International Review of Financial Analysis 23 (2012) 41–54

Contents lists available at ScienceDirect

International Review of Financial Analysis

Market fraction hypothesis: A proposed test

Michael Kampouridis [a,*], Shu-Heng Chen [b], Edward Tsang [c]

[a] School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, CO4 3SQ, UK
[b] AI-ECON Research Center, Department of Economics, National Chengchi University, 11623, Taiwan, ROC
[c] Centre for Computational Finance and Economic Agents, University of Essex, Wivenhoe Park, CO4 3SQ, UK

ARTICLE INFO

Available online 21 July 2011

Keywords:
Market Fraction Hypothesis
Genetic Programming
Self-Organizing Feature Map
Time-Invariant Self-Organizing Feature Map
Agent-based financial model

ABSTRACT

This paper presents and formalizes the Market Fraction Hypothesis (MFH), and also tests it under empirical datasets. The MFH states that the fraction of the different types of trading strategies that exist in a financial market changes (swings) over time. However, while such swinging has been observed in several agent-based financial models, a common assumption of these models is that the trading strategy types are static and pre-specified. In addition, although the above swinging observation has been made in the past, it has never been formalized into a concrete hypothesis. In this paper, we formalize the MFH by presenting its main constituents. Formalizing the MFH is very important, since it has not happened before and because it allows us to formulate tests that examine the plausibility of this hypothesis. Testing the hypothesis is also important, because it can give us valuable information about the dynamics of the market's microstructure. Our testing methodology follows a novel approach, where the trading strategies are neither static, nor pre-specified, as in the case in the traditional agent-based financial model literature. In order to do this, we use a new agent-based financial model which employs genetic programming as a rule-inference engine, and self-organizing maps as a clustering machine. We then run tests under 10 international markets and find that some parts of the hypothesis are not well-supported by the data. In fact, we find that while the swinging feature can be observed, it only happens among a few strategy types. Thus, even if many strategy types exist in a market, only a few of them can attract a high number of traders for long periods of time.

1. Introduction

In the agent-based financial literature, the proportion of different types of trading strategies in a market can be referred to as the *Market Fraction*. A common observation in many agent-based financial models is that the market fraction of the trading strategy types that exist in a market keeps swinging (i.e., changing). In other words, if for instance we have two types of agents in the market (e.g., fundamentalists and chartists), it has been found that the fraction of these two types of strategies keeps changing over time. If, for example at time *t* 90% of the market participants adopt the fundamental strategy and 10% of them adopt the chartist strategy, these fractions change continuously over time; therefore, in a future time period, we could observe that the fundamentalists occupy only 10% of the agents, and the chartists the other 90%. This swinging feature has been observed in many agent-based financial models (Amilon, 2008; Boswijk, Hommes, & Manzan, 2007; Brock & Hommes, 1998; Gilli & Winker, 2003; Kirman, 1991, 1993; Lux, 1995, 1997, 1998; Winker & Gilli, 2001).

Based on these observations about the swinging of the market fraction, Chen (2008) and Chen, Chang, and Du (2012) suggested a new hypothesis, called the Market Fraction Hypothesis (MFH). The MFH states that there is a constant swinging among the fractions of the types of trading strategies that exist in a market. However, although the term 'Market Fraction Hypothesis' was introduced and used by Chen, it has never been formalized as a hypothesis. This thus motivates us to formalize the MFH, by presenting its main constituents. Formalizing the hypothesis is very important, because it allows us to suggest and formulate tests that will examine its plausibility.

Furthermore, as we mentioned, the swinging feature that the MFH describes has been observed in several agent-based models. However, all of the above models assume that the trading strategy types are static and pre-specified. By this we mean that these models endow their agents with a specific number of trading strategy types which they have to choose from. To the best of our knowledge, the MFH has not been empirically examined under a more dynamic environment in which strategies are not static and are not exogenously given. Therefore, in this paper we will present a new agent-based financial model which incorporates this more general setting and test it.

In addition, motivated by the fact that the observations about the swinging of the market fraction have so far only taken place under artificial frameworks (Chen et al., 2012), we test the MFH under *empirical* datasets. We run tests for 10 international markets and hence provide a general examination of the plausibility of the MFH. One goal of our empirical study is to use the MFH as a benchmark and

* Corresponding author. Tel.: +44 1206 87 3975; fax: +44 1206 87 2684.
E-mail addresses: mkampo@essex.ac.uk (M. Kampouridis), chen.shuheng@gmail.com (S.-H. Chen), edward@essex.ac.uk (E. Tsang).

1057-5219/$ – see front matter. Crown Copyright © 2011 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.irfa.2011.06.009

examine how well it describes the empirical results which we observe from various markets. In particular, we are interested in knowing how this benchmark performs when we tune the key parameter, i.e., the number of types in the market. More details regarding tuning the number of trading strategies can be found in Section 7.

Therefore, the objectives of this paper can be summarized in the following way: (i) formalizing the MFH, (ii) suggesting a new agent-based financial model which does not assume pre-fixed types of trading strategies, (iii) suggesting a testing methodology for the MFH, and (iv) testing the hypothesis under empirical datasets.

The remainder of this paper is organized as follows. Section 2 formalizes the MFH by presenting its main constituents. Section 3 then presents a brief overview of the different types of agent-based financial models that exist in the literature, and also discusses their limitations. In order to address these limitations, we have created a new agent-based financial model, which is presented in Section 4. In addition, Section 4 presents the two techniques that our agent-based model uses, namely, Genetic Programming (GP) (Koza, 1992; Poli, Langdon, & McPhee, 2008) and Self-Organizing Maps (SOM) (Kohonen, 1982). Furthermore, Section 5 presents the experimental designs. Section 6 addresses the methodology employed to test the MFH and explains the technical approaches needed to be taken to facilitate the testing of the MFH. Section 7 presents the test results. It first starts by presenting the results over a single run for a single dataset. Then it continues by presenting the summary results over 10 runs for this dataset and it finally presents summary results for all datasets. Section 8 concludes this paper and briefly discusses possible directions for further research.

2. The Market Fraction Hypothesis

Within a market there exist different types of trading strategies. The Market Fraction Hypothesis (MFH) says that the fraction among these types of strategies keeps changing (swinging) over time. The following two statements are the basic constituents of the MFH, and are based on a summary of the empirical development of the agent-based financial models, presented in Chen et al. (2012).

1. In the short run, the fraction of different clusters of strategies keeps swinging over time, which implies a short dominance duration for any cluster. (Statement 1)
2. In the long run, however, different clusters are equally attractive and thus their market fractions are equal. (Statement 2)

The first statement means that it is not possible for a single strategy to dominate the market by attracting an overwhelming fraction of market participants for many consecutive periods. In other words, according to the MFH there is no such thing as a 'winner type'. Thus, an ex ante characterization of winners simply does not exist. The term 'dominance' will become technical for testing the MFH, and we shall make it precise in Section 7.

The second statement means that if for instance the market has two trading strategies (like the traditional fundamentalists–chartists model), their fraction should keep on swinging. Let us revisit the example we gave at the beginning of this paper. If at time t the fundamental strategy occupies 90% of the market participants and the chartist strategy occupies 10%, later on in another time period, these proportions could switch to 10% and 90%, respectively. So eventually *in the long run both types of traders will have enjoyed about the same market share*, i.e., about one half.[1]

As we can see, the implications of the MFH are very important. First of all, if the MFH holds, this means that all types of trading strategies that exist in a financial market will, at some point, become popular.[2] In other words, any type of trading strategy has an equal chance of attracting a significant amount of traders. Nevertheless, this seems unrealistic, because it means that even a bad strategy can become popular. It is thus interesting to investigate if this can happen under real data. In addition, another implication of the MFH is that no strategy can remain popular over a long period of time, as it will soon be succeeded by another popular strategy. It is therefore also interesting to examine if this is true, because it would give us a good picture of the microstructure dynamics of financial markets.

Hence, what we shall do in this paper is test the above MFH properties against our empirical dataset. More details about this will follow in the following sections. But before we talk about the tests, we first need to give a brief overview of the different designs of agent-based financial models, and explain how the limitations that exist in these models have led us to create a new agent-based model, which will be presented later, in Section 4. Section 3 thus presents this overview.

3. Agent-based financial models

Agent-based financial models are models of financial markets, where artificial agents can trade with each other. According to Chen et al. (2012), these models can be divided into two basic designs: the *N*-type design, and the autonomous agent design. The rest of this section presents these two designs.

3.1. Categories of agent-based financial models

In the *N*-type design, agents have beliefs regarding the price of a stock in the next time period. For instance, in the two-type design, there are two types of agent beliefs. Consequently, there are two types of *fixed and pre-specified* trading strategies. Each agent *can only choose between these two types*. These two types are usually the fundamentalists and the technical traders.[3] Many extensions of the *N*-type design exist. For example, a typical way to do this is by adding a memory factor to these rules. Other extensions can be to add an adaptive behavior, where the agents can learn from their previous experiences. Such examples can be found in (Brock & Hommes, 1998), where Brock and Hommes use 2-, 3-, and 4-type models. Other adaptive behavior examples include Kirman's ANT Model (Kirman, 1991, 1993) and Lux's Interactive Agent Hypothesis Model (Lux, 1995, 1997, 1998).

