Introduction:

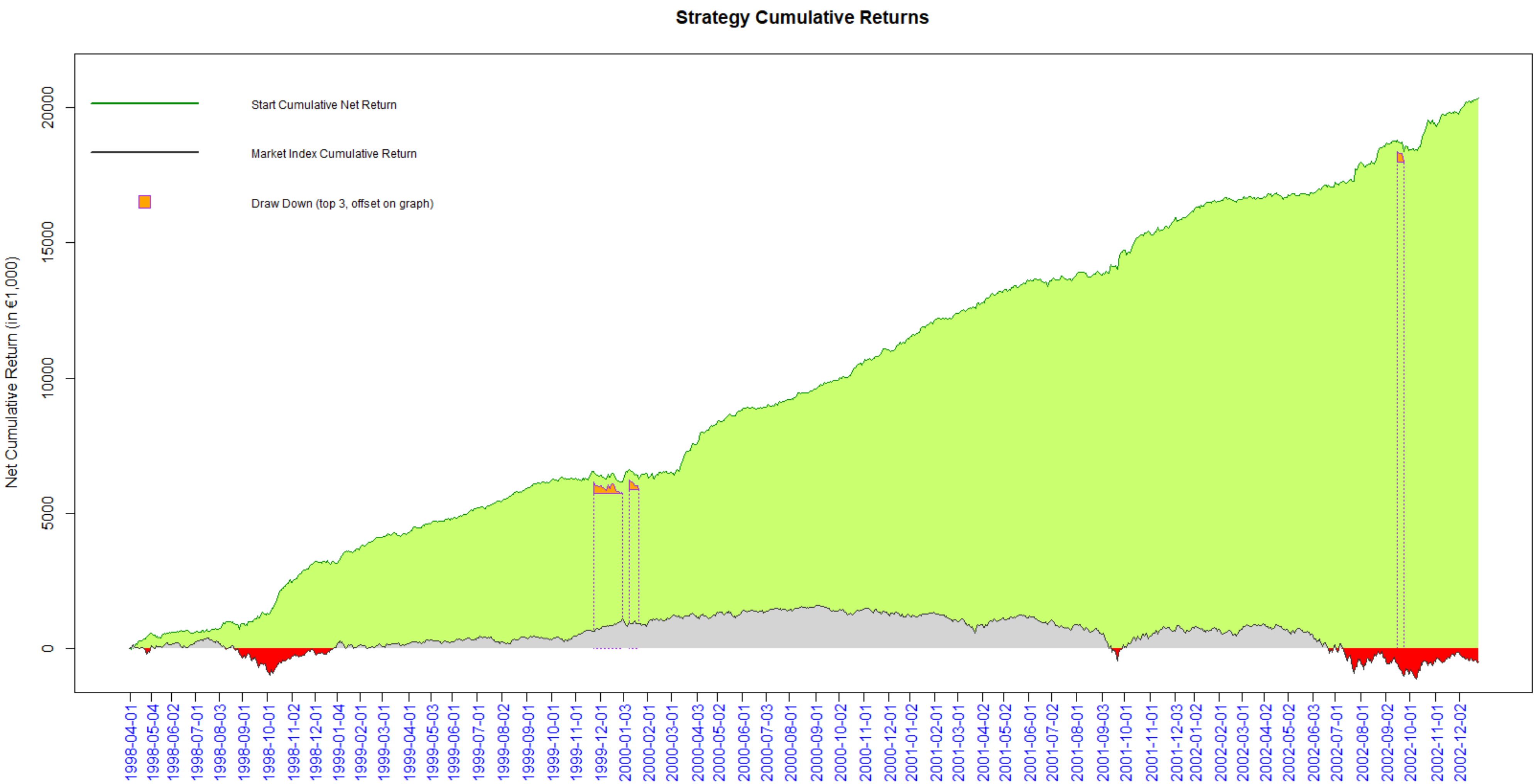
This is a daily statistical arbitrage strategy applied on European equities. Over the out of sample period the strategy achieved an annualized Sharpe ratio of 5.26 (more on this in the results section), virtually zero market beta and 100% long-short positions. The overall write up is designed to be brief with a redacted code base posted in the GitHub folder (this is not intended to be a fully comprehensive write-up). Underlined portions are pieces that are very relevant for the project. The write up is organized as follows: introduction, results, starting point and data, daily procedure and code development notes.

