GP-evolved Technical Trading Rules Can Outperform Buy and Hold

Lee A. Becker & Mukund Seshadri
Department of Computer Science
Worcester Polytechnic Institute
Worcester, MA 01609 USA
{lab | mukund}@cs.wpi.edu

Abstract

This paper presents a number of experiments in which GP-evolved technical trading rules outperform a buy-and-hold strategy on the S&P500, even taking into account transaction costs. Several methodology changes from previous work are discussed and tested. These include a complexity-penalizing factor, a fitness function that considers consistency of performance, and coevolution of a separate buy and sell rule.

1 Introduction

Previous attempts [1][2] to use GP for acquiring technical trading rules have not been able to establish that GP-evolved technical trading rules could outperform a buy-and-hold strategy if transaction costs were taken into consideration. This study describes an approach to the genetic programming of technical trading rules which has evolved rules that can outperform a buy-and-hold strategy, at least if dividends are excluded from stock returns.

Our approach has a number of significant changes from that of Allen & Karjalainen[1], which was adopted by Neely[2] with the additional measurement of risk-adjustment. These include

- using monthly as opposed to daily data,
- reducing the operator set and increasing the number of derived technical indicators,
- using a complexity-penalizing factor in the fitness function to avoid overfitting as well as improve comprehensibility,
- utilizing a fitness function which considers the number of periods a rule performs well, and not just its total return or average excess return, and
- using co-evolution of a specialized buy rule and a specialized sell rule.

The remainder of the paper is organized as follows. Previous work on GP-evolved technical trading rules for market-timing on the S&P500 and some issues it raises are discussed. Our approach is described in detail and contrasted to that of Allen & Karjalainen [1] and Neely[2] and the design of experiments involving each of the three latter changes is presented. The results of our experiments are then compared with those of previous studies and discussed.

2 Genetic Programming for Technical Trading Rules

Genetic Programming (GP) has been applied to a wide range of problems in finance[3]. There have been a number of attempts to use GP for acquiring technical trading rules, both for Foreign Exchange Trading [4][5] and for S&P500 market timing by Allen & Karljalainen [1] and Neely[2]. These latter studies were not able to establish that GP-evolved technical trading rules could outperform a buy-and-hold strategy if transaction costs were taken into consideration. Still there have been several studies [6],[7],[8] which have shown that technical trading rules do have forecasting ability. These results are relevant to the Efficient Market Hypothesis [9][10][11]. Neely[2], whose risk-adjusted, ex-ante GP-evolved technical trading rules did not outperform a buy-and-hold strategy on the S&P500, concluded that the results were consistent with market efficiency.

We now consider the approach of Allen & Karljalainen[1] in detail; it will be referred to as AK below. AK used daily data for the S&P500 index from 1929 through 1995. They divided up the data into successive in-sample periods, consisting of 7 years, further sub-divided into 5 years training and 2 years for selection for each set of ten runs. The selection period is used to avoid overfitting. Each generation the fittest rule in the population is applied to the selection period, and if its performance is better than the best rule so far, it becomes the new best rule. The out-of-sample periods go from the end of the in-sample through 1995.

The operators used by AK include arithmetic operators and real valued functions (+, -, /, *, average of a certain period, max and min of a certain period, and absolute difference) and Boolean operators and functions (<, >, and, or, not, if-then-else), the terminal nodes in the expression tree included real constants between 0..2, Boolean constants (true, false), price; and lag, which causes the price to be lagged by a certain number of days.

AK used a population of 500, with a size limit of 100 nodes and depth of 10, and with a maximum of 50 generations. AK used one-month T-bill rates as risk-free rate when out of the market, and used one-way transaction costs of .25%. This seems reasonable as an estimate of combined bid-ask spread and commissions [10]. AK did not consider dividend yield, which obviously underestimates returns more for the buy-and-hold than for the technical trading rules, which are sometimes out of the market. Dividend yield has been estimated to be 0.016% per day for the DJIA[10], but would be less for a broader index like the S&P500.

