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 market microstructure model, which investigated fraction dynamics of trading strategies. Our model consisted of two parts: Genetic Programming, which acted as an inference engine for trading rules, and Self-Organizing Maps (SOM), which was used for clustering the above rules into trading strategy types. However, for the purposes of the experiments of our previous work, we needed to make the assumption that SOM maps, and thus strategy types, remained the same over time. Nevertheless, this assumption could be considered as strict, and even unrealistic. In this paper, we relax this assumption. This offers a significant extension to our model, because it makes it more realistic. In addition, this extension allows us to investigate the dynamics of market behavior. We are interested in examining whether financial markets' behavior is non-stationary, because this implies that strategies from the past cannot be applied to future time periods, unless they have co-evolved with the market. The results on an empirical financial market show that its behavior constantly changes; thus, agents' strategies need to continuously adapt to the changes taking place in the market, in order to remain effective.

Keywords: Genetic Programming, Self-Organizing Maps, Market Microstructure, Market Behavior.

1 Introduction

There are several types of models in the agent-based financial markets literature. One way of categorizing them is to divide them into the N-type models and the Santa-Fe Institute (SFI) like ones [2]. The former type of models focuses on the mesoscopic level of markets, by allowing agents to choose among different types of strategies. A typical example is the fundamentalist-chartist model. Agents in this model are presented with these two strategy types and at any given time they have to choose between these two. A typical area of investigation of these models is fraction dynamics, i.e., how the fractions of the different strategy types change over time. However, what is not presented in most of these models is

C. Di Chio et al. (Eds.): EvoApplications 2011, Part II, LNCS 6625, pp. 91–100, 2011.
© Springer-Verlag Berlin Heidelberg 2011

novelty-discovering agents. For instance, in the fundamentalist-chartists example, agents can only choose between these two types; they cannot create new strategies that do not fall into either of these types. On the other hand, the SFI-like models overcome this problem by focusing on the microscopic level of the markets. By using tools such as Genetic Programming [7], these models allow the creation and evolution of novel agents, which are not constrained by pre-specified strategy types.[1] However, this kind of models tends to focus on price dynamics, rather than fraction dynamics [2].

In a previous work [3], we combined properties from the N-type and SFI-like models into a novel model. We first used Genetic Programming (GP) as a rule inference engine, which created and evolved autonomous agents; we then used Self-Organizing Maps (SOM) [6] as a clustering machine, and thus re-created the mesoscopic level that the N-type models represent, where agents were categorized into different strategy types. We then investigated the short- and long-term dynamics of the fractions of strategies that existed in a financial market. Nevertheless, that study rested upon an important assumption, i.e., the maps derived from each time period were comparable with each other. This comparability assumption itself required that the types (clusters), as well as their operational specification, would not change over time. If this were not the case, the subsequent study would be questioned. This was mainly due to one technical step in our analysis called translation. The purpose of translation was to place the behavior of agents observed in one period into a different period and to recluster it for the further cross-period comparison. We could not meaningfully have done this without something like topological equivalence, which could not be sustained without the constancy of the types.

However, this assumption can be considered as strict and unrealistic. Strategy types do not necessarily remain the same over time. For instance, if a chartist strategy type exists in time t, it is not certain it will also exist in $t+1$. If market conditions change dramatically, the agents might consider other strategy types as more effective and choose them. The chartist strategy would then stop existing.

In this paper, we relax the above assumption, since our current work does not require cross-period comparisons. Our model thus becomes more realistic. In addition, *we shift our focus from fraction dynamics to behavior dynamics*: we examine the plausibility of an observation made under artificial markets [1], which suggests that the nature of financial markets constantly changes. This implies that trading strategies need to constantly co-evolve with the markets; if they do not, they become obsolete or *dinosaurs* [1]. We hence test if this observation holds in the 'real' world, under an empirical financial market. This will offer important insights regarding the behavior dynamics of the markets.

The rest of this paper is organized as follows: Section 2 presents our model, and Sect. 3 briefly presents the GP algorithm we use. Section 4 then presents the experimental designs, Sect. 5 reviews the testing methodology, and Sect. 6 presents the results of our experiments. Finally, Sect. 7 concludes this paper.

[1] We refer the reader to [2], which provides a thorough review on both N-type and SFI-like models, along with a detailed list of them.

2 Model

2.1 Genetic Programming as a Rule-Inference Engine

The use of GP is motivated by considering the market as an evolutionary and selective process.[2] 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 will attract fewer traders. An advantage of GP is that it does not rest upon any prespecified 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.

Hence, GP can help us infer what are the rules the traders follow, by simulating the evolution of the microstructure of the market. Traders can then be clustered based on realistic, and possibly complex behavioral rules.The GP algorithm used to infer the rules is presented in detail, later, 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*.[3] 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.[4]

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 "enter the market" by "1" and "leave the market" 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.

