4

A Broad-Spectrum Computational Approach for Market Efficiency

Olivier Brandoy[1] and Philippe Mathieu[2]

[1] LEM, UMR CNRS-USTL 8179, France olivier.brandouy@univ-lille1.fr
[2] LIFL, UMR CNRS-USTL 8022, France philippe.mathieu@lifl.fr*

4.1 Introduction

The Efficient Market Hypothesis (EMH) is one of the most investigated questions in Finance. Nevertheless, it is still a puzzle, despite the enormous amount of research it has provoked. For instance, many recent results have shadowed the well-established belief that market cannot be outperformed in the long run (Detry and Gregoire [2]).

One other reason is that persistent market anomalies cannot be easily explained in this theoretical framework Shiller [11]. Additionally, one can also consider that some talented hedge-fund managers (like Jim Simons) keep earning excess risk-adjusted rates of returns since years. Nevertheless, there is no consensus on this last point today Malkiel [7].

Many versions of the EMH have been proposed since the founding works of Samuelson [10]. We concentrate in this paper on the weak form of efficiency Fama [3]: *"past informations are useless to predict future price changes"*. We, therefore, focus on the efficacity of simple technical trading rules, following Jensen and Benington [6] or more recently Brock et al. [1]. An extensive survey for this issue is proposed in Park and Irwin [8].

Nevertheless, we depart from previous works in many ways: we first have a large population of technical, virtual agents (more than 260.000) exploiting real-world data to manage a financial portfolio as chartists or technical traders would do. Very few researches have used such a large amount of calculus to examine the EMH. Our experimental design allows for agents selection based on past absolute performance, as well as consistency of performance. We take into account the data-snooping risk, which is an unavoidable problem in such broad-spectrum researches, using a rigorous *Bootstrap Reality Check* (BRC) procedure [12].

* This work has received a grant from European Community – FEDER – and *"Region Nord-Pas de Calais"* – CPER TAC –.

While market inefficiencies, after including transaction costs, cannot clearly be successfully exploited, our experiments present troubling outcomes like persistent (but not statistically significant at commonly admitted levels) over performance, inviting close re-consideration of the weak-form EMH.

This research is organized as follows: section 4.2 presents our multi-agent system (MAS) and experimental design, section 4.3 is dedicated to our results and section 4.4 concludes this research.

4.2 Methodology

The methodological section gives the main features of our experimental design, including the global MAS architecture, descriptions of the agents, and the statistical procedure aimed at detecting potential market inefficiencies.

4.2.1 MAS Architecture

In this experiment, agents represent virtual investors trading a single financial commodity called *"a tracker"*. As it is generally admitted [13], the agents' fundamental characteristics in this study are an idiosyncratic decision-making process, autonomy and reactivity to contextual changes. Our MAS is based on a three-stage architecture (see figure 4.2); at each stage, one can consider a particular kind of agents with specific aims or logic:

First stage: *Strategic Agents* are micro-agents always playing the same basic strategy through the entire simulation. Those basic strategies are known in the financial community as "technical trading rules". For instance, a *Strategic Agent:"5-days moving average"* cannot process any other operation and has to decide whether to trade or not on the basis of a single rule.

Second stage: *Family Agents* are general agents defining all the formal characteristics used in the instantiation of each *Strategic Agent*. Each *Family Agent* also has to perform a ranking between each of his "children" at each time step. The *Family Agent*, thus, has the capacity to select the most successful individuals among the *Strategic Agents*. For instance, the *Family Agent "Rectangle"* combines four parameters (n, m, p, and s, see figure 4.1).

Third Stage: *Meta Agents* are able to mimic the behavior of various *Strategic Agents* according to the circumstances and the ranking given by the relevant *Family Agent*. For instance, a *Meta Agent* based on the 2-uple $\{Momentum, Triangle\}$ will choose and mimic various instances of those *Family Agents*, after considering some signals. For instance, it can begin with replicating a *Strategic Agent:"5-days momentum"* and then keep on going with this for eight days, then switch to replicate a *Strategic Agent:"3-days triangle"* for the next six iterations and so on... We do not develop this point in this article, nor do we report results concerning this category of agents.