AK's out-of-sample results were an average yearly excess return over buy-and-hold of –0.0205 (-0.0323), average time in the market 57.5% (57.5%), average number of trades per year of 3.8, and an average spread in daily returns between in and out of market periods of +0.000714 (+0.001413). The values in parenthesis were calculated on the training periods after 1959, which would be the best comparison to our study.

Neely[2] added risk-adjustment, measured in a number of ways, and even though this improved the relative attractiveness of the rules, it was not enough to outperform buy-and-hold.

3 Our Approach and Experimental Design

We made a large number of changes to the regime used by AK, some of which are more significant than others. The changes are described below, and for three of them (complexity-penalizing factor, 'consistency-based' fitness, and coevolution) experiments were designed to determine their effectiveness.

We used monthly rather than daily data. This provides fewer opportunities to trade and might be expected to reduce the number of transactions and thus transaction costs. Related to this change, we used a single in-sample period rather than a number of shorter in-sample periods. This was necessary, given the reduced number of data points, to experience a variety of market and economic cycles.

Our language for expressing technical trading rules was significantly different from that of AK. In particular, we drastically reduced the number of operators and added a number of derived technical indicators. The non-leaf nodes in the expression tree allowed only the logical operators (AND, OR, NOT) and the arithmetic comparison operators (>, <). The leaf nodes include the following types of technical indicators.

Prices: Opening, Closing, High, Low of the current month.

Moving Averages: 2,3,6, and 10 month.

Rate of Change: 3 and 12 month.

Price Resistance Markers: two previous 3-month moving average minima and two previous 3-month moving average maxima.

Trend Line Indicators: a lower resistance line based on the slope of the 2 previous minima and a upper resistance line based on the slope of the 2 previous maxima.

Both changes serve to improve the comprehensibility of the evolved rules, while the addition of these commonly used [6], [12] [13] derived indicators is a way of incorporating domain knowledge. In comparison to the lower level arithmetic operators and functions, which perhaps in theory could evolve these derived indicators, these indicators bias the search, while at the same time producing rules which are more comprehensible. It should be noted that the trend line indicators differ from the more common trend lines [12/13] which are based on price. These are based on maxima and minima of 3-month moving averages, and one could use longer moving averages to capture longer term trends.

We used a standard GP algorithm[14]. The software for this work used the GAlib genetic algorithm package, written by Matthew Wall at the Massachusetts Institute of Technology[15]. We used a population of 500 trees and ran for 100 generations. It was a steady state GA with half the population being replaced each generation.

We used S&P500 data from January 1954 through December 2002. With the exception of the Prices, all the technical indicators are derived. The latter were pre-processed, and the need to assure the two previous minima and maxima required data from 1954-1959. We trained on data from 1960-1990 and tested on data from 1991-2002. For testing, we used 'expensive' transaction costs of .5% for each buy or sell. For months when we were out of the market, we credited 1/12 of the interest rate on 3-month T-bills, which we used due to data availability. For our baseline, we used a fitness function of the portfolio value produced by using the rule at the end of the in-sample period.

Experiment 1 contrasted the baseline with rules produced by a fitness function that incorporated a complexity-penalizing (C-P) factor. The purpose of this factor was two-fold: to increase the comprehensibility of the tree by reducing its size and to avoid overfitting and thereby increase performance on the out-of-sample period. These issues are discussed in more detail in [16].

Experiment 2 introduced a different fitness function. Instead of using the total portfolio value, modified by the complexity-penalizing factor, the fitness was calculated as the number of periods with well-performing returns, modified by the factor. The periods were 12 months and a rule was taken to be well-performing if it beat or were equal to both the buy-and-hold return and the return on the risk-free interest rate. We refer to this fitness function as consistency of performance.