[2] See [8] for his eloquent presentation of the *Adaptive Market Hypothesis*.

[3] Other similarity criteria could take place, too, such as risk averseness. However, in this paper we wanted to focus on the behavioral aspects of the rules.

[4] One might question the above similarity criterion, since very different rules might be able to produce the same signals. This does not pose a problem in this work, since we are interested in the behavior of the market (and thus the rules' behavior). We are not interested in the semantics aspect of the rules.

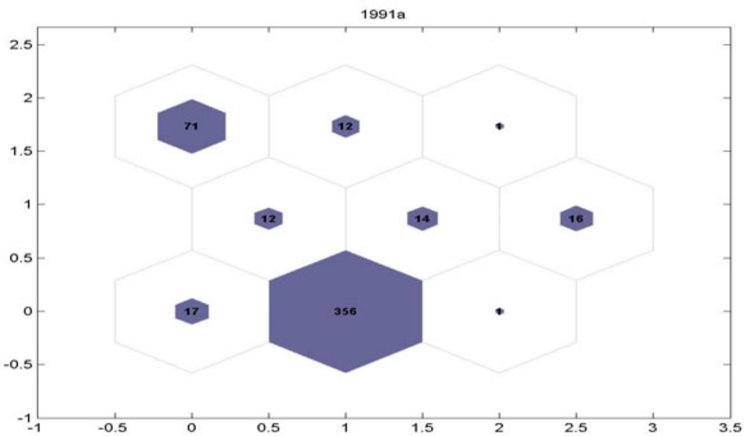

Fig. 1. Example of a 3 × 3 Self-Organizing Map

Figure 1 presents a 3×3 SOM. Here, 500 artificial traders are grouped into nine clusters. 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).

3 GP Algorithm

Our GP is inspired by a financial forecasting tool, EDDIE [4], which applies genetic programming to evolve 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).[5] Each indicator has two different periods, a short- and a long-term one (12 and 50 days). Figure 2 presents a sample GDT generated by the GP.

Depending on the classification of the predictions, there are four cases: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). We then use the following 3 metrics, presented in Equations (1)-(3):

[5] We use these indicators because they have been proved to be quite useful in previous works like [4]. However, the purpose of this work is not to provide a list of the ultimate technical indicators.

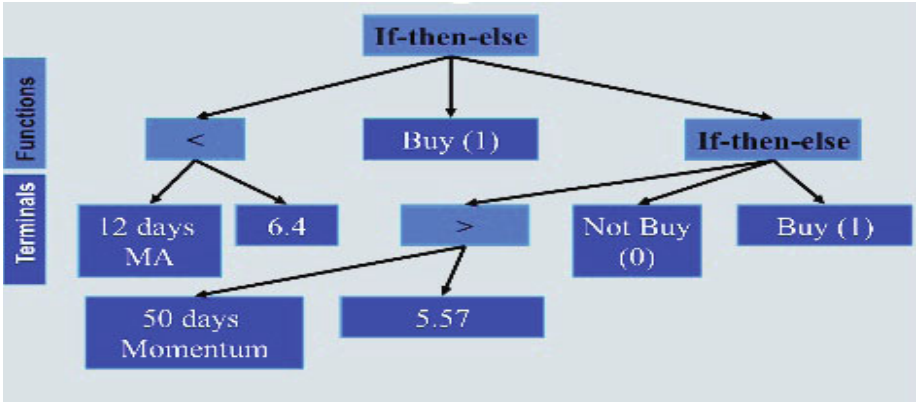

Fig. 2. Sample GDT generated by the GP

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.

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

4 Experimental Designs

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

Each year is split into 2 halves (January-June, July-December), so in total, out of the 17 years, we have 34 periods.[6] The first semester of a year is denoted with an ‘a’ at the end (e.g., 1991a), and the second semester of a year is denoted with a ‘b’ (e.g., 1991b). The GP systems is therefore executed 34 times, i.e., one time per period. Table 1 presents the GP parameters for our experiments. The GP parameters for our experiments are the ones used by Koza [7]. Only the tournament size has been lowered, because we were observing premature convergence. Other than that, the results seem to be insensitive to these parameters.

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

After generating and evolving strategies for each one of the 34 periods, we then use SOM to cluster these strategies into types. We do this for every one of the 34 periods. Thus, we end up with 34 different SOMs, one per semester, which represent the market in different time periods over the 17-year horizon.

Finally, we define as ‘base period’, the period during which GP creates and evolves GDTs. We also define ‘future period(s)’, as the period(s) which follow(s) the base period (in chronological order).

