Extended Abstract

Growth Stock Portfolio Optimization Using a Multi-Objective EA

Jeffrey Christian Alexander Goedhuys

Industrial Engineering and Management

Examination Committee

Chairperson: Prof. José Rui de Matos Figueira

Supervisor: Prof. Rui Fuentecilla Maia Ferreira Neves

Member of the Committee: Prof. João Carlos da Cruz Lourenço

June 2016

Abstract

This work presents a method to build a portfolio of growth stocks with an above index average performance. The goal was to find good performance with a portfolio of growth stocks that are stable in terms of price development and relatively cheap. This was done using a multi-objective genetic evolutionary algorithm. It focusses on fundamental analysis, using quarterly data obtained from financial statements of the companies on the S&P500 index. The use of technical analysis was minimal. The training and simulations were done between the periods of 15-02-2012 and 13-02-2015. A short term strategy and a long term strategy were tested. The most consistent results were obtained using a long term multi-objective approach. Revenue growth appeared to have below average return when used as a main indicator. The best results for a stable growth stocks portfolio were found using profit margin growth as the main indicator. When profit margin growth was used as a main indicator, the return was well above market average.

Keywords: genetic evolutionary algorithm, multi objective, S&P500 index, growth stocks, fundamental analysis

1 Introduction

The stock exchange market as we know it today goes back to as far as the early 1600s. It was in 1602 when the Dutch East India Company (Dutch: Vereenigde Oostindische Compagnie) was founded. It offered stocks to fund their trading travels to South East Asia. Even though it was not the first company to offer stocks, it was the first stock to experience lively trading. In 1611 the world's first real stock exchange was opened in Amsterdam, The Netherlands.

Stocks can be characterized as value or growth stocks. Studies have shown that in the past value stocks outperform growth stocks on many occasions. In fact, in the period of 1975-1995 value stocks outperformed growth stocks in twelve of thirteen major markets (Fama & French, 1998). At first thought this might seem rather strange as growth stocks gain in price more rapidly than value stocks. The decline of growth stocks can however be explained by expectation errors by traders about future earnings prospects of these stocks. Earnings surprises are systematically more positive for value stocks (La Porta, et al., 1997).

So how do professional investors make decisions on what stocks to invest in? In a survey conducted with 692 fund managers in the United States, Germany, Switzerland, Italy and Thailand, it was concluded that in each country the fund managers used three kinds of analysis. These are technical analysis, fundamental analysis and flow analysis. For the long term, investment decisions were based on fundamental analysis. This type of analysis dominates the forecasting for periods down to two months. Only in the United States, technical analysis became more important in the two to six month forecasting period. When the forecasting period was in the range of weeks (less than two months), technical analysis was predominantly used by fund managers. In this range, the fundamental analysis was barely used. On a forecasting period of days, the flow analysis dominates. Flow analysis is the analysis of the trading orders. From this it can be concluded that fundamental analysis is important for decisions on the long term, technical analysis for short term and flow analysis for the very short term (Menkhoff, 2010).This suggests that using a combination of analysis is the best method for stock trading in general, according to professionals.

2 System Architecture

2.1 General Investment Concept

Before discussing all the indicators it is important to understand the goals of the investment. The general goal is to achieve an above index average profit with a low risk of losing money on the investment. The growth stocks that are looked for to achieve this goal should be relatively cheap (no high Price Earnings Ratio (PER)) and be subject to a stable above index average growth rate.

A good example of such a growth stock is the stock of Cap Gemini (CAP.PA) on the French CAC40 (Cotation Assistée en Continu) index. The development of this stock can be found in figure 1. At the moment of writing this stock has a PER of around 21, which means it is not an expensive stock.

Figure 1. Price development of CAP.PA (source: finance.yahoo.com)

2.2 Technical Analysis

The technical analysis part of the trading system is important, because it is there to spot trend reversals and trend direction of the stocks. The indicators to be used will be described in more detail below.

