Recent developments in high frequency trading and more specifically statistical arbitrage increase demand for capable technology. Traders not only need to automate complex trading algorithms but also to have them executed within microseconds. Today’s commercially available algorithmic trading engines meet the flexibility and scalability requirements in statistical arbitrage. Since these development platforms are often also offered as hosted solutions, market participants can enter the segment and compete on similar terms as larger participants without major investments in proprietary technology.

With arbitrage becoming more and more competitive, trading firms seek new opportunities in statistical arbitrage and strategies not solely dependent on the fastest technology. In the early years of the 21st century, statistical arbitrage was still closely equivalent to pairs trading. Stocks perceived as undervalued were bought against stocks perceived as overvalued, generating a profit when the relation between the stocks’ prices returned to normal. Analysis of cheap/expensive instruments is still at the core of statistical arbitrage in the position strategies, but today additional scalping strategies profiting on price inefficiencies in the financial market have evolved thanks to the rapid development of automated trading.

Statistical arbitrage:
Low-latency automation of complex trading strategies

AUTHOR: MARKUS KÄMPE
STATISTICAL ARBITRAGE STRATEGIES

There is no commonly accepted definition of statistical arbitrage, but many trading firms active in statistical arbitrage frequently use scalping and position strategies. Both types of strategies are generally market neutral, but they differ in terms of holding period and primary success factor. The holding period is the expected time between entering and exiting positions for a specific opportunity, and the primary success factor is the intended main profit source of the strategy. In the figure below the strategies are seen in relation to each other and arbitrage strategies based on holding period and primary success factor.

Scalping strategies

Scalping strategies try to profit from price inefficiencies just like arbitrage strategies do. A major difference is that in scalping strategies the instruments are not fungible. This means there is no arbitrage relation forcing the involved instruments to trade at a certain price level against each other. Instead the strategy tries to trade around the market equilibrium between highly correlated instruments. By buying below and selling above market equilibrium the scalping strategy generates profits. As soon as the strategy has entered positions by buying (selling) below (above) market equilibrium it immediately tries to reverse the trades by selling (buying) at a higher (lower) price. The need to exit the trades in the market makes the holding period for scalping strategies longer than for arbitrage strategies, but still it is very short. Holding positions overnight is generally not part of a scalping strategy.

Execution and therefore low-latency trading is very important in scalping strategies since they are trading on price inefficiencies, but with changing market conditions the analysis of market equilibrium is also important.

Position strategies

Position strategies are based on statistical analysis of market prices in listed instruments. Pairs trading is an example of a position strategy where an under-valued stock is bought and an overvalued stock is sold at the same time. By doing this on a market neutral basis the strategy builds a position with an isolated exposure to the relation between the two stocks without taking a view on the direction of the market in general. When the stocks return to trading at normal levels, the strategy realizes a profit and reverses the positions in the market. In this way position strategies have longer holding periods (driven by the mean reversion time) and overnight positions are frequently part of the trading strategy.

Correlated assets have been used for a long time to improve trading strategies and this also seen in statistical arbitrage. On a larger scale a trader can use position strategies to build a market neutral portfolio strategy with opportunities for optimizing trading on a portfolio level rather than trade level. In this way position strategies on a larger scale are more about building a fairly market neutral long/short portfolio and trading it over time.

Compared to arbitrage and scalping strategies the importance of execution is significantly lower for position strategies. Instead, a good analysis of incorrectly priced instru-
ments and the building of a suitable long/short portfolio are what make positions strategies successful. If, in addition, the execution of the strategy is efficient, it will further improve the profitability of the strategy.

TECHNOLOGY REQUIREMENTS IN STATISTICAL ARBITRAGE

Traders involved in statistical arbitrage typically mix scalping strategies with position strategies. The strategies are highly proprietary and, therefore, firms look for software products with open interfaces to automate the trading strategies and integrate with other systems like analysis tools or internal systems. The sophistication of the trading strategies makes them complex and requires full flexibility when automating them. A second major area of concern in statistical arbitrage is the ability to scale the business without significant additional costs or introducing additional trading latency.

Automation of complex trading strategies

The higher the level of sophistication in trading strategies the more complex decisions that need to be automated. Still, the complex decisions need to execute with as little additional latency as possible compared to simpler strategies.

If traditional trading models were about trading instruments in isolated pairs, today’s statistical arbitrage strategies work on more instruments simultaneously. A trader today can turn around a good trade in a stock by using a correlated stock, an index future, an ETF or some other instruments for hedging. Depending on the current market prices at the time of trade and the level of risk the trader is willing to take the actual hedge instruments can (and likely will) fluctuate during the day.

Statistical arbitrage strategies used to be based purely on top of book information, but today more and more strategies look further into the depth. The more sophisticated strategies not only use more depth levels but also analyze single orders at each level to extract more market intelligence to base trading decisions on. In addition, since statistical arbitrage strategies today tend to work on more instruments simultaneously the complexity of the strategies has grown rapidly during the last years.

Trading strategies that have been profitable for quite some time might suddenly lose their edge. Probable causes are that the market behavior has actually changed or that other market participants have identified and started to compete for the same opportunity. In this way statistical arbitrage strategies have to be dynamic and traders need to update them to adapt to changing market conditions or increased competition.

The above considerations are examples of needs and challenges for traders doing statistical arbitrage. In addition, good integration capabilities to statistical analysis engines and flexible implementation in safety precautions are important features of software products used for statistical arbitrage. The most advanced algorithmic trading engines provided by ISVs are suitable to meet the requirements from statistical arbitrage traders. The flexibility to automate even the most complex strategies; the open data models to represent loose relations between instruments; the technical infrastructure to cope with huge amounts of data and co-location; and the short time to market for changes when adapting trading strategies to changing market conditions are all features that make algorithmic trading engines a perfect match for larger scale statistical arbitrage traders.

Scalability in trading

Scalability issues are frequently seen as a trading business evolves and this is a key concern in statistical arbitrage. The completely automated nature of statistical arbitrage strategies makes the effort to add new strategies, products or trading venues very low. If done in the right way, the profit potential is substantial, but there is a significant risk that in the hunt for new opportunities the already profitable strategies might suffer. When increasing the number of concurrent trading strategies running,
there is a risk that the response times of the strategies overall will increase. Especially for scalping strategies, an increased latency due to scalability issues in technology might drastically change the profitability of a strategy and could turn a winner into a loser.

When looking at software products for statistical arbitrage from a scalability perspective the technical infrastructure of today’s algorithmic trading engines makes them a good fit. The ability to run strategies on many servers makes it possible to have a large number of concurrent trading strategies active. With a distributed set up supporting co-location at multiple trading venues simultaneously, latency increases can be kept at a minimum when expanding the trading activity. If, in addition, the algorithmic trading engine offers extensive market reach, there are few limitations when evolving the trading business to capture new opportunities.

"Especially for scalping strategies, an increased latency due to scalability issues in technology might drastically change the profitability of a strategy and could turn a winner into a loser."