5 Testing Methodology

In order to investigate whether the behavior of markets is non-stationary, we recluster the GDTs of each base period, to all future periods’ clusters.[7] By applying the same GDTs (strategies) to clusters of future periods, we can observe how well these strategies fit in the new environments (clusters). The logic behind this is the following: when we first evolved and clustered the GDTs (base period), these GDTs were placed in clusters that represented their respective strategies. For instance, if there was a strategy type (cluster) that represented ‘chartists’, then all GDTs which followed a chartist strategy were placed in that cluster. When we then take the GDTs from a base period and recluster them to strategy

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

[7] The process of reclustering is explained later in this section.

types of future periods, it is not guaranteed that there will again be a cluster that represents chartists. If the market constantly changes, there is a possibility that this type of strategies does not exist any more in the future periods. Thus, the GDTs find themselves *unadapted* to the new environment (clusters) and have to choose another cluster, which represents them as closely as possible. This cluster will be the one that has the centroid with the smallest Euclidean distance[8] from the market-timing vectors of these GDTs. Of course, since now the SOM of the future period is formed by different clusters, the GDTs might not fit in as well as they did in the base period. In order to measure this 'unfitting', we use a 'dissatisfaction rate', i.e., how dissatisfied these GDTs will be when placed into a future period's cluster that does not represent their strategy. *If the market is non-stationary, the GDTs' dissatisfaction rate will be high*, as a result of the changes that took place in the market. The dissatisfaction rate is defined as the Euclidean distance of a GDT's market-timing vector to the centroid of the cluster in which it is placed, after the reclustering procedure. Under a non-stationary market behavior, the following statement should hold:

The average dissatisfaction rate of the population of GDTs from future periods should not return to the range of dissatisfaction of the base period.

Hence, we will test the above statement against the STI index.

Let us now explain the process of reclustering. We start with 1991a as the base period. Each evolved GDT is moved to the next period, 1991b, and reclustered into 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 to the centroid of each cluster; it is then placed into 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 will have been reclustered into the clusters of period 1991b. The same procedure is followed in all future periods. This means that the GDTs from 1991a are also reclustered into 1992a, 1992b, ..., 2007b. Finally, the same process is done for all other base periods (i.e., 1991b, 1992a, ..., 2007a).

Once the process of reclustering is complete, we calculate the dissatisfaction rate of each GDT in the population. Next, we calculate the population's average dissatisfaction rate. We do the same for all 34 periods. Given a base period, the population average dissatisfaction of all periods is normalized by dividing those population average dissatisfaction rates by the population average dissatisfaction rate in the base period. Hence, each base period has its normalized average dissatisfaction rate equal to 1. In order to prove that the market is non-stationary, we need to show that the normalized average dissatisfaction rate of the GDTs increases in the future periods, and never returns to its initial value of 1, which was during the base period. If, on the other hand, this rate reaches 1 or below, *it is an indication of a cyclic market behavior*, since the GDTs have found the same conditions with the base period, and as a result feel as 'satisfied' as before.

[8] One may wonder if the choice of the Euclidean distance as a distance metric, when the vectors of the GDTs are binary, is an appropriate one. However, this does not pose a problem, because the vectors of the clusters' centroids are real valued.

Finally, we define as *dinosaurs* the population of GDTs that has been reclustered from a base period to future periods. The reason of calling them in this way is because these GDTs have not adapted to the new market environment (clusters of the SOMs from future periods) and are thus ineffective. If these GDTs' normalized average dissatisfaction rate drops to *less than or equal to 1*, we call them *returning dinosaurs*, because they have become effective again.[9]

6 Results

As explained, returning dinosaurs denote a cyclic market behavior. To examine if dinosaurs return, we iterate through each base period and calculate the minimum normalized average dissatisfaction rate for each future period. This gives us an indication of how many returning dinosaurs, if any, exist. 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, and so on, until base period 2007a. We thus end up with a 1×33 vector, which presents the minimum dissatisfaction per base period and thus shows whether any returning dinosaurs exist. In addition, we are interested in investigating whether different number of clusters (strategy types) can affect the test results. We thus run tests under 2 to 9 clusters, for 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 minimum dissatisfaction vectors for the STI index are presented in Fig. 3. Each line represents the results of a different SOM dimension. The horizontal line indicates a dissatisfaction of 1, and is given as a reference.

What we can see from Fig. 3 is that there are no base periods with a minimum normalized dissatisfaction rate below 1. In fact, the closest to 1 this rate gets is around 2 (1998a). The first row of 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 away from 1, which as we mentioned is the threshold for a returning dinosaur.

In addition, the second row of Table 2 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 and Table 2), 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)

[9] In a previous work [5], where we investigated the markets' behavior dynamics by only using GP but not SOM, we did not use this 'strict' definition of returning dinosaurs. This led us to conclude that returning dinosaurs existed. However, if we had also used the current paper's definition, the results from [5] would not have dramatically differed from the current paper.

