Market Microstructure: Can Dinosaurs Return? A Self-Organizing Map Approach under an Evolutionary Framework

Michael Kampouridis[1], Shu-Heng Chen[2], and Edward Tsang[1]

[1] School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, CO4 3SQ, UK
[2] AI-Econ Center, Department of Economics, National Cheng Chi University, Taipei, Taiwan 11623

Abstract. This paper extends a previous model where we examined the markets' microstructure dynamics by using Genetic Programming as a trading rule inference engine, and Self Organizing Maps as a clustering machine for those rules. However, an assumption we made in that model was that clusters, and thus trading strategy types, had to remain the same over time. This assumption could be considered unrealistic, but it was necessary for the purposes of our tests. For this reason, in this paper we extend this model by relaxing this assumption. Hence our framework does not lie on pre-specified types, nor do these types remain the same throughout time. This allows us to investigate the dynamics of market behavior and more specifically whether successful strategies from the past can be successfully applied to the future. In the past, we investigated this phenomenon by using a simple fitness test. Nevertheless, a drawback of that approach was that because of its simplicity, it could only offer limited understanding of the complex dynamics of market behavior. With the extended model we can thus have a more realistic view of the markets and hence draw safer conclusions about their behavior. Empirical results show that market behavior is non-stationary, and thus agents' strategies need to continuously co-evolve with the market, in order to remain effective.

Key words: Genetic Programming, Self-Organizing Maps, Market Microstructure, Market Behavior

1 Introduction

In a previous work [3], we investigated the dynamics of fractions of trading strategy types that exist in financial markets. That study used Genetic Programming (GP) [11] to infer trading rules and Self Organizing Maps (SOM) [10] for clustering these rules into *trading strategy types*. However, an important assumption of that study was that maps (and thus strategy types) among different time periods remained the same. The reason for doing this was because we wanted to allow cross-period comparison among clusters from different SOMs. We could not meaningfully have done this without something like topological equivalence, which could not sustain without the constancy of the types. In this paper, we relax the above assumption, since our current work does not require any cross-period comparisons. Thus one contribution of this paper is that we extend our previous model by using a more *dynamic* approach, where we allow the types of trading strategies to change throughout time. This modification allows us to investigate the market behavior dynamics in more details.

Our motivation behind this investigation is inspired by observations made under artificial market simulations [1, 4], which suggest that the behavior of financial markets is non-stationary. This basically means that this behavior cannot cycle, but instead it follows a linear path. An important implication of this is that trading strategies need to continuously co-evolve with the markets; if they don't, then they become obsolete or *dinosaurs* [1, 9].

Recently [9], we investigated the plausibility of this phenomenon under empirical datasets and found that after applying a strategy from the past to a future time period, this strategy could not perform as successfully as it originally did. This thus verified the previous observations under artificial markets. However, in that work [9], we only used a GP framework, where we inferred trading rules and then observed the fitness of those rules over time. A pitfall of that approach was that because of the simplicity of that methodology, those results offered limited understanding of the complex dynamics of market behavior.

In this paper, we present preliminary results from a more *rigorous* approach, where we apply the extended framework of [3], which as we have said uses both GP and SOM. This brings us to the second contribution of this paper. Using the above techniques offers a more realistic modeling of the market, and allows us to draw more accurate conclusions about the behavior of financial markets.

The rest of this paper is organized as follows: Section 2 presents the methods used for our tests, namely GP and SOM, and also explains our motivation for using them. Section 3 briefly presents the GP algorithm we have used. Section 4 then presents the experimental designs, and Sect. 5 presents and discusses the results of our experiments. Finally, Sect. 6 concludes this paper.

2 Methods

2.1 Genetic Programming as a Rule-Inference Engine

In this paper, we assume that traders' behavior, including price expectations and trading strategies, is either not observable or not available. Instead, their behavioral rules have to be estimated by the observable market price. In order to estimate these rules, we use Genetic Programming (GP).