2.2.1 Moving Average

The moving average (MA) is an indicator aimed at spotting trend reversals. It performs well when the market is moving in a trend. Only the 200 day MA will be used in this system. The current price of the stock is compared to this 200 day MA. This is to check if a stock is in an upward trend when it is considered for the portfolio. If it is not the case, the MA will filter this stock out of consideration. The MA can be calculated with the following formula:

$$MA_t = \frac{P_t + P_{t-1} + \cdots + P_{t-n}}{n} \quad (1)$$

Where: P_t = Price of current day

n = number of days in the MA

2.2.2 Stop of Losses

The system must be protected by a trading rule against heavy sudden losses. The maximum allowable limit of loss depends on the investment strategy (short term, medium term or long turn). If the loss exceeds the limit, the system must sell the stock to protect itself from even greater losses. The limit of the stop-loss trigger can be linked to the entire index movement. This is to not lose large amounts of money on very short sudden drops in the macro-economic environment caused by, for example, a terrorist attack. The basic formula for the stop-loss trigger can be found in equation 2. The dynamic formula can be found in equation 3.

$$Sell\ Stock, if:\ P_{stock\ current} < (1 - stoploss) * \ P_{stock\ initial}\ (2)$$

$$if: P_{Index\ current} < P_{Index\ initial}$$

$$Sell\ stock, if: P_{stock\ current} < \left(\frac{P_{index\ current}}{P_{index\ initial}} - stoploss\right) * P_{stock\ \ initial}(3)$$

Where:

$P_{stock\ initial}$ = Price of stock when added to portfolio

$P_{stock\ current}$ = Price of stock at the current day

$P_{index\ initial}$ = Price of index (S&P500) at start of trading quarter

$P_{index\ current}$ = Price of index (S&P500) at the current day

Stoploss = maximum allowable loss

2.3 Fundamental Analysis

The fundamental analysis uses financial ratios to help understand the financial state of a company. The fundamental indicators used in this work are presented in table 1. To keep the results that are presented later more organized, the fundamental indicators are numbered as they are on the chromosomes used in the Evolutionary Algorithm (EA).

Gene Number	Fundamental Indicator
1	Revenue Growth
2	Earnings per Share Growth
3	Profit Margin Growth
4	Adjusted Asset Turnover Growth
5	Debt Growth
6	Revenue vs. Debt Growth
7	Company Sector
8	Profit Margin (Sector Specific)
9	Return on Equity (Sector Specific)
10	Asset Turnover Ratio (Sector Specific)
11	Debt Ratio (Sector Specific)

Table 1. Fundamental Indicator number table

3 Evolutionary Algorithm

A genetic EA will be used to optimize the use of the fundamental indicators. Based on this the best stocks will be selected to enter the portfolio.

3.1 Chromosome Structure

A solution can be represented on a chromosome. The chromosome contains all the parameters to get to the solution. These parameters are stored on the genes of that chromosome. In this trading system there are eleven indicators, which could form the first part of the chromosome. The rest of the chromosome will contain the other trading system's parameters, such as: PER-filter, MA-filter, Stop-loss and cash distribution.

The trading system used for this research, however, focusses on using only four of the eleven weighted indicators. The thought behind this concept is, that it is more practical as a trader to only focus on a few indicators. Since there are only four indicators, the minimum weight of a gene will be 1, with a total weight of 100 for all four genes used. An example of what a random chromosome of the trading system would look like, can be found in figure 2.

Weighted Indicators				Trading System Parameters			
9 Weight: 61	8 Weight: 6	6 Weight: 10	2 Weight: 23	PER-Filter	MA-Filter	...	Stop-Loss

Figure 2. Chromosome with four indicators

3.2 Evolutionary Process

To optimize the profit of the portfolio, the EA must find the best combinations of genes (weighted indicators). This is done during a training period, where the chromosomes are evaluated by their fitness.

3.2.1 Individual Fitness

The fitness of an individual chromosome is measured by the profit of the portfolio this individual leads to. Each indicator (I) has a score for each stock ranging between -2 and 2, with 2 being the best score possible. The indicator's score is then multiplied with its Weight (W). The full calculation for a stock's score can be found in equation 4 below. The fitness of an individual is the average profit of the 20 stocks with the highest scores.

$$Score\ Stock_n = I_v * W_v + I_x * W_x + I_y * W_y + I_z * W_z \quad (4)$$

3.2.2 Crossover & Reproduction

After the creation of the initial population, it must form a new and more optimal generation. This is done by combining the indicator genes of two parent individuals from the current population with each other. They can be combined in different ways to form the new children chromosomes. In this case, the goal is to form two children to have a population pool with a total of 256 individuals. From this pool the 128 individuals with the highest fitness will survive and form the population of the next generation. To illustrate the crossover process, an example is shown in figure 3.