Experiment 3 involved the coevolution of a pair of trading rules, a buy rule and a sell rule. When out of the market, only the buy rule is considered; when in the market, only the sell rule is considered. The rules are considered to be separate species; buy rules only crossover with other buy rules, while sell rules only crossover with other sell rules. This is an example of cooperative evolution[17], and we discuss different approaches to the coevolution of technical trading rules elsewhere [18].

4 Results and Discussion

The first table below compares the technical trading rules generated by the baseline to those with the C-P factor. It includes the best technical trading rule generated by each of ten runs of 100 generations. The last row gives the average values. It is obvious that the size in terms of the number of nodes as well as the depth of the trees is much larger when the factor is not used.

Baseline: w/o C-P Factor				With C-P Factor			
Size	Dpth	1960 - 1990	1991 - 2002	Size	Dpth	1960 - 1990	1991 - 2002
224	19	15195	3491	15	5	10976	3128
542	88	13856	3300	15	5	13690	2006
530	45	20265	1469	3	2	8762	3377
41	11	16832	2023	15	5	13981	2003
200	38	12818	1835	3	2	8762	3377
967	87	14984	1834	12	5	14788	1685
178	28	19605	1626	15	4	15078	2096
45	11	16367	1977	3	2	8762	3377
596	68	23702	2089	15	5	9128	3697
111	20	17637	1764	3	2	8762	3377
343	***42***	***17126***	***2141***	***10***	***4***	***11269***	***2812***

In considering the performance of the trees, one should note that with a buy-and-hold strategy if one invested $1000 in the S&P500 at the beginning of 1960, it would have grown to $5457 at the end of 1990. This was the in-sample period on which the GP was trained. For the out-of-sample period, $1000 would have grown to $2638. In comparing the second table with the first, we can see that without the C-P factor the performance was significantly better in the in-sample period, but worse in the out-of-sample period. This indicates overfitting. In contrast, the rules learned with the C-P factor were less overfit and their average performance exceeded the buy-and-hold strategy for the out-of-sample period.

In the following tables (in-sample and out-of-sample) are the results using the consistency of performance fitness function. The additional fields are given for comparison with AK. They indicate the number of roundtrip trades per year, the number of months in vs. out of the market, and the average monthly return of the market when the trading rule was in vs. out of the market.

RT trans. per year	Mnths in mrket X/372	Avg. market return when IN	Avg. market return when OUT	Prtflio value
0.87	337	0.75	-1.45	8860
0.84	336	0.72	-1.12	8204
0.32	346	0.68	-1.25	8436
0.74	342	0.72	-1.51	8501
0.13	363	1.12	0.18	6927
0.23	359	0.66	1.45	7775
0.19	360	0.63	-2.21	7205
0.84	336	0.72	-1.12	8204
0.32	324	0.66	-0.27	8442
0.32	353	0.65	-1.45	7341
0.48	**346**	**0.73**	**-0.88**	**7990**

RT trans. per year	Mnths in the mrket X/144	Avg. market return when in	Avg. market return when out	Prtflio value
0.83	133	0.67	2.06	2127
0.67	134	0.84	-0.06	2719
0.17	133	1.09	-3.01	4001
0.75	134	0.69	1.88	2226
0.08	139	0.83	-0.81	2957
0.17	137	0.96	-2.78	3450
0.17	138	0.90	-2.04	3195
0.67	134	0.84	-0.06	2719
0.5	126	0.94	-0.36	3034
0.25	134	1.04	-2.70	3712
0.43	**134**	**0.88**	**-0.79**	**3014**

Most importantly, the mean of the return outperforms the buy-and-hold of 2638 for the out-of-sample period (ignoring dividends) at 95% significance. Average spread between in vs. out of market periods was 1.67% per month or 0.0557% per day, compared to 0.001413 per day reported by AK. The mean of the returns of the market when the rules put us in it is greater than the mean of the returns of the market when the rules put us out of it at almost 99% significance. Average number of trades per year was 0.83, compared to 3.8 for AK. The rules put us in the market 93% of the time, compared to 57% for AK.