The second design of agent-based financial models is the autonomous agent design. In this type of model, we can have artificial agents who are autonomous and thus have the ability to discover new strategies, which have never been used before. An example of this is the well-known Santa Fe Institute (SFI) model (Arthur, Holland, LeBaron, Palmer, & Tayler, 1997; Palmer, Arthur, Holland, LeBaron, & Tayler, 1994), where a Genetic Algorithm (GA) (Holland, 1975) was used. Thus, a fixed number of strategies does not exist; on the contrary, each artificial agent can have a different trading strategy which is "customized" by a GA. SFI is of course not the only application of GA in artificial stock markets. Another example is AGEDASI TOF[4] (Izumi & Okatsu, 1996; Izumi & Ueda, 1999). If the reader is interested in these topics, a very good literature review can be found in (Chen, Huang, & Wang, 2009).[5]

[1] This idea is first made rigorous by Kirman (1993), who attempted to solve a puzzling entomological problem, i.e., ants swinging among themselves within two identical sources of food.

[2] Popularity is equivalent to dominance, which means that a strategy type occupies many market participants (strategies).

[3] Other equivalent names for technical traders are chartists, trend-followers and noisy traders.

[4] It stands for A GEnetic-algorithmic Double Auction SImulation in the TOkyo Foreign exchange market.

[5] It should also be said that apart from GA, other population-based learning models have been used, such as GP.

3.2. Limitations of the agent-based financial models

In the previous section, we described the two main agent-based financial designs: the *N*-type design and the autonomous agent design. The former design consists of *N* pre-specified strategy types, and the agents have to choose among these *N* types. An advantage of this design is that it allows us to observe the changes in the market fraction dynamics of the above strategy types. However, as we saw, a disadvantage of this type of model is that the agents are restricted in choosing from the given *N* strategy types. In addition, another limitation of this type of model is the lack of heterogeneity. Agents that belong in the same trading strategy type *must follow exactly the same behavioral rule*. Nevertheless, in the real world, the behavior of each trader is expected to be heterogeneous; even if some traders are following a certain trading strategy type, it does not mean that they behave in exactly the same way.

As we saw, the issue of heterogeneity is addressed by the autonomous agent models. This type of model allows the creation of autonomous and heterogeneous agents. Nevertheless, even under the autonomous agent models, agents have to choose among a pre-specified number *N* of trading strategy types. To the best of our knowledge, there is no model that uses autonomous agents that are not restricted to predefined, fixed strategy types.

This thus motivated us to create such a model. In order to do this, we used Genetic Programming as a rule inference engine, and Self-Organizing Maps as a clustering tool. The next section presents our model in detail.

4. Model

In this section, we present our agent-based financial model. This model first allows the creation of novel, autonomous and heterogeneous agents by the use of GP. The reason for using GP is because the market is regarded as an evolutionary process; this is inspired by Andrew Lo's Adaptive Market Hypothesis (AMH) (Lo, 2004, 2005), where Lo argued that the principles of evolution (i.e., competition, adaptation, and natural selection) can be applied to financial interactions. Thus, agents can be considered to be organisms that learn and try to survive.

After creating and evolving novel agents, we cluster them into types of trading strategies via SOM. These types are thus not pre-specified, but depend on the strategies of the agents.

The advantages of this approach are thus twofold: first of all, agents can create autonomous and heterogeneous trading strategies. Thus, even if two trading strategies belong to the same type of trading strategy (e.g., fundamental), it does not mean that these two strategies have to follow exactly the same trading rule, as it happens in the traditional agent-based model literature. In addition, when these trading strategies are categorized into types of trading strategies, they are not clustered into pre-specified, fixed types; on the contrary, the types depend on the existing trading strategies. T*his thus makes our model more realistic*.

The process we presented above is depicted in Fig. 1. Next, we present the two techniques of our model, GP and SOM. We first start by giving a brief introduction to each technique, then explain how these techniques were employed in our model, and finally provide some information on the algorithm of each technique.

4.1. Genetic Programming (GP)

Genetic Programming (GP) (Koza, 1992; Poli et al., 2008) is an evolutionary technique inspired by natural evolution, where computer programs act as the individuals of a population. The GP process has several steps. To begin with, a random population is created, by using terminals and functions appropriate to the problem domain, where the former are variables and constants of the programs, and the latter are responsible for processing the values of the system (either terminals or other functions' output).

In the next step, each individual is measured in terms of a pre-specified fitness function. The purpose of assigning a fitness to each individual is to measure how well it solves the problem. In the following step, individuals are chosen to produce new offspring programs. A typical way of doing this is by using the fitness-proportionate selection, where an individual's probability of being selected is equal to its normalized fitness value (Koza, 1992). The individuals chosen from the population are manipulated by genetic operators such as crossover and mutation, in order to produce offspring. The new offspring constitute the new population. Finally, each individual in the new population is

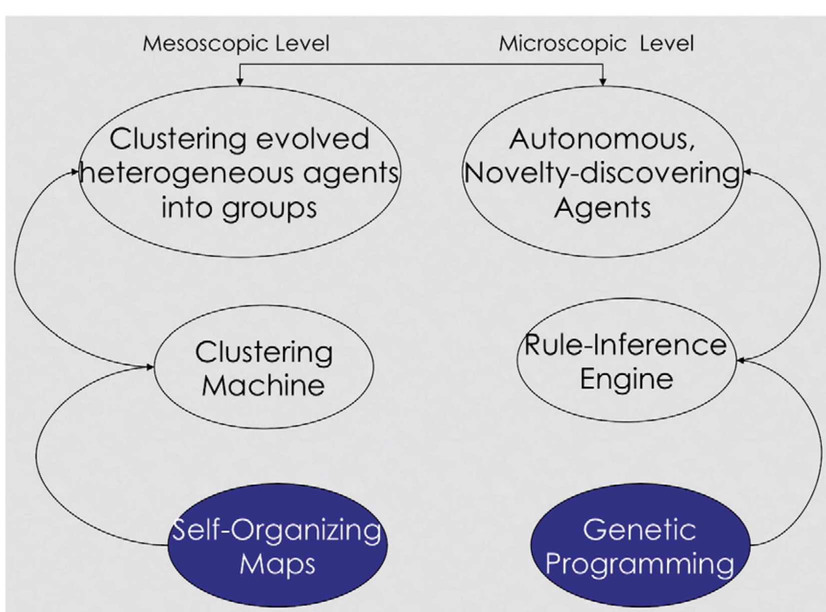

Fig. 1. Process followed in our agent-based financial model. First, novel agents are created through a Genetic Programming process. Then, these agents are clustered into types of trading strategies by the use of Self-Organizing Maps.

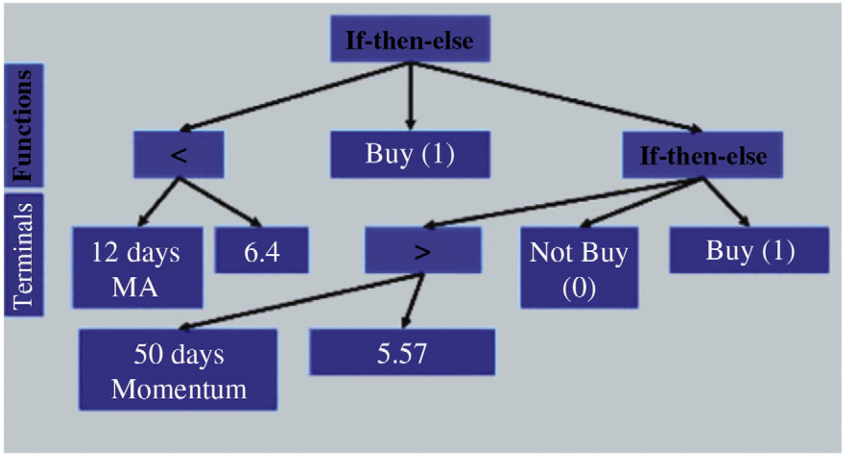

Fig. 2. Sample GDT generated by the GP.

again assigned a fitness and the whole process is repeated again, until a termination criterion is met (usually a number of generations). At the end of this procedure (last generation), the program with the highest fitness is regarded as the result of that run.

Next, we explain how GP was employed in our model.

4.1.1. GP as a rule-inference engine

First of all, 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. Using macro data to estimate micro behavior is not new, as many empirical agent-based models have already performed such estimations (Chen et al., 2012). However, such estimations are based on very strict assumptions, as we saw earlier (e.g., having pre-specified trading strategy types is considered to be a strict and unrealistic assumption). Since we no longer keep these assumptions, an alternative must be developed, and in this paper we recommend Genetic Programming (GP).

As we have already mentioned, the use of GP is motivated by considering the market as an evolutionary and selective process.[6] 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 pre-specified class of behavioral rules, like many other models in the agent-based finance literature (for instance, Brock & Hommes, 1998; Kirman, 1991, 1993). 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.[7]

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 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.[8] The GP algorithm used to infer the rules is presented in the next section.