3.2.3 Mutation

After the best 128 individuals are selected, about 5% of them will be given a mutation. This is done to attempt to generate more genetic diversity among the population. A mutation here means changing one of the indicator genes to another random value, which is not yet present on that particular chromosome. The gene to be changed is selected at random. The concept of elitism is used to preserve the best solutions. The best 5% of the population will be excluded from selection for mutation. The algorithm will pick the individuals for mutation at random from the remaining available individuals (position 7 to 127 on the list).

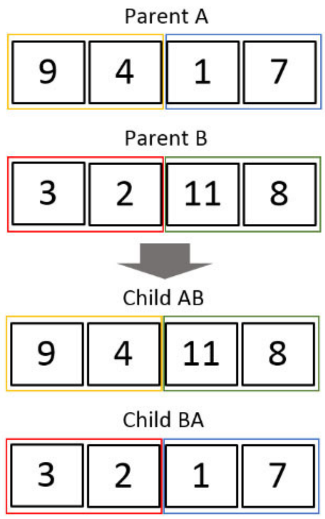

Figure 3. Crossover Parent A with Parent B

4 Simulation Results

In simulation, the portfolio can hold a maximum of 20 different stocks. There are two simulation strategies, one is short term and the other is long term. The short term system will form a totally new portfolio each quarter and has multiple training periods of one quarter. The long term system is looking to build a portfolio over time and depending on their evaluation, it will hold stocks in the portfolio after each quarter and add new ones to the existing portfolio. The long term system has a single training period of one year. Evaluating the trading of the systems will be done using the ROI and the Information ratio. The benchmark used for calculating the Information ratio is the S&P500 index. For the multi objective approach, the Pareto dominance concept is used to select which chromosomes will be simulated.

4.1 Single Objective Short Term

The first simulations were done for the short term system, optimizing for the single objective ROI. Training periods had a length of one quarter and each new trading period, the training results of the preceding quarter were added to the training results. From this the new indicators were chosen based on which four indicators were the most present in terms of weight. In total three simulations were done, creating the dynamic chromosomes S1 to S3.

Simulation of these chromosomes started on 15-05-2012 and ended on 15-02-2015. The results are presented in growth rates for each quarter in table 2. The total growth, tracking error and Information ratio can be found in table 3.

	Q1 ended 15-8-2012	Q2 ended 15-11-2012	Q3 ended 15-2-2013	Q4 ended 15-5-2013	Q5 ended 15-8-2013	Q6 ended 15-11-2013	Q7 ended 15-2-2014	Q8 ended 15-5-2014	Q9 ended 15-8-2014	Q10 ended 15-11-2014	Q11 ended 15-2-2015
S1	0.033	-0.023	0.137	0.131	0.004	0.117	0.042	0.021	0.044	0.027	0.067
S2	0.032	-0.019	0.137	0.091	-0.019	0.104	0.030	0.056	0.049	0.008	0.048
S3	0.034	-0.015	0.055	0.121	-0.017	0.127	0.009	0.052	0.042	0.050	0.074
S&P500	0.056	-0.037	0.123	0.091	0.002	0.082	0.022	0.018	0.045	0.043	0.028

Table 2. Simulation growth rate results single objective, short term

Chromosome Combination	Total ROI	Tracking Error	IR
S1	0.772	0.024	8.078
S2	0.637	0.023	2.710
S3	0.664	0.035	2.510
S&P500	0.576	0.000	X

Table 3. Simulation results single obj. short term

The results show that generally speaking each chromosome outperformed the index. In the first quarter each chromosome performed worse than the index, but this was expected because the chromosomes only had training data of a single quarter and that quarter was not likely to be similar to the next. Overall the short term strategy with rebuilding an entire portfolio each quarter does not perform greatly. For a period of eleven quarters, the total ROI of each chromosome compared to the index seems a bit on the low side, especially that of S2 and S3.