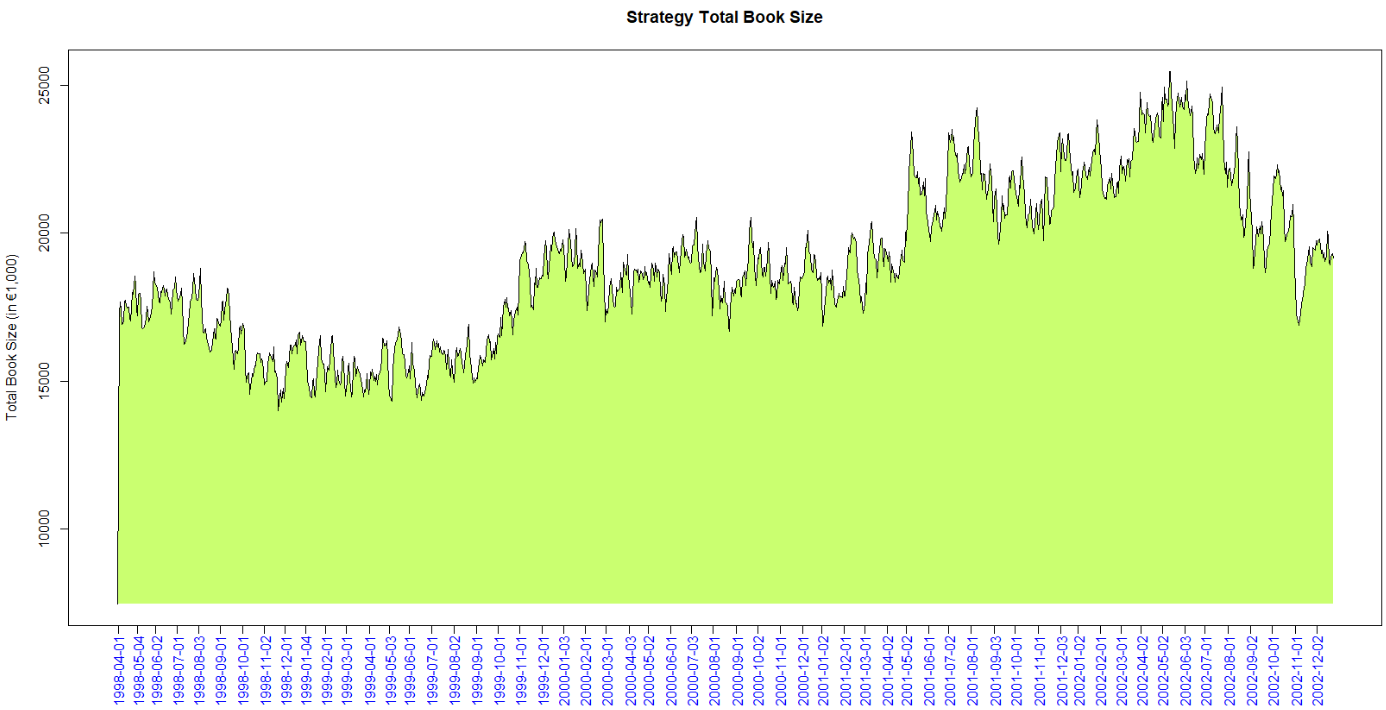
Results:

