash flows on the last day of the trading interval when all positions are closed out. All positions are marked-to-market daily to analyze daily returns.
What is the Denominator?

To compute returns you must measure how much capital you need. They assume it costs $1 to go long and nothing up front to go short. Thus *committed capital* is equal to the number of pairs monitored. *Employed capital* is equal to the number of open positions at each time, and thus less than committed capital. Employed capital makes sense if the uninvested money can be used elsewhere in the hedge fund.
# Results

<table>
<thead>
<tr>
<th>Pairs portfolio</th>
<th>Top 5</th>
<th>Top 20</th>
<th>101 to 120</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. monthly return</td>
<td>0.01308</td>
<td>0.01436</td>
<td>0.01081</td>
<td>0.01104</td>
</tr>
<tr>
<td>Return on committed capital</td>
<td>0.00784</td>
<td>0.00805</td>
<td>0.00679</td>
<td>0.00614</td>
</tr>
<tr>
<td>Avg. monthly return (1 day delay)</td>
<td>0.00745</td>
<td>0.00895</td>
<td>0.00795</td>
<td>0.00715</td>
</tr>
<tr>
<td>Return on committed capital</td>
<td>0.00463</td>
<td>0.00520</td>
<td>0.00503</td>
<td>0.00396</td>
</tr>
</tbody>
</table>

Performance is sensitive to execution time or bid/ask spread and denominator.
Performance is relatively insensitive to which pairs are used.
Much of the return disappears when trading one day after optimal, which was studied to capture some aspect of transaction costs.
### Trading Statistics

<table>
<thead>
<tr>
<th></th>
<th>Top 5</th>
<th>Top 20</th>
<th>101120</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average price deviation trigger</td>
<td>0.04758</td>
<td>0.05284</td>
<td>0.07560</td>
<td>0.16888</td>
</tr>
<tr>
<td>Average pairs traded per period</td>
<td>4.81</td>
<td>19.30</td>
<td>19.41</td>
<td>1944.22</td>
</tr>
<tr>
<td>Average round-trip trades per pair</td>
<td>2.02</td>
<td>1.96</td>
<td>1.78</td>
<td>1.62</td>
</tr>
<tr>
<td>Average time pairs are open</td>
<td>3.75</td>
<td>3.76</td>
<td>3.98</td>
<td>3.97</td>
</tr>
<tr>
<td>Average weight of stocks in top five size deciles</td>
<td>0.91</td>
<td>0.91</td>
<td>0.79</td>
<td>0.62</td>
</tr>
<tr>
<td>Average weight of pairs from different deciles</td>
<td>0.66</td>
<td>0.69</td>
<td>0.75</td>
<td>0.82</td>
</tr>
<tr>
<td>Utilities</td>
<td>0.72</td>
<td>0.71</td>
<td>0.32</td>
<td>0.08</td>
</tr>
<tr>
<td>Transportation</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Financials</td>
<td>0.11</td>
<td>0.13</td>
<td>0.26</td>
<td>0.16</td>
</tr>
<tr>
<td>Industrials</td>
<td>0.15</td>
<td>0.14</td>
<td>0.40</td>
<td>0.73</td>
</tr>
<tr>
<td>Mixed sector pairs</td>
<td>0.20</td>
<td>0.22</td>
<td>0.44</td>
<td>0.33</td>
</tr>
</tbody>
</table>

The most highly correlated pairs provided little potential return (4.76%), because two $\sigma$ represents a small absolute spread.

Returns from randomly selected pairs were negative, so the initial correlation was important.
Most of the returns came from the short position, the effect is not simple mean reversion or reward for unrealized bankruptcy risk. Most pairs occurred within a sector because the systematic risk is so highly correlated. Certain sectors (particularly utilities) generated higher-correlated pairs.
Returns are positive and quite regular, but become smaller and more volatile after the pairs trading became a popular strategy. Of the $137 billion in hedge fund assets in 2000, $119 billion were employed in market neutral strategies.
Pairs Trading: Conclusions

Pairs trading is an important type of statistical arbitrage technique. There is good theoretical, experimental and practical experience that it can be a profitable strategy. Such strategies become much harder to employ once they become popular; hedge funds deploy them on a massive scale as part of algorithmic trading systems. The original Morgan Stanley group was initially extremely profitable but disbanded with losses only a few years later.
Types of Buy/Sell Orders

Brokers can typically perform the following buy/sell orders for exchange traded assets:

- *Market orders* request the trade happen immediately at the best current price.

- *Limit orders* demand a given or better price at which to buy or sell the asset. Nothing happens unless a matching buyer or seller is found.

- *Stop* or *stop-loss order* becomes a market order when a given price is reached by the market on the downside. This enables an investor to minimize their losses in a
market reversal, but does not guarantee them the given price.

- *Market-if-Touched order* (MIT) becomes a market order when a given price is reached by the market on the upside. This enables an investor to take profits when they are available, but does not guarantee them the given price.

The volume and distribution of stop and limit orders in principle contains information about future price movements. Theory argues against making such orders as giving away an option for no payoff, however, such orders are useful particularly for modest-sized investments.