4.2 Single Objective Long Term

The next simulations were done using single objective optimization for ROI. This time the simulation system is looking for stocks on the long term. The training period of one year was done from 15-05-2012 till 15-05-2013. The simulation period started on 15-5-2013 and ended on 15-2-2015. The simulation results are presented in growth rates in table 5.

Five chromosomes were picked from the population which went through 20 iterations in the genetic EA. The chromosome selection is based on the maximum value of ROI, but the chromosomes must be at least clearly different from another in terms of indicators or weight distribution. The composition of the five selected chromosomes can be found in table 4.

It must be noted however that the chromosome with the highest ROI was not selected here, because this is a non-dominated solution for the multi-objective approach. The simulation results of that chromosome can be found in the next section as chromosome M13.

Chromo-some #	ROI	1st Gene	2nd Gene	3rd Gene	4th Gene	Weight 1st Gene	Weight 2nd Gene	Weight 3rd Gene	Weight 4th Gene
L1	0.549117	5	1	4	2	27	54	8	11
L2	0.538764	1	5	2	6	56	18	10	16
L3	0.526488	6	1	3	2	30	38	18	14
L4	0.519741	1	5	2	6	29	29	15	27
L5	0.508123	1	5	10	11	41	33	22	4

Table 4. Chromosome composition, single objective long term

Chromo-some #	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Compounded Growth	Tracking Error	IR
L1	-0.003	0.086	-0.155	0.034	0.047	-0.015	0	-0.025	0.077	-3.727
L2	0.011	0.099	-0.051	-0.013	0.014	0.025	0.041	0.126	0.037	-3.723
L3	0.016	0.074	-0.04	0.007	0.076	-0.007	0.008	0.138	0.037	-3.441
L4	-0.029	0.081	-0.036	-0.012	0.005	-0.072	0.071	-0.002	0.06	-4.407
L5	-0.027	0.086	-0.011	0.04	0.053	0.093	0.05	0.313	0.03	1.628
SP500 Index	0.002	0.082	0.022	0.018	0.045	0.043	0.028	0.264	0	X

Table 5. Simulation quarterly growth rates, single objective long term

The results show that four out of five chromosomes were beaten by the index in terms of compounded growth. Each chromosome has revenue growth as its main indicator, but only L5 has a positive ROI compared to the benchmark index. Possibly this indicator comes with added risk. Before drawing any conclusions, the multi-objective simulation results will be checked. This may give a better insight in the characteristics of certain indicators.

4.3 Multi Objective Long Term

For the Multi Objective Long Term approach, the training period was done from 15-05-2012 to 15-05-2013. Generating the Pareto front was done by extracting the population from different generations of the EA. The population was taken from the initial population (generation zero), the first generation, the fifth generation, the tenth generation and finally the twentieth generation. All individuals/chromosomes were then plotted into a graph. The result can be seen in figure 4.

The red dots in the graph represent the non-dominated individuals and the green dots are the dominated individuals. The non-dominated individuals that appear to be the most interesting because of their ROI relative to their variance, are connected by a grey line. These are the individuals that will be simulated over the period of 15-05-2013 till 15-02-2015. More details about the composition of these individuals can be found in table 6.