The use of GP is motivated by considering the market as an evolutionary and selective process.[3] In this process, traders with different behavioral rules participate in the markets. Those behavioral rules which help traders gain lucrative profits will attract more traders to imitate, and rules which result in losses

[3] See [13, 14] for his eloquent presentation of the *Adaptive Market Hypothesis*

will attract fewer traders. An advantage of GP is that it does not rest upon any pre-specified class of behavioral rules, like many other models in the agent-based finance literature [2]. Instead, in GP, a population of behavioral rules is randomly initiated, and the survival-of-the-fittest principle drives the entire population to become fitter and fitter in relation to the environment. In other words, given the non-trivial financial incentive from trading, traders are aggressively searching for the most profitable trading rules. Therefore, the rules that are outperformed will be replaced, and only those very competitive rules will be sustained in this highly competitive search process.[4]

Hence, even though we are not informed of the behavioral rules followed by traders at any specific time horizon, GP can help us infer what these rules are approximately by simulating the evolution of the microstructure of the market. Traders can then be clustered based on realistic, and possibly complex behavioral rules.[5] The GP algorithm used to infer the rules is presented in Sect. 3.

2.2 Self Organizing Maps for Clustering

Once a population of rules is inferred from GP, it is desirable to cluster them based on a chosen similarity criterion so as to provide a concise representation of the microstructure. The similarity criterion which we choose is based on the *observed trading behavior*. Based on this criterion, two rules are similar if they are *observationally equivalent* or *similar*, or, alternatively put, they are similar if they generate the same or similar market timing behavior.

Given the criterion above, the behavior of each trading rule can be represented by its series of market timing decisions over the entire trading horizon, for example, 6 months. Therefore, if we denote the decision "buy" by "1" and "not-to-buy" by "0", then the behavior of each rule is a binary vector. The dimensionality of these vectors is then determined by the length of the trading horizon. For example, if the trading horizon is 125 days long, then the dimension of the market timing vector is 125. Once each trading rule is concretized into its market timing vector, we can then easily cluster these rules by applying Kohonen's Self-Organizing Maps to the associated clusters.

The main advantage of SOMs over other clustering techniques such as K-means is that the former can present the result in a visualizable manner so that we can not only identify these types of traders, but also locate their 2-dimensional position on a map, i.e., a distribution of traders over a map. dynamics of the microstructure directly as if we were watching the population density on a map over time.

What we have discussed so far is presented in Fig. 1, where a 3×3 SOM is presented. Here, 500 artificial traders are grouped into nine clusters. The

[4] It does not mean that all types of traders surviving must be smart and sophisticated. They can be dumb, naive, randomly behaved or zero-intelligent. Obviously, the notion of rationality or bounded rationality applying here is *ecological* [15, 6].

[5] [5] provides the first illustration of using genetic programming to infer the behavioral rules of human agents in the context of ultimatum game experiments. Similarly, [7] uses genetic algorithms to infer behavioral rules of agents from market data.

parameter value '500' refers to the population size used in GP, i.e., the rule-inference stage, whereas the parameter value '9' is due to a 3×3 two-dimensional SOM employed in the rule clustering stage. In a sense, this could be perceived as a snapshot of a nine-type agent-based financial market dynamics. Traders of the same type indicate that their market timing behavior is very similar. The market fraction or the size of each cluster can be seen from the number of traders belonging to that cluster. Thus, we can observe that the largest cluster has a market share of 71.2% (356/500), whereas the smallest one has a market share of 0.2% (1/500).