To get it clearer, let's consider one *Family Agent: "Periodical Trader"*. This agent buys and sells the trackers at fixed intervals. It is similar to speculators buying on Mondays and selling on Fridays. This agent has at least two parameters coding the dates on which it will buy and sell the trackers. If it decides to generate all possible *Strategic Agents* using all possible delays between 1 and 100, 10.000 "children" will be processed. In this study, we have 10 *Family Agents* generating more than 260.000 *Strategic Agents*. One can imagine that the number of *Meta Agents* is, therefore, really huge and, despite computer power or parallel computing facilities, cannot be investigated exhaustively.

4.2.2 Agent's Design

Agent's design is specified in terms of the decision making process and operations allowed in the market.

Agents Population:

As has been presented previously, we have implemented a large population of heterogeneous agents (267.069 agents, see Table 4.1).

Among these strategic agents, 264.117 (98.89%) are never bankrupted during the whole process.

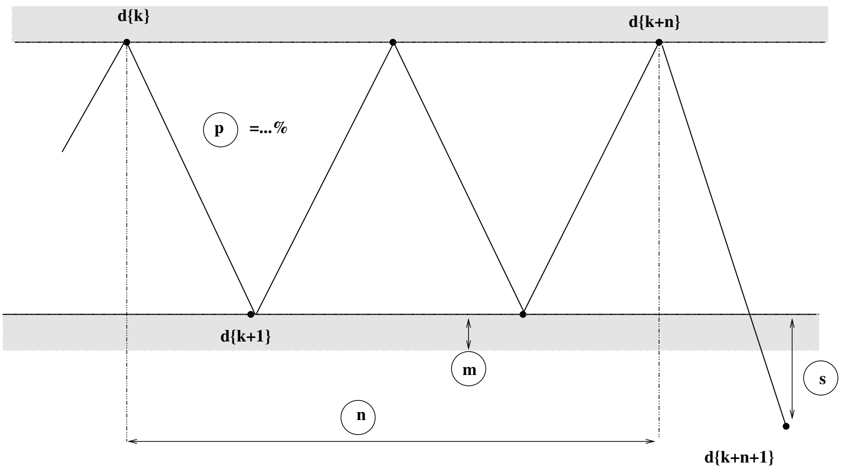

Fig. 4.1. Family Agent "Rectangle"

Table 4.1. Agents Population

Family Agent	Num. *Strategic Agents*
Periodical $\{n, m\}$	250.000
Indicator $\{n, m, p\}$	1.470
Rectangle $\{n, m, p, s\}$	7290
Triangle $\{n, m\}$	1.547
Variation $\{n, m, p\}$	2.000
Momentum $\{n, m\}$	220
Moving Average $\{n\}$	200
Weighted Moving Average $\{n\}$	200
RSI $\{n, p\}$	4.141
Buy & Hold	1
Total	267.069

Allowed Operations and Behavior:

Each agent is allowed to trade n trackers ($n \in R^+$), that is, one financial commodity replicating exactly one market index (like CAC40, Dow-Jones or Nikkei). If it has not decided to hold such commodities, the agent holds cash. Therefore, each agent is in one of these situations:

- it possesses a number of trackers > 0; in this situation we say that the agent is *"in the market"*. Its wealth fluctuates along with the market.
- it does not have any tracker or fraction of a tracker, all its wealth been converted into cash; the agent is *"out-of the market"* and its wealth is

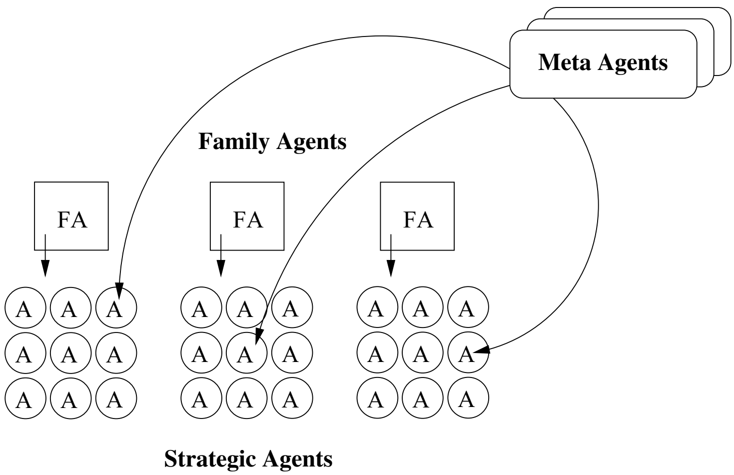

Fig. 4.2. Multi-Agent System design

stable. There is no risk-free asset paying a low interest rate available is our simulations.