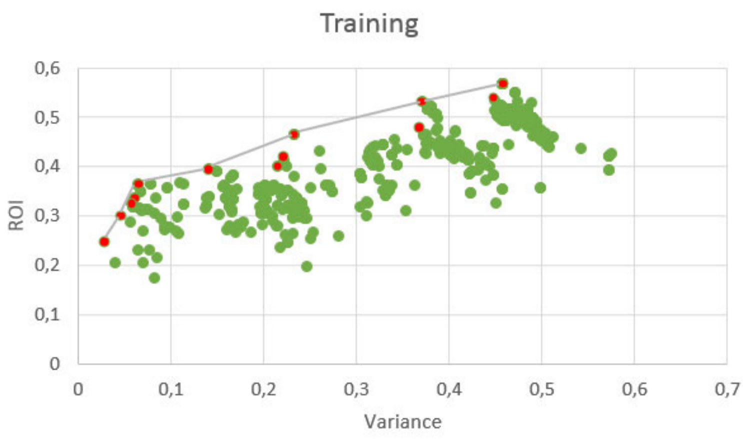

Figure 4. Training results plotted

	Variance	ROI	1st Gene	2nd Gene	3rd Gene	4th Gene	Weight 1st Gene	Weight 2nd Gene	Weight 3rd Gene	Weight 4th Gene
M1	0.027623	0.247262	3	1	8	10	83	1	3	13
M2	0.044887	0.299454	3	7	2	6	77	18	2	3
M5	0.064502	0.365611	2	1	3	6	12	24	53	11
M6	0.140695	0.394648	7	5	1	6	55	8	6	31
M9	0.232294	0.464276	5	1	2	3	13	32	17	38
M11	0.371682	0.532761	1	6	2	3	68	15	9	8
M13	0.456967	0.567565	5	1	2	4	44	32	19	5

Table 6. Non-dominated Solutions from Training period

The results of simulating all these chromosomes can be found in table 7. The results are presented in growth rates, to make them easy to compare to other results. The tracking error and Information ratio (IR) can be found in the last columns of the table. The growth rates of the S&P500 index are shown at the bottom, which was the benchmark to compare the chromosomes performances with.

Chromo-some #	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Compounded Growth	Tracking Error	IR
M1	-0.018	0.079	0.049	0.029	0.033	0.064	0.072	0.348	0.025	3.374
M2	-0.005	0.077	0.032	0.029	0.071	0.049	0.050	0.340	0.016	4.888
M5	0.004	0.082	0.020	0.050	0.104	0.032	0.042	0.382	0.028	4.151
M6	0.021	0.080	0.014	0.042	0.036	0.029	0.040	0.291	0.016	1.675
M9	0.022	0.076	0.040	0.048	0.064	0.050	0.045	0.398	0.020	6.746
M11	0.008	0.085	-0.086	0.008	0.073	-0.021	-0.014	0.045	0.056	-3.950
M13	0.032	0.092	-0.020	0.051	0.033	0.060	0.029	0.309	0.027	1.637
S&P500	0.002	0.082	0.022	0.018	0.045	0.043	0.028	0.264	0.000	X

Table 7. Multi Objective simulation results, presented in growth rate

The results show that almost every chromosome has a positive Information ratio and thus outperformed the S&P500 index. Only one chromosome performed poorly, which was chromosome M11.

5 Conclusions

This work presented a trading system with the goal to manage a portfolio of stocks with above average return. The desired stocks were described as stable growth stocks that are relatively cheap. The focus was on using fundamental indicators, which were calculated using quarterly financial data from companies on the S&P500 index. The combination of fundamental indicators was optimized with a genetic EA, which always used a combination of four weighted indicators. Optimization was done during training periods of one quarter or one year, depending on the investment strategy. The two applied investment strategies were a short term and a long term one. The conclusions that can be drawn from the work are presented below.

- The multi-objective long term approach gave the most consistent results, with six out of seven chromosomes beating the index. The best chromosome outperformed the index with 39.8% versus 26.4% of the index.
- Single objective long term simulation gave the worst results, with only two out of six chromosomes beating the index.
- Single objective short term performed decently, with all three simulated chromosomes beating the index. The best performing chromosome had a ROI of 77.2% versus 57.6% ROI of the index. The results were less consistent than the multi-objective long term approach though.
- It is possible to find stable growth stocks using only four fundamental indicators in an optimized combination.
- Profit margin growth works the best as a main indicator for stable growth stocks.
- Revenue growth gave inconsistent results as a main indicator when selecting stocks to enter the portfolio.

References

Fama, E. & French, K., 1998. Value versus Growth: The International Evidence. *The Journal of Finance,* pp. 1975-1999.

La Porta, R., Lakonishok, J., Shleifer, A. & Vishny, R., 1997. Good new for value stocks: Further evidence on market efficiency. *The Journal of Finance,* pp. 859-874.

Menkhoff, 2010. The use of technical analysis by fund managers: International evidence. *Journal of Banking & Finance,* pp. 2573-2586.