4.1.2. GP algorithm

Our GP is inspired by a financial forecasting tool, EDDIE (Kampouridis & Tsang, 2010), which applies genetic programming to evolve a population of market-timing[9] 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: 12 and 50 days Moving Average (MA), 12 and 50 days Trader Break Out (TBR), 12 and 50 days Filter (FLR), 12 and 50 days Volatility (Vol), 12 and 50 days Momentum (Mom), and 12 and 50 days Momentum Moving Average (MomMA).[10] Each indicator has two different periods, a short- and a long-term one (12 and 50 days). Fig. 2 presents a sample GDT generated by the GP. As we can observe, this tree suggests that the trader should buy if the 12 days Moving Average is less than 6.4. If, however, this is not the case, the tree examines the 50 days Momentum; if it is greater than 5.57, the then GDT recommends not-to-buy. If, finally, the 50 days Momentum is less than or equal to 5.57, then the GDT recommends to buy.

Therefore, what the GP basically does is to generate similar trees to the one in Fig. 2. All these trees start with an If–Then–Else statement, and then the first branch will take a “comparison statement”, which checks whether a technical indicator (12 days MA in the Fig. 2 example) is greater than/less than/equal to a threshold. This threshold is a real number, which has been randomly created by the GP and is optimized through a hill-climbing process.[11] Then, depending on whether the comparison statement (e.g., 12 days MA $<$ 6.4) is true or false, the second or third branch of the tree is visited, respectively. Both of these branches can be either a suggestion (buy or not-buy,

[6] See Lo (2004, 2005) for his eloquent presentation of the *Adaptive Market Hypothesis*.

[7] This 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* (Gigerenzer & Todd, 1999; Simon, 1956).

[8] Duffy and Engle-Warnick (2002) provide the first illustration of using genetic programming to infer the behavioral rules of human agents in the context of ultimatum game experiments.

[9] ‘Market timing’ refers to the strategy of making buy or sell decisions regarding stocks, by attempting to predict future price movements.

[10] We use these indicators because they have been proved to be quite useful in previous works like Garcia Almanza (2008); Kampouridis and Tsang (2010); Martinez-Jaramillo and Tsang (2009). However, the purpose of this work is not to provide a list of the ultimate technical indicators. A brief presentation of these indicators, along with their formulas, is provided in Appendix A.1.

[11] Hill climbing is an optimization technique. It is an iterative algorithm that starts with an arbitrary solution to a problem. After evaluating this solution (in our case the threshold number), it then attempts to find a better solution by incrementally changing the solution (e.g., increasing the threshold by a certain value). If the change produces a better solution, an incremental change is made to the new solution, repeating until no further improvements can be found.

Table 1
Confusion matrix.

	Actual positive	Actual negative
Positive suggestion	True Positive (TP)	False Positive (FP)
Negative suggestion	False Negative (FN)	True Negative (TN)

denoted by 1 and 0, respectively), as is the case in the second branch, or even another tree, as is the case in the third branch.

After creating a GDT, we need to evaluate its performance. In order to do this, we first evaluate whether the GDT's suggestions were successful, e.g., where in reality the prices went up and thus the GDT correctly identified this as a buy opportunity. Depending on the classification of the GDT's suggestions, there are four cases: True Positive (TP), for a successfully identified buy opportunity, False Positive (FP), for a not-buy opportunity falsely identified as buy, True Negative (TN), for a successfully identified not-buy opportunity, and False Negative (FN), for a buy opportunity falsely identified as not-buy. These four together give the familiar confusion matrix (Provost & Kohavi, 1998), which is also presented in Table 1.

We then use the following 3 metrics, presented below:

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. We then use SOM to cluster these strategies into types of trading strategies.

4.2. Self Organizing Maps (SOM)

Self-Organizing Maps (SOM) (Kohonen, 1982) represent a type of artificial neural network which takes in input data with high dimensionality,[12] and returns a low-dimensional representation of these data, along with their topological representation. This representation is called a map. A self-organizing map consists of components called neurons. Associated with each neuron is a weight vector, which has the same dimensions as the input data. During the SOM procedure, the weight vector of each neuron is dynamically adjusted via a competitive learning process. Eventually, each weight vector becomes the center (a.k.a. centroid) of a cluster of input vectors. Thus, at the end of the SOM procedure, all input vectors have been assigned to different clusters of a map.

The next section presents how SOM was applied to our model.

4.2.1. SOM as a clustering machine

Once a population of rules is inferred from GP, it is desirable to cluster them based on a chosen similarity criterion. As we have already discussed at the beginning of Section 4, this allows us to cluster heterogeneous agents into different types of trading strategies, which are neither fixed, nor pre-specified.

The similarity criterion which we choose is based on the *observed trading behavior*.[13] 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.[14]

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, when we denote the decision "buy" by "1" and "not-buy" by "0", then the behavior of each rule (GDT) 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. Thus, each GDT can be represented by a vector which contains a series of 1 s and 0 s, denoting the tree's recommendations to buy or not-buy on each day. 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 (MacQueen, 1967) 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. This provides us with a rather convenient grasp of the dynamics of the microstructure directly as if we were watching the population density on a map over time.

4.2.2. SOM algorithm

For our experiments, we use MathWorks' Neural Network toolbox,[15] which is built in the MATLAB environment. This algorithm follows the standard SOM procedure presented earlier at the beginning of Section 4.2.

Fig. 3 presents the results after running 3×3 SOM for a population of 500 individuals for the daily TAIEX[16] index for the first and second half of 2007, respectively. Here, 500 artificial traders are grouped into nine clusters (types of trading strategies). 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. As we can see, there are usually a few strategies that are occupying the majority of the population, whereas the rest of the strategies have significantly fewer members. For instance, we can see that in the left map, 193 trading strategies have been clustered into the bottom-right cluster, 152 strategies have been clustered into the top-left cluster, 92 into the bottom-left cluster, and so on. Similar observations can be made for the second map, on the right of Fig. 3. Having different maps

[12] In this work the input data consist of the market-timing vectors of the GDTs.

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

[14] 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.

[15] http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/self_or4.html.

[16] Taiwan Stock Exchange Capitalization Weighted Stock Index. Available from http://finance.yahoo.com.

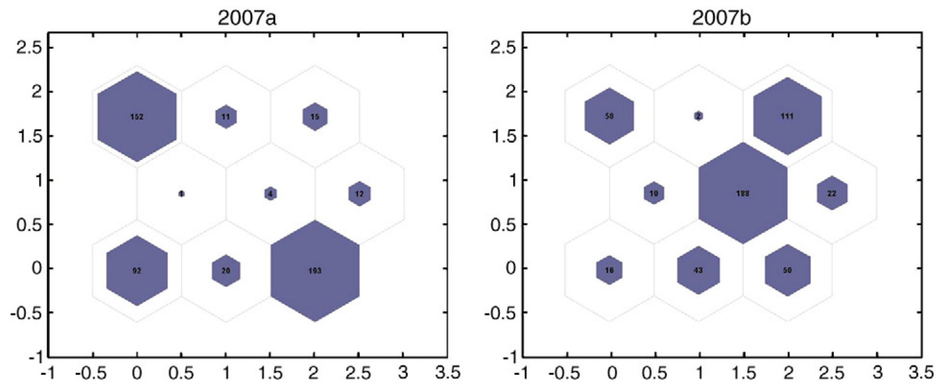

Fig. 3. 3 × 3 Self-Organizing Feature Maps. Two self-organized maps constructed from the rules inferred using the daily data of the TAIEX market, the first half (the left panel) and the second half (the right panel) of 2007, separately.

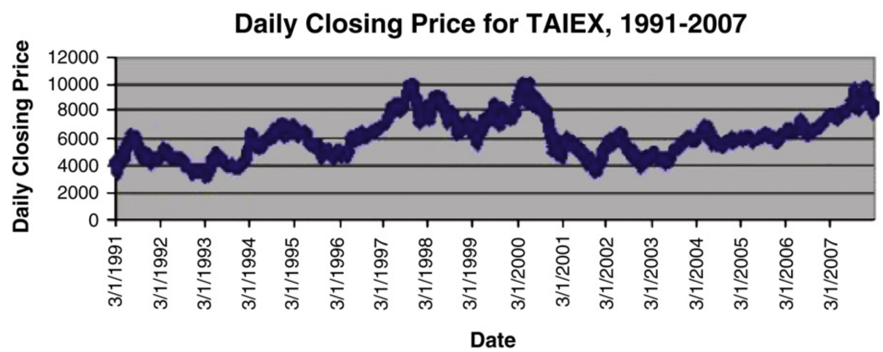

Fig. 4. Daily Closing Price for the TAIEX:1991–2007.

for different periods in time allows us to observe how the market fraction dynamics change from period to period, e.g., whether a cluster which occupies a high number of trading strategies will continue doing this in the future periods, and for how long.

This concludes the presentation of our agent-based financial model. Next, we present the experimental designs.

5. Experimental designs

The experiments were conducted for a period of 17 years (1991–2007) and the data were taken from the daily closing prices of 10 international market indices. These 10 indices are the CAC 40 (France), DJIA (USA), FTSE 100 (UK), HSI (Hong Kong), NASDAQ (USA), NIKEI 225 (Japan), NYSE (USA), S&P 500 (USA), STI (Singapore) and the TAIEX (Taiwan). For each of these indices, we run each experiment 10 times. To make it easier for the reader, we will first present the testing methodology and results for a single run of the TAIEX index. Fig. 4 presents the daily closing price of the TAIEX. We then proceed by presenting summary results over the 10 runs for all indices.

Each year was split into 2 halves (January–June, July–December), so in total, out of the 17 years, we have 34 periods.[17] The GP was therefore implemented 34 times. Table 2 presents the GP parameters for our experiments.