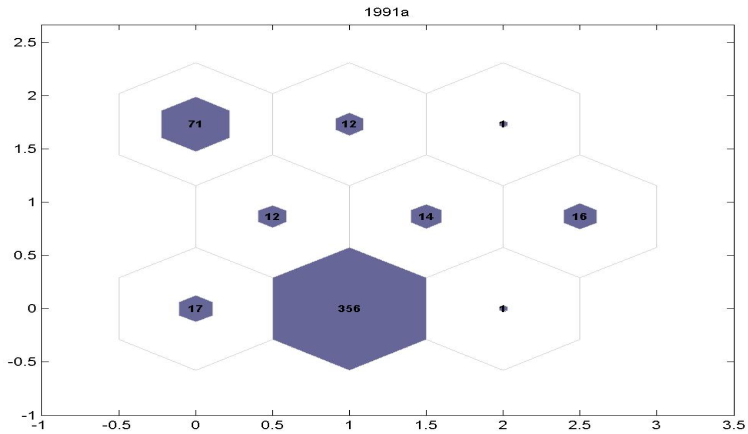

Fig. 1. Example of a 3×3 Self-Organizing Map

3 GP Algorithm

Our GP is inspired by a financial forecasting tool, EDDIE [12, 8], which applies genetic programming to evolve a population of financial advisors, or, alternatively, a population of market-timing strategies, which guide investors on when to buy or hold. These market timing strategies are formulated as decision trees, which, when combined with the use of GP, are referred to as *Genetic Decision Trees* (GDTs). Our GP uses indicators commonly used in technical analysis: Moving Average (MA), Trader Break Out (TBR), Filter (FLR), Volatility (Vol), Momentum (Mom), and Momentum Moving Average (MomMA).[6]. Each indicator has two different periods, a short- and a long-term one, 12 and 50 days respectively.

[6] We use these indicators because they have been proved to be quite useful in previous works like [8]

Each of these market-timing strategies (GDTs) is syntactically (grammatically) produced by the Backus Naur Form (BNF). Figure 2 presents the BNF grammar of the GP. As we can see, the root of the tree is an If-Then-Else statement. Then the first branch is a Boolean (testing whether a technical indicator is greater than/less than/equal to a value). The 'Then' and 'Else' branches can be a new GDT, or a decision, to buy or not-to-buy (denoted by 1 and 0). Thus, each individual in the population is a GDT and its recommendation is to buy (1) or not-buy (0). Depending on the classification of the predictions we can have

<Tree> ::= If-then-else <Condition> <Tree> <Tree> | Decision
<Condition> ::= <Condition>"And"<Condition>|<Condition>"Or"<Condition>|
"Not" <Condition> | Variable <RelationOperation> Threshold
<Variable> ::= MA_12 | MA_50 | TBR_12 | TBR_50 | FLR_12 |FLR_50 | Vol_12 | Vol_50 | Mom_12 | Mom_50 |MomMA_12 | MomMA_50
<RelationOperation> ::= ">" | "<" | "="
Decision is an integer, Positive or Negative implemented
Threshold is a real number

Fig. 2. The Backus Naur Form of the simple GP uses to construct trees

are four cases: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). As a result, we can use the following 3 metrics:

Rate of Correctness

$$RC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Rate of Missing Chances

$$RMC = \frac{FN}{FN + TP} \tag{2}$$

Rate of Failure

$$RF = \frac{FP}{FP + TP} \tag{3}$$

The above metrics combined give the following fitness function:

$$ff = w_1 * RC - w_2 * RMC - w_3 * RF \tag{4}$$

where w_1, w_2 and w_3 are the weights for RC, RMC and RF respectively, and are given in order to reflect the preferences of investors. For instance, a conservative investor would want to avoid failure; thus a higher weight for RF should be used. For our experiments, we chose to include GDTs that mainly focus on correctness and reduced failure. Thus these weights have been set to 1, $\frac{1}{6}$ and $\frac{1}{2}$ respectively, and are given in this way in order to reflect the importance of each performance measure for our predictions.

Given a set of historical data and the fitness function, GP is then applied to evolve these market-timing strategies in a standard way. After evolving a number of generations, what stands (survives) at the end (the last generation) is, presumably, a population of financial agents whose market-timing strategies are financially rather successful.

4 Experimental Designs

This section presents the experimental designs. But before we do this, let us first present some frequently used terms:

- *Base period*, is the period during which GP creates and evolves GDTs
- *Future period(s)*, is a period(s) which follow(s) the base period (in chronological order)

The experiments were conducted for a period of 17 years (1991-2007) and the data was taken from the daily closing prices of the STI market index (Singapore). For statistical purposes, we run the experiments for 10 times.