I assumed a margin requirement of 20% for any portfolio holding, a risk-free interest rate of 5% throughout the period and zero cost for shorting equities. I don't know the exact costs and rates for the period which warranted the assumptions. Including transactions costs, the Sharpe ratio is 5.26 using static trading and risk penalty parameters and not explicitly controlling for book size (Sharpe ratio is 7.25 when transaction costs are not included). A major portion of generating such a high ratio is the assumption that trading costs are half of the bid ask spread. For a full bid ask spread the Sharpe ratio is 3.26. The out of sample period is from 04/01/1998 – 12/31/2002. This was before the wide spread appearance of quantitative strategies and augmented by the fact that European equities were less efficient. The factors are very basic and have since been highly used in practice which makes me doubtful that the exact same strategy would work as well in current markets. However, I am well aware of newer and more robust factors that can be implemented in the strategy and intelligently blende. [Elaboration omitted on purpose]. The final correlation between the strategy and the market is .07 – a very negligible amount. Additionally, there is built functionality to control risk, trading and different asset classes correlation with the strategy on an ongoing basis. Lastly market impact and price movement weren't modeled which will certainly bring down the Sharpe ratio when implemented in practice.

Below are charts and information showing strategy performance, book size and trading throughout the out of sample period.

	Strategy	Market
Annual Sharpe ratio	5.26	-0.012
Market correlation	.07	1
Max drawdown	-411.19	-509.73
Longest drawdown	27 days	744 days
Average trade size	4,792.70	NA
Skewness	.6137	.0489
Kurtosis	5.9141	4.9326


Strategy Cumulative Returns
Start Cumulative Net Return
Market Index Cumulative Return
Draw Down (top 3, offset on graph)
Net Cumulative Return (in €1,000)
0
5000
10000
15000
20000



<u>Starting Point and Data:</u>

The universe consists of the constituents of 11 major European indices over the period from 1997 – 2002 with the constituents being updated at the end of every calendar year. For every day and equity combination I have data for market capital, active matrix, market to book ratio, price level, changes in analyst recommendation, monthly average transaction cost, total return index, industry classification and volume. The active matrix indicates which equites satisfied certain criteria to be part of the actively traded universe. The data structure is a matrix with date X equity identifier dimensions for each variable (market capital, price etc.) <u>The reason for the date period is because the data is easily accessible for me (getting recent data is proving to be a challenge).</u>

Alpha Signals:

There are a total of 4 signals that are used for relative performance prediction:

1. Regular 12 months momentum
2. Changes in analyst recommendation
3. Short term mean reversion and
4. Market to book

All the signals, except for analyst recommendation are linearly decayed and windsorized.

Covariance Matrix:

Using empirical covariance matrices can produce unstable and skewed results. I'm using a statistical method to estimate stable covariance matrices. While this is not essential for the strategy it does produce significantly better results. [Precise description of the estimation is omitted on purpose].

Daily Procedure:

On the first of every month in the out of sample period, which is 15 months from the first date in the sample, the active universe is determined for the month. The alpha signals are statically blended based on the data that is available up to but not including the day in question. Using quadratic programming techniques, the blended equity signals are optimized into portfolio weights with risk and trading considerations. All the trading is assumed to instantaneously take place at the very last second of the previous day (when the information is known). More advanced techniques for trading estimation and subsequent true return estimation are beyond the scope of the project due to time and data constraints. When an equity goes out of the active universe its holding is liquidated at the maximum allowable rate.

There are several constraints in the optimization that are crucial for realistic backtesting:

1. The maximum trade size on any given day per equity is the lesser of €50,000 or 1% of the daily average daily volume but not to exceed a maximum position size of the lesser of either €150,000 or 10% of the daily average volume.
2. The entire portfolio must sum up to zero.
3. The beta of the portfolio with respect to the market must be close to zero (a little off since this is a historical estimate).
4. The overall absolute net exposure per industry is not to exceed €100,000.

Code Development:

The entire development took just shy of a month. For 57 months, the code takes roughly 40 minutes to run due to stock universe expansion which makes it tougher to solve the optimization function. I tried to make the backtest as realistic as possible. As such I chased down every possible source of error I encountered. I am not claiming that this strategy will work just as well in current markets. Rather this is to showcase my ability and implementation work.