After generating and evolving GDTs for each one of the 34 periods, we used 3 × 3 SOM[18]; therefore, we obtained a total of 34 SOMs (one per period), with 9 clusters each. In other words, in every period the 500 GDTs were placed in one of the nine clusters (i.e., the categories of the trading strategies) of that SOM. Thus, we ended up with 34 different SOMs, one per semester, which represent the market in different time periods over the 17-year horizon. From this point on, whenever we use the term ‘trading strategy type’ we will be referring to one of the nine clusters and each GDT will be a member of one of these nine clusters.

Finally, it is important to say that the GP was only used for creating and evolving the trading strategies. No validation or testing took place, as is the case in the traditional GP approach. The reason for this is because we were not using the GP for forecasting purposes; instead, what we were interested in was to use the GP as a *rule inference engine* so that it could help us to see what the strongest species were during a certain period. To be more specific, the GP was used for each of the 34 periods and each time created and evolved trading strategies. After the evolution of the strategies under a specific period, these strategies (GDTs) were not tested against another set. This approach is consistent with Lo's AMH, as it states that the heuristics of an old environment are not necessarily suited to the new ones.[19] Furthermore, our no-testing approach is also consistent with the well-tested *overreaction hypothesis* (De Bondt & Thaler, 1985, 1987), which essentially states that top-ranked portfolios are outperformed by bottom-ranked portfolios during the next period.

6. Testing methodology

After having presented the necessary tools and the experimental designs, we can now proceed to present the testing methodology. Our

[17] At this point the length of the period was chosen arbitrarily as 6 months. We are aware that dividing the data in fixed semesters might “hide” some bias in the results, such as a seasonality effect. We leave the investigation of this to future research. For instance, one potential approach of this investigation could be the use of sliding-windows.

[18] The number of clusters (types of strategies) at this point was set arbitrarily. Later in this work we examine the sensitivity of the results if we tune this number.

[19] Lo refers to this possible situation as “maladaptive”. He also uses the example of the flopping of a fish for a better understanding of behaviors under different environments: on dry land the flopping might seem meaningless, but under water, it is the flopping that protects the fish from its enemies.

Table 2
GP Parameters. The GP parameters for our experiments are the ones used by Koza (1992). Only the tournament size has been changed (lowered), and the reason for that was because we were observing premature convergence. Other than that, the results seem to be insensitive to these 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
$\{w_1, w_2, w_3\}$	$\left\{1, \frac{1}{6}, \frac{1}{2}\right\}$

methodology consists of three parts: GP, SOM and the time-invariant SOM.

Let us start with GP. As we have already seen, we have used GP in order to generate and evolve trading strategies. However, there is a problem with comparing trading strategies from different periods. This happens because we cannot compare the fitness function of a trading strategy (GDT) from one period with the fitness function of a strategy (GDT) for another period, since they were presented with different datasets (environments).

This also applies to the clusters' comparison among different SOM maps. *Maps are not directly comparable*. In order to better understand this, consider Fig. 3. The way SOM works is that it creates the clusters after it is given a specific population of, in our cases, GDTs. When we have different periods with different populations, the nine clusters from different periods will generally be different, because they represent different populations of investment behaviors generated by different data environments. For example, if we name the bottom-left cluster of each SOM (Fig. 3) as 'Cluster 1', then we are saying that 'Cluster 1' of the SOM derived using the data for 2007a will in general not be the same as 'Cluster 1' of the SOM derived from the data using 2007b. It is quite likely that they will have different centroids (weighting vectors), representing different investment behaviors. This, therefore, makes the strategy types incomparable crossing different periods.

In order to tackle this problem, we introduce a time-invariant SOM based on the idea of *emigrating* and *reclustering*, which is the third and last part of our testing methodology. The following section thus presents these "translations" needed in order to make SOMs from different periods comparable so as to facilitate our tests for the MFH.

6.1. Translations

6.1.1. Emigrating

As we just mentioned, after obtaining the trading strategies from GP, we cannot directly compare them among different periods, because the dataset for each period is different. What, therefore, needs to be done is to apply the same dataset as a base to all periods. In other words, all GDTs that are derived from each period need to be applied to the dataset of the base period. For convenience, we call these *emigrant GDTs*. Therefore, after applying these emigrant GDTs to the base period, new signals are created. In this way, the market timing vectors of all GDTs originally derived from different periods are re-built based on the same grounds and hence become comparable. In this paper, we choose the second half of 2007 (2007b) as the base period.[20]

[20] The base period was chosen arbitrarily; however, we found that the results are insensitive to the choice of the base period.

6.1.2. Reclustering

Reclustering or time-invariant SOM is the second part of the translation, which allows the SOM clusters to be compared throughout different periods. We again use 2007b as the common base period. This time, we keep the centroids of the clusters originally derived from the common base period fixed and assign the market timing vectors from other periods (derived through emigrated GDTs) to one of the fixed centroids. This reclustering is conducted in the following way: the market timing vector of each emigrated GDT is compared with each centroid of the nine clusters and it will then be assigned to the one with the minimum Euclidean distance. We do this period by period from 1991a to 2007a. 33 SOMs are constructed in this way,[21] and now these SOMs can be directly compared with each other, given that they all share the same centroids. Fig. 5 presents 4 out of these 34 SOMs, where we can examine how the fraction of the clusters changes over time. These SOMs are now directly comparable over time. For instance, we can observe that although the bottom-right cluster of the top-left map (2006a) occupied a high number of trading strategies (390), this did not continue to take place in the future periods. Only 1, 6 and 50 trading strategies were clustered in this cluster in periods 2006b, 2007a, and 2007b, respectively. This figure thus gives a clear picture of what we mean by market fraction dynamics. As we can observe, the distribution over the clusters is uneven over time. In each period of time, some clusters obviously dominate others, but that dominance changes over time. This can be seen from the constant renewing of the major blocks.

We will thus use all 34 SOMs that have been generated for the years 1991–2007 to test Statements 1 and 2, and thus examine how the market fraction dynamics changes in the short and in the long run.

7. Results

This section presents the results of our tests. We ran two tests, one per each Statement of the MFH (see Section 2). Therefore, Test 1 investigates the plausibility of Statement 1, and Test 2 investigates the plausibility of Statement 2. In the following sections, we present the results of these tests first for a single run of a single dataset, TAIEX, then for 10 runs of TAIEX, and finally, for 10 runs of all 10 indices.

7.1. Results of a single run of a single dataset

7.1.1. Test 1: the short-run test

The first test regards the short-run behavior of market fractions. In the short run, the fraction of different clusters of strategies is expected to keep swinging over time, which implies a *short* dominance duration for any cluster. To be operational, a type of strategy is said to be *dominant* if its fraction is greater that the threshold,

$$TH = \frac{1+p}{N+p}, \tag{5}$$

where TH denotes a threshold, N is the number of clusters and p is a free parameter to manipulate the degree of dominance. For example, as $N = 9$, the threshold of being a dominant type changes with p as follows. It is 11.11% when $p = 0$, 20% when $p = 1$, and 27.27% when $p = 2$. Clearly, the higher the value of p, the higher the threshold. If all clusters were having the same number of members, then each cluster would be occupying 11% (1/9) of the population. Hence, the case where $p = 0$ corresponds to a threshold that just breaks the tie. However, to be dominant, we may expect a value of p to be higher than just breaking the tie. Hence, in this paper, p is set to be 2.

Furthermore, we need to be precise as to what we mean by *short duration* for a dominant type. Here, any specific number may be arbitrary; after all, short is only a matter of degree. We, therefore, first

[21] 2007b does not need reclustering, since we use it as the base period.