Each year was split into 2 halves (January-June, July-December), so in total, out of the 17 years, we have 34 periods[7]. The first semester of a year will be denoted with an 'a' at the end (e.g. 1991a), and the second semester of a year will be denoted with a 'b' at the end (e.g. 1991b).The GP system was therefore executed 34 times. Table 1 presents the GP parameters for our experiments. As we have already mentioned, after generating and evolving strategies for each

Table 1. GP Parameters

GP Parameters	
Max Initial Depth	6
Max Depth	17
Generations	50
Population size	500
Tournament size	2
Reproduction probability	0.1
Crossover probability	0.9
Mutation probability	0.01

one of the 34 periods, we then use SOM to cluster them into strategy types. Then, in order to investigate whether the behavior of markets is stationary, we re-cluster the GDTs of each base period, to all future periods' clusters. We want to investigate is how *"dissatisfied"* these GDTs will be when they are moved to future periods.

[7] At this point the length of the period was chosen arbitrarily to 6 months. We leave it to a future research to examine if and how this time horizon can affect our results.

The process of this re-clustering is quite simple; let us give an example when 1991a is the base period. Each evolved GDT would first be moved to the next period, 1991b, and be re-clustered to one of the clusters of that period. In order to 'decide' which cluster to choose, the GDT compares the euclidean distance of its market timing vector with each cluster; it is then placed to the cluster with the smallest euclidean distance. The same procedure follows for all GDTs of the population. At the end, the population of evolved GDTs from the base period of 1991a has been reclustered to the clusters of period 1991b. We also follow the same procedure with all future periods. This means that the GDTs from 1991a are also re-clustered in 1992a, 1992b, ..., 2007b. Finally, the same process is done for all other base periods (i.e. 1991b, 1992a, ..., 2007a).

Once these processes are complete, we can then move to calculate the dissatisfaction rate per period. In order to calculate it, we again use the euclidean distance as our metric. Dissatisfaction rate of a GDT is therefore defined as the euclidean distance of this GDT's behavior vector from the centroid of the cluster it was placed during the re-clustering procedure. By repeating this process for each GDT per period, we can then calculate how dissatisfied all GDTs are as a whole, and thus obtain their average dissatisfaction rate, which will act as a measure of how dissatisfied the population is every period. From now on, this average dissatisfaction rate of the population of GDTs is going to be called as *dissatisfaction rate*.

The logic of using the dissatisfaction rate is the following: we want to investigate whether this dissatisfaction increases, because the GDTs have not adapted to any changes that have happened in the market. In other words, we are interested in observing if unadapted strategies can still fit in the market environment (clusters) as well as they did in the past. If they do not, this means that their new environment does not represent them as effectively as it did in the past; this thus causes these strategies to be 'dissatisfied'. If this dissatisfaction is quite high, we consider these strategies to be *dinosaurs*, because their rules are not up-to-date with the new market environment and as a result of this, they have become obsolete. The implications of this is that markets' behavior is non-stationary, but constantly changes, as Arthur first observed in [1]. Thus, strategies need to constantly co-evolve with the market and the changes happening in it, in order to remain effective and survive.

Given a base period, the population dissatisfaction of all periods is normalized by dividing those population dissatisfaction rates by the population dissatisfaction rate in the base period. Hence, each base period has its normalized dissatisfaction rate equal to 1 and a *returning dinosaur* is a population of strategies from future periods that has its normalized dissatisfaction rate less than or equal to 1.

5 Results

In order to examine how often dinosaurs return, we iterate through each base period and calculate the minimum dissatisfaction rate among its future periods.

If, for instance, 1991a is the base period, then there is a series of 33 population dissatisfaction values for its future periods. We obtain the minimum value among these 33 values, in order to check how close to 1 this future period is. This process is then repeated for 1991b and its 32 future periods, 1992a, and so on, until base period 2007a. We thus end up with a 1×33 vector, called *DissatisfVect* which shows the potential returning dinosaur per base period. In addition, we repeat the above procedure under the following SOM dimensions: 2×1, 3×1, 2×2, 5×1, 3×2, 7×1, 4×2, and 3×3. The graphs of the *DissatisfVect* vectors are presented in Fig. 3. Each line represents the results on a different SOM dimension. What we can see from these graphs is that there are no base