At each time step, agents receive new information and have to decide if it is worth staying in the market or getting out: they follow systematically the signals given by their technical rules. For instance, a *Strategic Agent* designed as a *"moving average 5, 5"* analyzes at each iteration if the five days moving average of past prices has crossed the price process in the top-down direction, which correspond to a "sell-signal" (*resp.* bottom-up, "buy-signal"). In this situation, if the gap between the five days moving average and the price is greater than 5%, it will "sell" (resp. "buy"). If the gap is under 5%, it will keep its portfolio unchanged. Each *Strategic Agent* follows the same kind of behavior with various charts or technical rules. Nevertheless, one has to notice than one singular agent follows systematically a *"Buy & Hold"* strategy (B&H), that is, it enters the market at $time = 0$ (buys one tracker) and lets the situation remain unchanged until the end of the simulation. This agent is our "benchmark" agent in terms of risk and return and stands for a "passive investor".

Theoretically, no one can outperform this agent when considering the risk-adjusted performance in the long run, assuming the EMH holds. In other words, despite it is obvious than anyone can construct a portfolio or adopt a strategy that will outperform the B&H strategy, however, this assumes a higher risk level for the investor and, generally speaking, cannot be qualified as an outstanding behavior.

Each agent is endowed with the same amount of cash at the beginning of the simulations. If an agent looses all its endowment during an experiment, since borrowing is not allowed, it is withdrawn from the market.

In the simulations, agents are considered as *"price takers"*, that is, their behavior has no effect on the price of the asset they trade. This is a very commonly accepted hypothesis in finance, whereas it can be debated in MAS dealing with artificial stock markets. Trading is subjected to transaction costs at a 0.5% level.

	Real–Worl Universe *(RWU)*	**Artificial–World Universe** *(AWU)*
Subperiod 1 *in–sample selection*	*CAC 40* 01-1988 : 07-1996	*i.i.d Random–Walk*
Subperiod 2 *out–of sample test*	*CAC 40* 08-1996 : 04-2005	*i.i.d Random–Walk*

Fig. 4.3. Experimental design

The simulations are based on real daily data from the Euronext Paris Stock Exchange between 1988 and 2005. The traded tracker perfectly replicates CAC40 index. Agents have access only to past values of this tracker and the information they receive at each time-step is the price of the tracker corresponding to the current iteration (no agent is "cheating" and none behaves like knowing what the "future" will be).

4.2.3 Organization of the Simulations

Our experimental design is organized in two steps on two *"universes"* (see Fig. 4.3):

1. *"Universes"* are sets of data used to perform the simulations. Simulations are parallelized over the universes, each of them being useful for understanding what happens in the other.
   a) Real-World Universe: consists in historical CAC40 observations (see fig. 4.4), split into two subsamples, RWU_1 (01/1988-07/1996) and RWU_2 (08/1996-04/2005).
   b) Artificial Universe: consists in computer-generated data using an *iid* random-walk process[3]. This universe is also split into two subsamples AU_1 and AU_2 and includes the same number of observations as in $RWU_1$and RWU_2. These sets of data are intended to provide a universe in which it is actually impossible to outperform the market since it is artificially generated (assuming the random generator is good enough).
2. Over these universes, the simulations are organized as follows:
   a) Step 1, *"in-sample selection"*: is the selection of the best performing agents, compared to the benchmark agent. This test consists of 10 simulations based on random subsamples picked in RWU_1 and AU_1. These subsamples will be called *windows*. At the beginning of each simulation ($t = 0$), *Family Agents* create *Strategic Agents*. Then *Strategic Agents* begin to compete against the *Buy & Hold Agent*. Then *Family Agents* rank their respective sub-populations of *Strategic Agents*, comparing their performances with that of the benchmark agent. Once the 10 simulations have been processed, *Strategic Agents* that have out-performed the benchmark at least in 50% of the simulations are selected.
   Performance is always appreciated in terms of risk-return: a *Strategic Agent* outperforms the *Buy & Hold Agent* if and only if it achieves a more than proportional return considering the risk it has been exposed to during the simulation. "Risk" is calculated as the standard deviation of the agent's portfolio, "return" being the average rate of growth of its wealth.