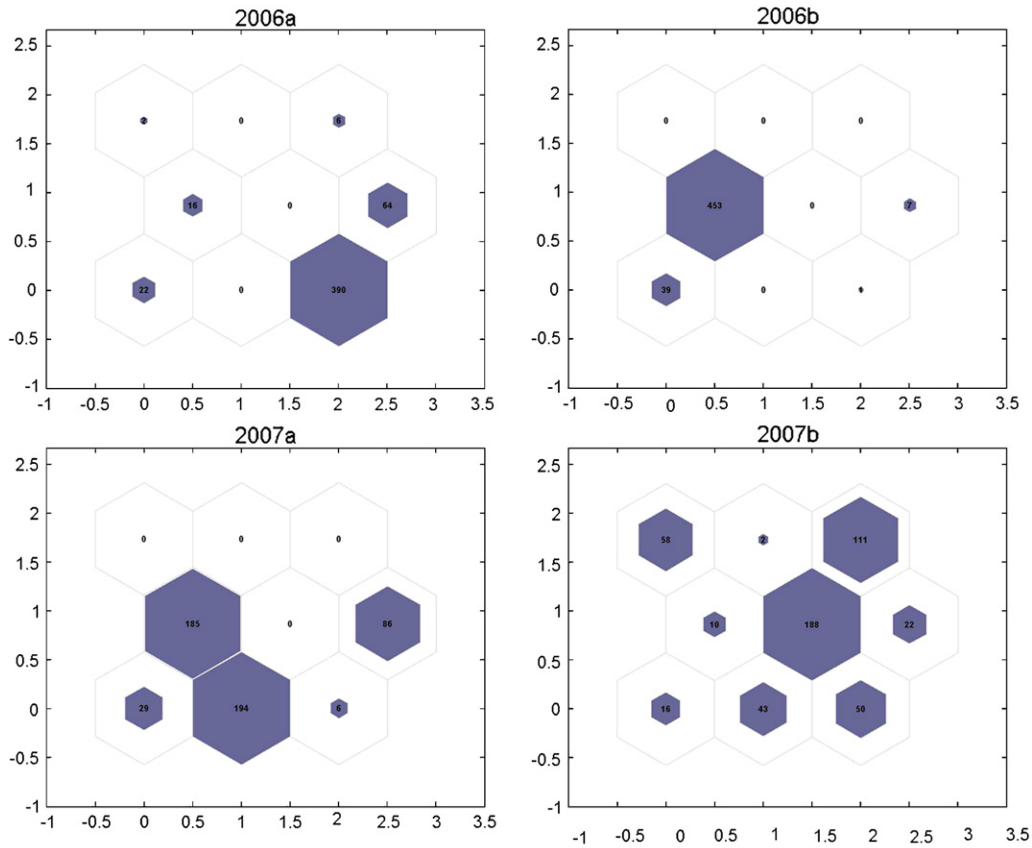

Fig. 5. The four SOMs above are constructed using the daily data of the TAIEX from 2006 to 2007. From the top-left panel to the bottom-right panel, they correspond to the first half and second half of the year 2006 (2006a, b) and the first half and second half of the year 2007 (2007a, b). Except for the last one, 2007b, the other three are reconstructed by using 2007b as the base period.

present the statistics of duration observed for each type. Fig. 6 summarizes the dominance results over the 34 periods. It presents the minimum, average and maximum of the duration times of each type. What we can observe from Fig. 6 is that the longest duration observed is nine periods (four and a half years) for type 6. For other types, the longest duration is barely over two periods. Hence, if we look at the average duration, with the exception of type 6, no type remains dominant for more than 2 consecutive periods, i.e., a year. Nevertheless, because of the long dominance duration of Cluster 6, we can argue that *evidence for the support of Test 1 is quite weak.*

7.1.2. Test 2: the long-run test

The second statement concerns the long-run behavior of market fractions. It says that, in the long run, different clusters are equally attractive and thus their market fractions are equal. As we said earlier,

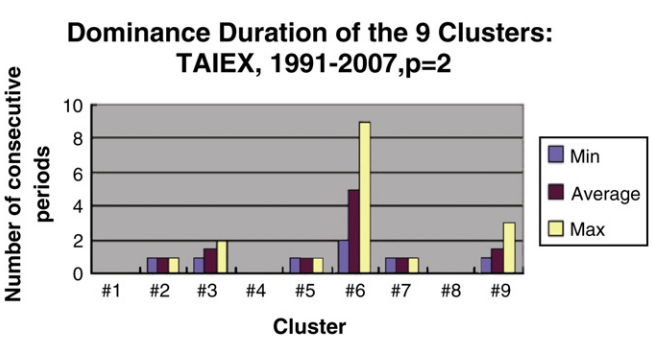

Fig. 6. Min, average and max number of consecutive times that a strategy remains dominant over the 34 periods for p = 2 (Daily Closing Price for the TAIEX:1991–2007).

we expect to see that the fraction of strategies keeps changing. In the left SOM of Fig. 3, for instance, we can see that three strategies are occupying a quite large fraction of the population (around 39%–193 members out of a total of 500, 30%–152 out of 500, and 18%–92 out of 500). The rest of the strategies have lower percentages. According to the MFH, these percentages should keep changing from period to period so that, in the long run, these percentages should be close to each other. In other words, if we have N types of traders, their long-term frequency of appearance should be close to $\frac{1}{N}$. Let $Card_{it}$ be the number (cardinality) of traders in Cluster i in time period t.

$$\sum_{i=1}^{N} Card_{it} = M, \forall t \tag{6}$$

In our current setting, M, the total number of traders is 500. The long-term histogram can be derived by simply summing up the number of traders over all periods and dividing it by a total of $M \times T$ (# of periods),

$$w_i = \frac{\sum_{t=1}^{T} Card_{it}}{M \times T} \tag{7}$$

Fig. 7 gives the long-term histogram of these clusters, $\{w_i\}$. Obviously, they are not equal and thus we present them in descending order from the left to the right. Cluster 6 has the largest market fraction of up to almost 60%, whereas Cluster 4 has the smallest market fraction, which is not even up to 1%.

Of course, it is obvious that this distribution is very different from the uniform one. In order to provide a measure of how far it is away from the uniform distribution, we use the familiar *entropy* as a metric.

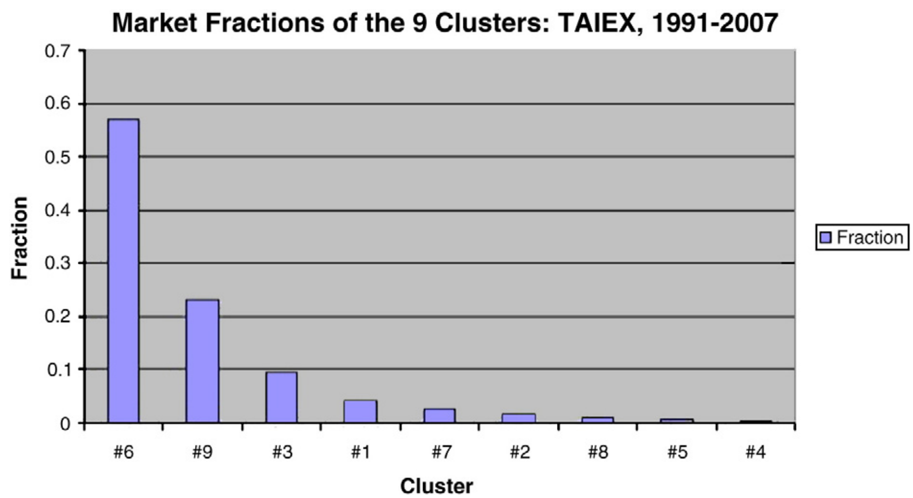

Fig. 7. Market Fractions of the nine clusters: TAIEX, 1991–2007.

Let us denote the empirical distribution presented in Fig. 7 as f_X, and the uniform distribution as f_Y. By definition, $f_Y = \frac{1}{N}$, where N is the number of clusters, which in this case is 9. In order to measure how close f_X is to the uniform distribution f_Y, we calculate the entropy of both distributions. For the discrete random variable, entropy is defined as

$$H = -\sum_{i=1}^{N} p_i \ln p_i, \quad (8)$$

where p_i is the fraction of each cluster. It is well known that for the uniform distribution $H(Y) = \ln N$. When $N=9$, it is $\ln 9 \approx 2.2$. The closer $H(X)$ is to 2.2, the closer X is to the uniform distribution. After calculating X's entropy, we find it equal to 1.3, which is only 59% of the entropy of the uniform distribution. This thus allows us to argue that f_X is far away from the uniform distribution, and hence *Test 2 is not supported by the TAIEX*.

Now that we have seen the test results of a single run for one dataset, it is interesting to see if these results can be generalized for more runs and more datasets. The next part of this section presents and discusses these summary results.

7.2. Summary results for all datasets under 9 clusters (3×3 SOM)

As we saw in the previous section, the experimental results of the two tests seem to deviate from what the MFH predicts to some extent. Test 1 has one cluster that dominates the market for 9 consecutive periods, which appears to be too long. In addition, Test 2 shows an even larger deviation since the long-term market fraction is very different from the uniform distribution. Altogether, the evidence for the MFH is weak. However, so far we have only presented a single run for a single dataset. Table 3 thus presents the results over 10 runs for all the datasets tested. The first two numeric columns are related to Test 1. They present the averages over the 10 runs for the average and maximum dominance duration of the 9 clusters. Furthermore, the last column is related to Test 2 and shows the ratio of the average realized entropy (over the 10 runs) over the base entropy (equal to 2.2).

The first observation we can draw from Table 3 is that homogeneity exists across the majority of the results. Let us first start with Test 1. We can see that on average there is no cluster that remains dominant for 2 consecutive periods. This is in line with Test 1. However, the second column tells us that even though on average no cluster dominates for more than 1 period, there is always an outlier that can remain dominant for longer, e.g., 8 consecutive periods for the TAIEX. Regarding Test 2, the entropy ratios for all datasets are somewhat distant from their maximum values. All entropy ratios are in the range 0.55–0.75, which is basically a 25–45% difference from the entropy of the uniform distribution. This essentially means that the distributions are on average different from the uniform distribution and therefore the clusters, in the long run, are not equally attractive, as Test 2 requires. *Overall, the MFH seems to be relatively weak for all 10 indices tested under the 3×3 SOM.*

7.3. Changing the number of clusters

So far, all of our tests have been performed for 3×3 SOMs. It is, however, interesting to investigate how sensitive the results are if we tune the number of clusters. Therefore, we repeat the whole procedure mentioned above for different SOM dimensions: 2×1, 3×1, 2×2, 5×1, 3×2, 7×1 and 4×2, i.e., from 2 clusters to 8 clusters.