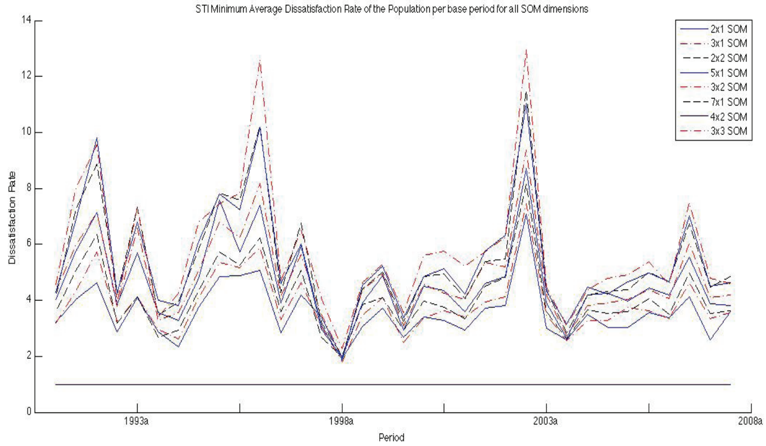

Fig. 3. Minimum normalized population dissatisfaction rate among all future periods for each base period. Each line represents a different SOM dimension.

periods with a minimum normalized dissatisfaction rate below 1. In fact, the closest to 1 this rate gets is around 2 (1998a). Table 2 presents the average of the minimum dissatisfaction rate per cluster and verifies this observation. As we can see, the minimum dissatisfaction rate is on average 3.56 for the 2×1 SOM and it gradually increases as the number of clusters increases, reaching 5.79 for the 3×3 SOM. Hence, the minimum dissatisfaction rate is on average quite far from 1, which as we said is the threshold for a returning dinosaur.

In addition, Table 3 informs us that the average dissatisfaction rate per cluster is even higher, and ranges from 5.17 (2 clusters) to 8.88 (9 clusters). It is thus obvious that on average no dinosaurs return. But even if we want to take into account the outliers (minimum dissatisfaction rate-Fig. 3), we can see that while the rate can get relatively low, it never reaches 1. This leads us to argue that *dinosaurs do not return or return only as lizards*. More specifically, the strategies (GDTs) found the new environments (clusters) very different from the ones in their base period and were very 'dissatisfied'. The strategies that had not adapted to the market changes could not fit in the new environment.

Furthermore, Table 3 also informs us that the standard deviation of the dissatisfaction rate is around 19-20%. This indicates that there are big upwards and downwards movements of the dissatisfaction rate. This is also verified by Fig. 3. This swinging reminds us of the Market Fraction Hypothesis (MFH) [3], which requires a *constant swinging (change) of the fractions of strategy types that exist in a market*. This is a very important observation, because it gives us an understanding of why the dissatisfaction rate swings: similar clusters (i.e. strategy types), to the ones of the base period, appear in the market. However, as these clusters are not exactly the same as in the past, the dissatisfaction rate cannot reach its minimum level and thus dinosaurs can only return as lizards, as we saw in Fig. 3. The above observation leads us to conclude that

Table 2. Average Minimum Dissatisfaction Rate per Cluster

	2×1	3×1	2×2	5×1	3×2	7×1	4×2	3×3
Mean	3.56	3.83	4.09	4.61	4.79	5.34	5.41	5.79

Table 3. Summary Statistics for Dissatisfaction Rate for STI per cluster

	2×1	3×1	2×2	5×1	3×2	7×1	4×2	3×3
Mean	5.17	5.65	6.11	6.98	7.19	8.30	8.33	8.88
Stand. Deviation (%)	18.71	19.67	19.70	19.44	19.73	20.37	20.13	19.88

market behavior constantly changes. However, it can sometimes resemble old environments. When this happens, old strategies might perform relatively well again (i.e. dinosaurs returning as lizards). Nevertheless, unadapted old strategies cannot reach performance levels equal to the ones they once had in their base period (i.e. no returning dinosaurs). Market conditions have changed and unless these strategies follow these changes, they become dinosaurs and thus ineffective.

One final observation we can make is that the dissatisfaction rate increases as the number of clusters increases. This should not surprise us, since the increased number of clusters has increased the sum of the dissatisfaction rates among the 500 GDTs, and thus has increased the average rate, too. What is important to state, however, is that no returning dinosaurs are observed, under all trading strategy types. The number of clusters does not affect the test's results.