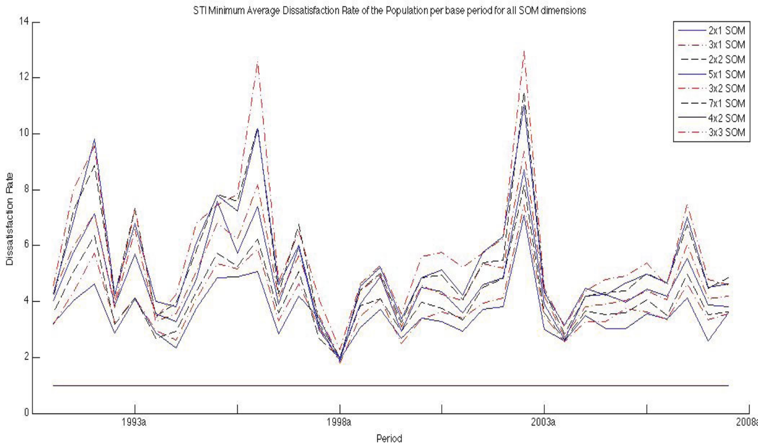

Fig. 3. Minimum normalized population dissatisfaction rate among all future periods for each base period for the STI index. Each line represents a different SOM dimension.

Table 2. Average Minimum Dissatisfaction (A.M.D.-row 1) and Average Dissatisfaction (A.D.-row 2) Rate per Cluster

	2×1	3×1	2×2	5×1	3×2	7×1	4×2	3×3
Mean of A.M.D.	3.56	3.83	4.09	4.61	4.79	5.34	5.41	5.79
Mean of A.D.	5.17	5.65	6.11	6.98	7.19	8.30	8.33	8.88

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. The above observation allows us to conclude that STI's behavior constantly changes. However, this behavior can sometimes resemble older ones. When this happens, old strategies might perform relatively well again (i.e., dinosaurs return as lizards). Nevertheless, strategies that have not co-evolved with the market, cannot reach performance levels as the ones they once had in their base period (i.e., no returning dinosaurs). Market conditions have changed and unless these strategies follow the changes, they become dinosaurs and thus ineffective.

One final observation we can make is that the number of clusters does not affect the test's results. The dissatisfaction rate of each market follows always the same pattern, regardless the number of clusters. No returning dinosaurs are observed, under any number of the trading strategy types tested.

7 Conclusion

To conclude, this paper presented a significant extension to a previous market microstructure model [3], and also discussed preliminary results on the behavior dynamics of financial markets. Our experimental work was inspired by an an observation made under artificial agent-based financial markets [1]. This observation says that the nature and constituents of agents, and thus their strategies, constantly change; if these strategies do not continuously adapt to the changes in their environments, then they become obsolete (dinosaurs). The results showed that on average, the dataset tested in this paper, STI (Singapore), did not demonstrate the existence of returning dinosaurs, and thus *verified the existence of the non-stationary property in financial markets' behavior.* The implications of this are very important. Strategies from the past cannot be successfully re-applied to future periods, unless they have co-evolved with the market. If they have not, they become obsolete, because the market conditions change continuously. They can occasionally return as lizards, meaning that these strategies can sometimes demonstrate relatively good performance, but they cannot become again as successful, as they once were. The next step of our research is to explore other markets and see if the above 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, working paper 92-07-038, Santa Fe Institute (1992)
2. Chen, S.H., Chang, C.L., Du, Y.R.: Agent-based economic models and econometrics. Journal of Knowledge Engineering Review (2010) (forthcoming)
3. Chen, S.H., Kampouridis, M., Tsang, E.: Microstructure dynamics and agent-based financial markets. In: Bosse, T., Geller, A., Jonker, C.M. (eds.) MABS XI 2010. LNCS (LNAI), vol. 6532, pp. 121–135. Springer, Heidelberg (2011)
4. 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, Barcelona, Spain, pp. 2019–2026 (2010)
5. Kampouridis, M., Chen, S.H., Tsang, E.: Testing the dinosaur hypothesis under empirical datasets. In: Schaefer, R., Cotta, C., Kołodziej, J., Rudolph, G. (eds.) PPSN XI. LNCS, vol. 6239, pp. 199–208. Springer, Heidelberg (2010)
6. Kohonen, T.: Self-organized formation of topologically correct feature maps. Journal of Biological Cybernetics 43, 59–69 (1982)
7. Koza, J.: Genetic Programming: On the programming of computers by means of natural selection. MIT Press, Cambridge (1992)
8. Lo, A.: The adaptive market hypothesis: market efficiency from an evolutionary perspective. Journal of Portfolio Management 30, 15–29 (2004)