7.3.1. Test 1

Fig. 8 presents the averages, over 10 runs, for the average (left panel) and maximum (right panel) dominance duration for numbers of clusters 2–9. The x-axis presents the number of types of trading strategies (clusters), and the y-axis the dominance duration. What we observe in these graphs, especially in Fig. 8(a), is that the dominance duration decreases as the number of clusters increases. Nevertheless, we can again see from Fig. 8(b) that there are always clusters with strong dominance, even under the 3×3 SOM. *Test 1 thus is not supported under any number of clusters.*

Table 3
Summary results over 10 runs, for all datasets, for a 3×3 SOM.

Summary statistics	Test 1		Test 2
	Average	Max	Entropy ratio
CAC 40	1.81	5.5	0.68
DJIA	1.93	5.78	0.66
FTSE 100	1.77	6	0.64
HSI	1.71	4.6	0.7
NASDAQ	1.59	4.11	0.69
NIKEI 225	1.51	3.4	0.79
NYSE	1.93	6.56	0.6
S&P 500	2.16	6.89	0.64
STI	1.67	3.7	0.75
TAIEX	2.02	8.25	0.55

The first two numeric columns are related to Test 1 and present the averages over the 10 runs for the average and maximum dominance durations of the 9 clusters, respectively. The last column presents the ratio of the average realized entropy (over the 10 runs) over the base entropy under the null of the uniform distribution. This ratio is maximized when it is equal to 1.

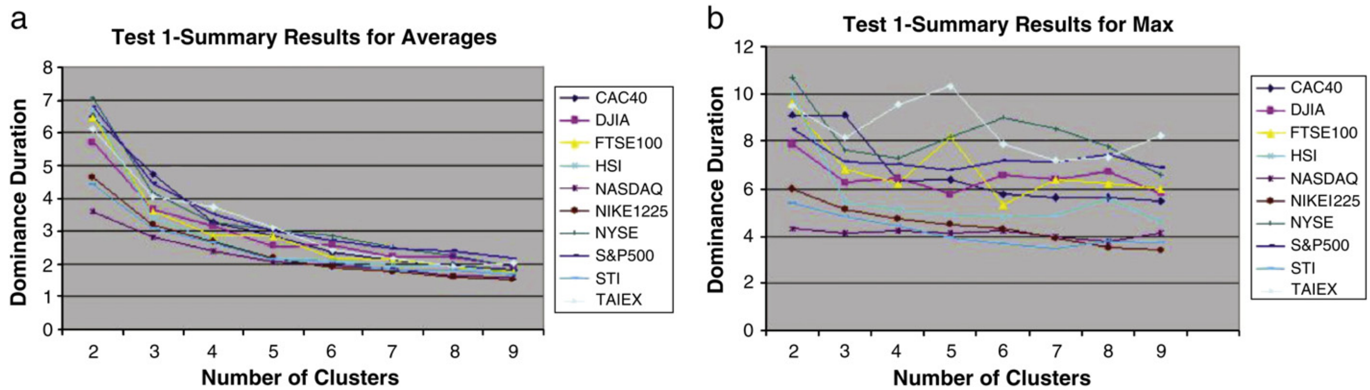

Fig. 8. Summary results for Test 1-Averages and Maximums, Financial Data.

To see how significant or how interesting this pattern is, we run a Monte Carlo simulation as follows. Starting with two clusters, we randomly assign a winner (dominant cluster) to either Cluster 1 or Cluster 2. We then conduct this binomial experiment 34 times. Considering this to be one run, we do it for ten runs. Hence, we have 10 artificial series of dominant clusters, with each series lasting for 34 runs. We then conduct the same analysis as above by figuring out the average duration and maximum duration of each series, and the average of the whole. We then apply this Monte Carlo experiment with an additional number of clusters, from three to nine incrementally. A comparable result is then drawn in Fig. 9.

By comparing Fig. 8 with Fig. 9, we can see that the behavior of the real markets is very different from that of the multinomial experiment. For the latter, the average of the maximum duration decays from above 6 to below 3, but for the former this decaying tendency is shown in none of the ten indices. Instead, they all fluctuate slightly around a horizontal line, and, depending on the market, the line is situated at an interval from four to eight. For the average duration, while both figures feature a decaying tendency, the one with financial data decays much more slowly than the one based on the artificial data. Therefore, our result cannot be treated as an incident from a random draw of the multinomial experiments, and in this sense this pattern is not spurious.

One last thing that we would like to point out is the low average dominance duration that can be observed in Fig. 8(a) for the high number of clusters. One might wonder what the reason for this 'phenomenon' is, especially since the average dominance duration started at quite high levels (for the low number of clusters). We believe that this can be better explained if we also take in to account Fig. 8(b). What seems to happen for all datasets is that there are always a few clusters that have strong (long) dominance over the 34 periods, whereas the rest have very low dominance. The low average dominance duration we see in Fig. 8(a) for the high number of clusters can therefore be explained by the extremely low dominance duration of the majority of clusters.[22]

7.3.2. Test 2

As we said earlier, we are interested in obtaining the distance of the entropy of the empirical distribution f_X (fractions of clusters) from the uniform distribution (benchmark). We have also said that the closer the entropy of distribution f_X is to the entropy of the uniform distribution, the closer distribution f_X is to the uniform one. After obtaining the entropies over 10 runs for each dataset, we first calculated the average of these runs. We then divided each one of these averages by the benchmark entropy and thus obtained 10 different ratios (one per dataset). Of course, this ratio is maximized when the two entropies are equal, and therefore their ratio is equal to 1. Hence, the higher the ratio, the closer to the uniform distribution the empirical distribution will be. Fig. 10 presents these ratios for all datasets.

What we observe from this figure is that the ratios tend to decrease as the number of clusters increases, and hence *the support for Test 2 gets weaker*. Such a divergence of the two distributions indicates again that the strong dominance of a few clusters continues to exist, even in the long run. Therefore, after combining Tests 1 and 2, we can have a quite clear picture. *Clusters tend to dominate for long periods and this dominance is usually interchanged among a few clusters*.

To make this argument even clearer, we also present Fig. 11, which shows the cumulative fractions for the TAIEX for different number of clusters. A graph named 'Number of Clusters: 2' means that the strategies are allocated to 2 clusters. When 'Number of Clusters: 3', the strategies are allocated 3 clusters, and so on. The clusters have been sorted by their size (fraction) in descending order. Therefore, Cluster 1 on the x-axis denotes the cluster with the highest fraction, Cluster 2 the cluster with the second highest fraction, etc. The y-axis presents the cumulative fraction of the clusters.

An observation we can make from Fig. 11 is that the contribution of each ranked cluster decreases when the number of clusters increases, which causes the entire cumulative curve to shift down.[23] For instance, when the number of clusters is 2, the largest cluster has a size of about 70%. However, while we move to a higher number of clusters, we can see that this size gradually decreases and finally falls below 60%, when the number of clusters is 9. The same happens for the contribution of the rest of the clusters. Hence, each graph moves a bit below, when the number of clusters increases. Nevertheless, what is important is that even when the number of clusters is 9, the clusters with the highest two or three ranks occupy a very large fraction (approximately 90%) of market participants.

The above observation leads us to consider that maybe there is a minimum number of clusters that covers a certain fraction of the market. Table 4 gives the result of the minimum number of clusters required to cover a targeted fraction of market participants. The three targeted values given in the table are 90%, 95% and 99%. Since the purpose was to see whether only a small number of clusters is required, we started with a larger number of clusters, namely, nine, and saw how large a reduction we could make. If the target is to cover 90% of the market participants, then most indices need four to five types, and if the target rises even higher to 95%, then most indices

[22] For instance, if 8 out of the 9 clusters have a dominance duration of 2 periods, and only 1 cluster has a dominance duration of 9 periods, then the average dominance duration is driven down to $\frac{(8*2)+9}{9} \approx 2.7$ periods.

[23] The graphs for the other indices can be found in Appendix A.2.

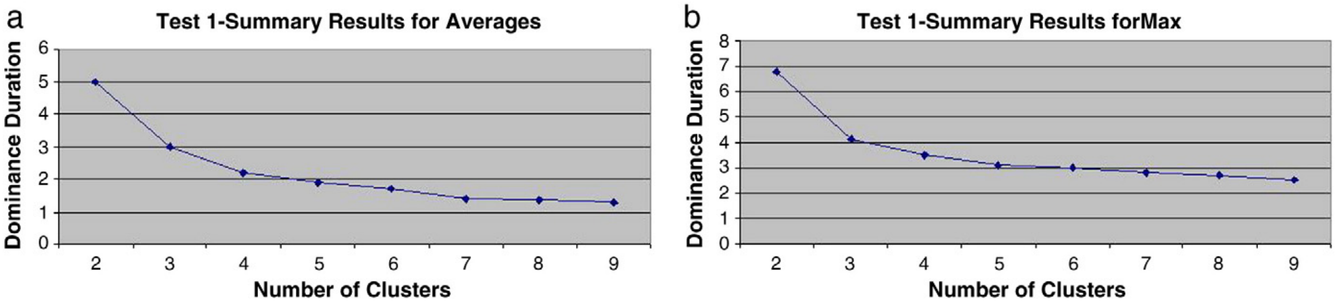

Fig. 9. Summary results for Test 1-Averages and Test 1-Maximums under Monte Carlo simulation.