6 Conclusion

To conclude, this paper presented preliminary results under a model which used Genetic Programming to infer trading strategies and Self Organizing Maps to cluster these strategies. This allowed us to create a realistic model of market behavior, where we could investigate an important property of artificial agent-based financial markets. This property basically supports that the nature and constituents of agents, and thus strategies, constantly changes; if these strategies do not continuously adapt to the changes in their environments, then they become obsolete (dinosaurs). In a previous work [9], we investigated this property under empirical data. However, because of the simplicity of the test used in that work, we also tested it in this paper, by using our more realistic model, which combines GP and SOM. The model in this paper also served as an extension of another framework presented in [3]. Results showed that on average, the dataset tested in this paper, STI, did not demonstrate the existence of returning dinosaurs. The implications of this are very important. Old strategies cannot successfully be re-applied to future periods, unless they have co-evolved with the market. If they have not, they become obsolete, because the market conditions have changed. They can occasionally return as lizards, meaning that they can show some relatively good performance, but they cannot become again as successful, as they initially were.

Finally, it should be stated that the above results are only under a single dataset (STI). The next step of our research is to explore other financial markets and see if these results are a universal phenomenon.

Acknowledgments

This work was supported by an EPSRC grant (No. EP/P563361/01) and by AI-Econ Center of National Cheng Chi University, Taiwan.

References

1. Arthur, B.: On learning and adaptation in the economy (1992), working paper 92-07-038, Santa Fe Institute
2. Chen, S.H., Chang, C.L., Du, Y.R.: Agent-based economic models and econometrics. Journal of Knowledge Engineering Review (to appear) (2010)
3. Chen, S.H., Kampouridis, M., Tsang, E.: Microstructure dynamics and agent-based financial markets. In: Multi-Agent-Based Simulation XI, 11th International Workshop, Revised Papers, LNAI. Springer-Verlag (2011), forthcoming
4. Chen, S.H., Yeh, C.H.: Evolving traders and the business school with genetic programming: A new architecture of the agent based artificial stock market. Journal Of Economic Dynamics & Control 25, 363–393 (2001)
5. Duffy, J., Engle-Warnick, J.: Using symbolic regression to infer strategies from experimental data, pp. 61–82. Springer (2002)
6. Gigerenzer, G., Todd, P.: Fast and Frugal Heuristics: The Adaptive Toolbox, pp. 3–34. Gigerenzer, G. Todd, P. and the ABC Research Group, Oxford University Press (1999)

7. Izumi, K., Ueda, K.: Analysis of dealers' processing financial news based on an artificial market approach. Journal of Computational Intelligence in Finance 7, 23–33 (1999)
8. Kampouridis, M., Tsang, E.: EDDIE for investment opportunities forecasting: Extending the search space of the GP. In: Proceedings of the IEEE Conference on Evolutionary Computation. pp. 2019–2026. Barcelona, Spain (2010)
9. Kampouridis, M., Chen, S.H., Tsang, E.: Testing the dinosaur hypothesis under empirical datasets. In: Schaefer, R., Cotta, C., Kołodziej, J., Rudolph, G. (eds.) Parallel Problem Solving from Nature – PPSN XI. Lecture Notes in Computer Science, vol. 6239, pp. 199–208. Springer (2010)
10. Kohonen, T.: Self-organized formation of topologically correct feature maps. Journal of Biological Cybernetics 43, 59–69 (1982)
11. Koza, J.: Genetic Programming: on the programming of computers by means of natural selection. Cambridge, MA: MIT Press (1992)
12. Li, J., Tsang, E.: Reducing failures in investment recommendations using genetic programming. In: The 6th International Conference on Computing in Economics and Finance, Society for Computational Economics. Barcelona (2000)
13. Lo, A.: The adaptive market hypothesis: market efficiency from an evolutionary perspective. Journal of Portfolio Management 30, 15–29 (2004)
14. Lo, A.: Reconciling efficient markets with behavioral finance: The adaptive markets hypothesis. Journal of Investment Consulting 2, 21–44 (2005)
15. Simon, H.: Rational choice and the structure of environments. Psychological Review 63, 129–138 (1956)