In the final table we see the results for the coevolved specialized buy and sell rules. Because the variance is low, the mean of the return outperforms the buy-and-hold (ignoring dividends) at 99.5%. These results seem to indicate that there is value in the specialization gained by separating the buy and sell rules and cooperatively coevolving the rules.

IN	10816	14957	22074	18467	19861
OUT	2807	3476	3434	3475	2856
IN	12374	11530	12667	10646	11818
OUT	2823	3258	2911	3541	2837

IN Sample Avg 14521
OUT Sample Avg 3141.8

In summary, we have presented a number of experiments in which GP-evolved technical trading rules outperform a buy-and-hold strategy on the S&P500. The results call into question the Efficient Market Hypothesis, even in its weakest form. In the interest of preserving efficiency in future markets, we include the GP-evolved technical trading rule with the best excess return in the out-of sample period: (12-month RateOfChange < 3-month RateOfChange) OR (1st local maximum of the 3-month Moving Average > 2nd preceding local minimum of the 3-month Moving Average). If you are out of the market and the rule becomes true, buy. If you are in the market and the rule becomes false, sell.

References

1 Allen, F. & Karjalainen, R. 1999. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics* 51:245-271.
2 Neely, C.J. 1999. Risk-Adjusted, Ex Ante, Optimal, Technical Trading Rules in Equity Markets. Working Paper 199-015D. Federal Reserve Bank of St. Louis. To appear in *International Review of Economics and Finance.*
3 Chen, S.-H. 2002. *Genetic Algorithms and Genetic Programming in Computational Finance.* Boston, MA: Kluwer.
4 Neely, C., Weller, P., & Dittmar, R. 1997. Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach. *Journal of Financial and Quantitative Analysis* 32:405-26.
5 Thomas, J. & Sycara, K. 1999. The Importance of Simplicity and Validation in Genetic Programming for Data Mining in Financial Data. *Proceedings of the joint AAAI-1999 and GECCO-1999 Workshop on Data Mining with Evolutionary Algorithms*, July, 1999.
6 Brock, W., Lankonsihok, J. & LeBaron, B. 1992. Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *Journal of Finance* 47:1731-1764.
7 Neftci, S.N. 1991. Naïve Trading Rules in Financial Markets and Wiener-Kolmogorov Prediction Theory: A Study of "Technical Analysis." *Journal of Business* 64:549-571.
8 Lo, A.W., Mamaysky, H. & Wang, J. 2000. Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *The Journal of Finance* 55:1705-1770.
9 Fama, E.F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work .*The Journal of Finance* 25:383-417.
10 Fama, E.F. & M. Blume. 1970. Filter Rules and Stock Market Trading Profits. *Journal of Business* 39:226-241.
11 Bessembinder, H. & Chan, K. 1998. Market Efficiency and the Returns to Technical Analysis. *Financial Management* 27: 5-17.
12 Pring, M.J. 1980. *Technical Analysis Explained.* New York: MCGraw-Hill.
13 Pring, M.J. 1993. *Martin Pring on Market Momentum*. New Delhi: Vision Books.
14 Koza, J. R. 1992. *Genetic Programming: On the Programming of Computers by means of Natural Selection*. Cambridge, MA: MIT Press.
15 http: //lancet.mit.edu/ga/
16 Becker, L.A. & Seshadri, M. 2003. Comprehensibility & Overfitting Avoidance in Genetic Programming for Technical Trading Rules, Worcester Polytechnic Institute, Computer Science Technical Report WPI-CS-TR-03-09.
17 Potter, M.A. & DeJong, K. 2000. Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents. *Evolutionary Computation* 8(1):1-29.
18 Becker, L.A. & Seshadri, M. 2003. Cooperative Coevolution of Technical Trading Rules. Worcester Polytechnic Institute, Computer Science Technical Report WPI-CS-TR-03-15.