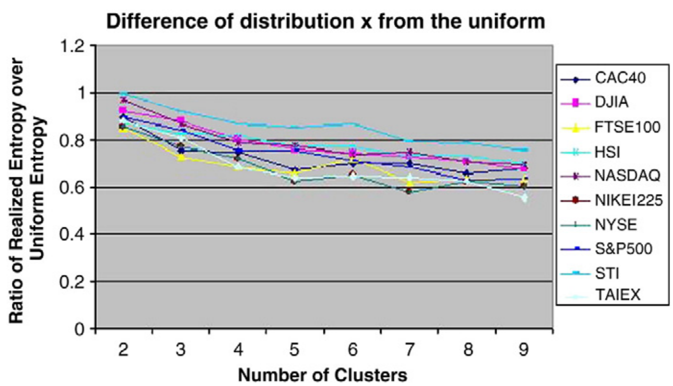

Fig. 10. Test 2: difference of the empirical distribution x (fractions of clusters) from the uniform distribution.

need five to six types. The percentage 95% is also the parameter used in Aoki (2002).[24] Hence, our finding regarding the minimum number required is also larger than Aoki's suggested two or three. Two or three types can be sufficient if we have a target somewhat lower than 90%.

8. Conclusion

To summarize, this paper has presented the Market Fraction Hypothesis (MFH), which states that the market fraction of the types of trading strategies that exist in a financial market changes over time. However, this hypothesis had never been formalized in the past. Our first contribution was thus to formalize the hypothesis. This then allowed us to suggest a testing methodology and test the statements of the hypothesis under 10 financial markets. The latter was very important, because until now the observations of the MFH had only been made under artificial market frameworks. In addition, in order to test the hypothesis we proposed a new agent-based financial model. The novelty of this model was that it did not assume pre-fixed types of trading strategies, as is typically the case in the agent-based financial literature (Chen et al., 2012). Finally, another contribution of this paper was the introduction of time-invariant SOM, a novel tool that allowed the comparison of SOMs among different time periods.

The experimental results showed that the MFH seems to be weak for the majority of the datasets we tested. More specifically, we found that, even in the long run, *the market tends to favor few types of agents, hence the property of the long-term uniform distribution (Test 2) does not hold*. We also found that while the results above are qualitatively insensitive to the number of types, a parameter set in the test, we only need four to five (five to six) types, to account for the behavior of 90% (95%) of market participants. Finally, while most types of agents cannot be dominant consecutively for more than 2 years, few exceptions can sustain up to 4 years. Therefore, *the property of short-term dominance duration (Test 1) also does not hold*.

The above observations lead us to valuable conclusions about market fraction dynamics. From our experimental results we can argue that popular (i.e., dominant) types of trading strategies can remain popular over long time periods. In addition, there can indeed be a swinging among the fractions of the different types of trading strategies; however, this swinging exists only among the few popular

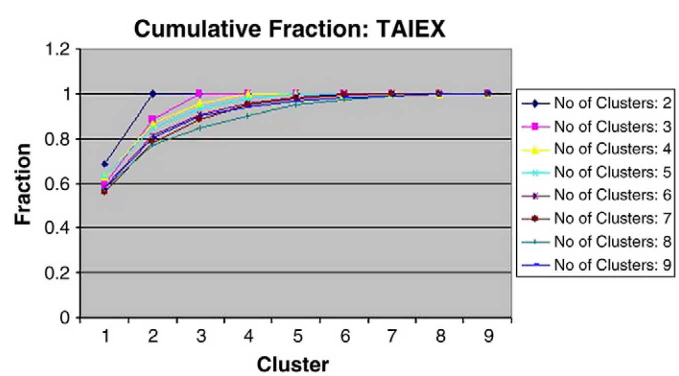

Fig. 11. Cumulative Fraction for TAIEX.

Table 4
Minimum number of clusters whose cumulative fraction is above the required threshold of 90%, 95% and 99%, respectively.

Minimum number of clusters			
	Threshold		
	90%	95%	99%
CAC 40	4	5	7
DJIA	5	6	8
FTSE 100	4	5	8
HSI	5	6	8
NASDAQ	4	6	8
NIKEI 225	5	6	8
NYSE	7	8	9
S&P 500	7	8	9
STI	5	6	8
TAIEX	4	6	7

[24] Aoki (2002) is probably the only paper known to us that deals with a number of types of agents in the multi-agents system. Using the Ewens–Pitman–Zabell induction method, Aoki applied the result from the evolution of biological species and population genetics to determine the minimum number of types of behavior required to capture multi-agent economic systems. He showed that it would be enough to characterize the market behavior by a few types, say, two to three. Others were rather marginal.

types. Thus, even if many strategy types exist in a market, only a few of them become and remain popular over time. It is therefore enough to characterize the market behavior by using a few types of trading strategies, even if many more types exist in this market.

Future research can include some changes in our model. Can the results be affected by the periods' window? In this work, the 17-year dataset was divided into semesters and thus each window was 6 months. It would be interesting to see whether a shorter or longer window can affect our results. Finally, we are interested in examining the sensitivity of our results to the tools used to process the data. Can any features which we obtained using GP or SOM be valid if different rule-inference machines, clustering techniques, or simply just different settings of GP or SOM are employed? For example, the use of *standard hierarchical clustering* (Xu & Wunsch, 2008) or the growing hierarchical self-organizing map (Dittenbach, Rauber, & Merkl, 2001) can provide us with much finer details of the hierarchical structure of the market participants, which has not yet been well exploited in the literature.

Acknowledgments

An early version of this paper was presented at the Econophysics Colloquium 2010, Academia Sinica, Taipei, Taiwan, November 4–7, 2010. The authors benefited significantly from the discussions with conference participants. This version has been substantially revised in light of two anonymous referees' very painstaking reviews, for which we are most grateful. The Taiwan NSC grant 98-2410-H-004-045-MY3 and EPSRC grant EP/P563361/01 are also gratefully acknowledged.

Appendix A. Appendix

Appendix A.1. Technical indicators

The following section presents the technical indicators that the GP is using, along with their formulas. We performed a sort of standardization in order to avoid having a very large range of numbers generated by GP, because this would increase the size of the search space even more.

Moving Average indicator (MA)

$$\mathrm{MA(L,t)} = \frac{P(t) - \frac{1}{L}\sum_{i=1}^{L} P(t-i)}{\frac{1}{L}\sum_{i=1}^{L} P(t-i)}.$$

By using the MA, traders are able to observe any changes in the trend of the prices of a stock. Typically, when a short-term MA (e.g., 12 days) goes above a long-term MA (e.g., 60 days), this indicates an upward momentum. On the other hand, when a short-term MA goes below a long-term one, this indicates a downward momentum.

Trade Break Out indicator (TBR)

$$\mathrm{TBR(L,t)} = \frac{P(t) - max\{P(t-1),\ldots,P(t-L)\}}{max\{P(t-1),\ldots,P(t-L)\}}.$$

In order to understand this indicator better, we first need to explain two terms: *support* and *resistance*. Support is the point where the price stops going any further down, whereas resistance is the point where the price does not go up any further. Technical analysts suggest that price downward trends tend to reverse at support points, whereas upward trends tend to reverse at resistance points. However, when these points are breached (break out), perhaps because of some new information regarding the market, it is likely that the price will continue in the same direction. Hence, traders tend to observe these breakouts and when a stock goes above its point of resistance, they buy; when on the other hand the stock price goes below its point of support, traders sell.

Filter indicator (FLR)

$$\mathrm{FLR(L,t)} = \frac{P(t) - min\{P(t-1),\ldots,P(t-L)\}}{min\{P(t-1),\ldots,P(t-L)\}}.$$

This indicator is used to indicate buy or sell actions, depending on whether the price movement goes in the opposite direction by a predefined percentage. For instance, if the price reverses from a downward trend and rises by a specific percentage from the low price that it was previously, then the trader will perform a 'buy' action.

Volatility indicator (Vol)

$$\mathrm{Vol(L,t)} = \frac{\sigma(P(t),\ldots,P(t-L+1))}{\frac{1}{L}\sum_{i=1}^{L} P(t-i)}.$$

A period of increasing volatility could indicate a reverse in the trend or strong downward trends. This would thus give an indication to a trader that he should be cautious. On the contrary, when there is a period of decreasing volatility, this indicates upward trends and traders should buy.

Momentum indicator (Mom)

$$\mathrm{Mom(L,t)} = P(t) - P(t-L).$$

The Momentum indicator measures the acceleration or speed at which a stock's price is changing. Traders use this indicator because they believe that a strong trend will likely persist for a period of time.

Momentum Moving Average indicator (MomMA)

$$\mathrm{MomMA(L,t)} = \frac{1}{L}\sum_{i=1}^{L} Mom(L, t-i).$$

This is basically a calculation of the moving average of the momentum, which was presented above.

Appendix A.2. Figures of cumulative fractions

In this section we present the figures of cumulative fractions for the remaining 9 market indices (apart from the TAIEX which was presented earlier in Section 7.3.2). We can again see here that most of the markets tend to have a few gigantic clusters. The only exceptions are the NYSE and S&P500.

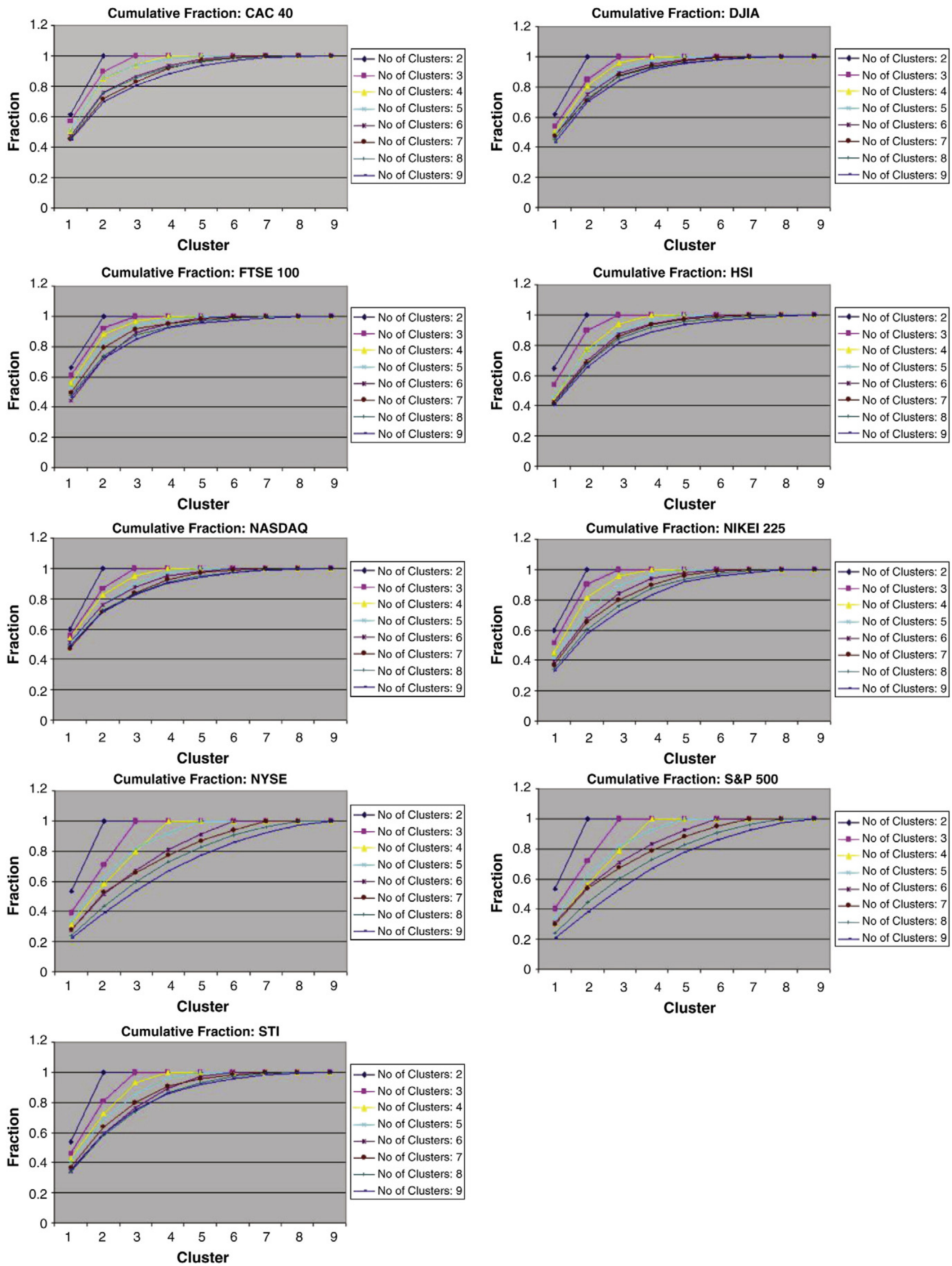

Fig. A.12. Cumulative fractions for the 9 datasets.

References

Amilon, H. (2008). Estimation of an adaptive stock market model with heterogeneous agents. *Journal of Empirical Finance*, *15*(2), 342–362.

Aoki, M. (2002). Open models of share markets with two dominant types of participants. *Journal of Economic Behavior and Organization*, *49*(2), 199–216.

Arthur, W., Holland, J., LeBaron, B., Palmer, R., & Tayler, P. (1997). Asset pricing under endogenous expectations in an artificial stock market. In B. Arthur, S. Durlauf, & D. E. Lane (Eds.), *The economy as an evolving complex system II* (pp. 15–44). Reading, MA: Addison-Wesley.

Boswijk, P., Hommes, C., & Manzan, S. (2007). Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control*, *31*(6), 1938–1970.

Brock, W., & Hommes, C. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, *22*, 1235–1274.

Chen, S. H. (2008). Financial applications: stock markets. *Wiley encyclopedia of computer science and engineering* (pp. 481–498). John Wiley & Sons, Inc.

Chen, S. H., Chang, C. L., & Du, Y. R. (2012). Agent-based economic models and econometrics. *Journal of Knowledge Engineering Review*, *27*(1).

Chen, S. H., Huang, Y. C., & Wang, J. F. (2009). Bounded rationality and the elasticity puzzle: An analysis of agent-based computational consumption capital asset pricing models. In S. Zambelli (Ed.), *Computable, constructive and behavioural economic dynamics: Essays in honour of Kumaraswamy (Vela) Velupillai*. Routledge.

De Bondt, W., & Thaler, R. (1985). Does the stock market overreact? *Journal of Finance*, *40*, 793–805.

De Bondt, W., & Thaler, R. (1987). Further evidence on investor overreaction and stock market seasonality. *Journal of Finance*, *42*(3), 557–581.

Dittenbach, M., Rauber, A., & Merkl, D. (2001). Recent advances with the growing hierarchical self-organizing map. In N. Allinson, H. Yin, L. Allinson, & J. Slack (Eds.), *Proceedings of the 3rd workshop on Self-Organizing Maps* (pp. 140–145). Lincoln, England: Springer.

Duffy, J., & Engle-Warnick, J. (2002). Using symbolic regression to infer strategies from experimental data. In S. -H. Chen (Ed.), *Evolutionary computation in economics and finance* (pp. 61–82). Springer.

Garcia Almanza, A., 2008. New classification methods for gathering patterns in the context of Genetic Programming. Ph.D. thesis. Department of Computer Science, University of Essex.

Gigerenzer, G., & Todd, P. (1999). In G. Gigerenzer, P. Todd, & the ABC Research Group (Eds.), *Fast and frugal heuristics: The adaptive toolbox* (pp. 3–34). Oxford University Press.

Gilli, M., & Winker, P. (2003). A global optimization heuristic for estimating agent-based models. *Journal of Computational Statistics and Data Analysis*, *42*(3), 299–312.

Holland, J. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.

Izumi, K., & Okatsu, T. (1996). An artificial market analysis of exchange rate dynamics. In L. J. Fogel, P. J. Angeline, & B.T.E. (Eds.), *Evolutionary programming V* (pp. 27–36). MIT Press.

Izumi, K., & Ueda, K. (1999). Analysis of dealers' processing financial news based on an artificial market approach. *Journal of Computational Intelligence in Finance*, *7*, 23–33.

Kampouridis, M., & Tsang, E. (2010). EDDIE for investment opportunities forecasting: Extending the search space of the GP. *Proceedings of the IEEE Conference on Evolutionary Computation, Barcelona, Spain* (pp. 2019–2026)..

Kirman, A. (1991). Epidemics of opinion and speculative bubbles in financial markets. *M. Taylor (Hrsg.), "Money and Financial Markets"* (pp. 354–368). London: Macmillan.

Kirman, A. (1993). Ants, rationality and recruitment. *Quarterly Journal of Economics*, *108*(1), 137–156.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Journal of Biological Cybernetics*, *43*, 59–69.

Koza, J. (1992). *Genetic Programming: On the programming of computers by means of natural selection.* Cambridge, MA: MIT Press.

Lo, A. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, *30*, 15–29.

Lo, A. (2005). Reconciling efficient markets with behavioral finance: The adaptive markets hypothesis. *Journal of Investment Consulting*, *2*, 21–44.

Lux, T. (1995). Herd behavior, bubbles and crashes. *The Economic Journal*, *105*, 880–896.

Lux, T. (1997). Time variation of second moments from a noise trader/infection model. *Journal of Economic Dynamics and Control*, *22*, 1–38.

Lux, T. (1998). The socio-economic dynamics of speculative markets: Interacting agents, chaos and the fat tails of return distributions. *Journal of Economic Behavior and Organization*, *33*, 143–165.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297). University of California Press.

Martinez-Jaramillo, S., & Tsang, E. (2009). A heterogeneous, endogenous and co-evolutionary GP-based financial market. *IEEE Transactions on Evolutionary Computation*, *13*(1), 33–55.

Palmer, R., Arthur, W., Holland, J., LeBaron, B., & Tayler, P. (1994). Artificial economic life: A simple model of a stock market. *Physica D*, *75*, 264–274.

Poli, R., Langdon, W., & McPhee, N. (2008). A field guide to Genetic Programming. http://Lulu.com

Provost, F., & Kohavi, R. (1998). Glossary of terms. *Journal of Machine Learning*, *30*(2–3), 271–274.

Simon, H. (1956). Rational choice and the structure of environments. *Psychological Review*, *63*, 129–138.

Winker, P., & Gilli, M. (2001). Indirect estimation of the parameters of agent-based models of financial markets. *FAME Research paper, University of Geneva*, *38*.

Xu, R., & Wunsch, D. (2008). *Clustering: Wiley-IEEE Press.*