[3] $p_t = p_{t-1} + \varepsilon_t$ with $\varepsilon_t \rightarrow N(\mu, \sigma)$, μ and σ being chosen to fit as closely as possible the corresponding parameters in RWU_1 and RWU_2

b) Step 2, *"out-of sample tests"*: consists of generalization of the first stage using the relevant second subsamples (*i.e. Strategic Agents* having out-performed the benchmark at least in 75% of the simulations.

4.2.4 How Do We Decide if a *Strategic Agent* Outperforms the *B&H* Agent?

Three performance indices are calculated providing information on risks and returns of the *Strategic-Agents*:

1. Return: $\overline{r_i}$ is the daily return earned by each agent i, for *windows* $t = [1, n]$, using the following formula:

$$\overline{r_i} = [Port_{i,t=n} - Port_{i,t=1}]^{1/n} \tag{4.1}$$

   In equation 4.1, $Port_{i,t}$ is the agent's i portfolio on date "t".
2. Risk: is calculated as the standard deviation of $r_{i,t}$ on each corresponding *window.*
3. Synthetic Index: combines the preceding indices and provides an aggregated measure for the absolute performance of one specific Agent i:

$$SI_i = r_i / \sigma_i \tag{4.2}$$

   One can notice the Synthetic Index reported in equation 4.2 is very similar to a Sharpe Index.

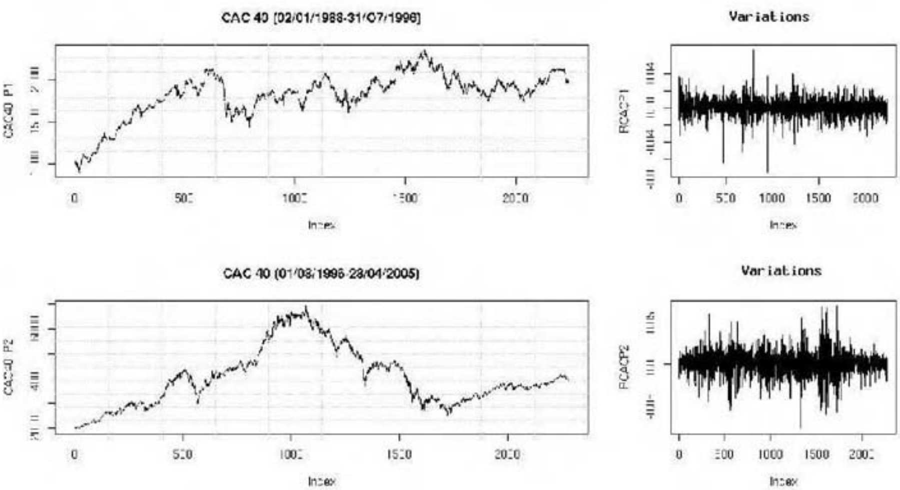

Fig. 4.4. RWU_1 and RWU_2 data – level / variations –

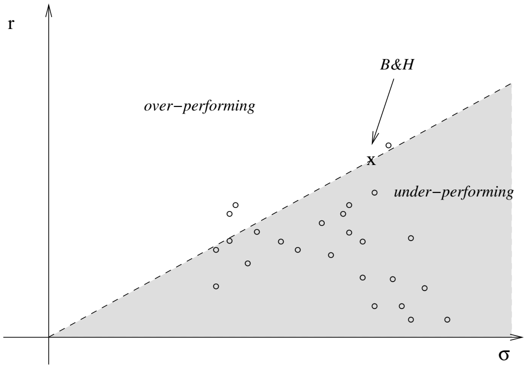

Fig. 4.5. Outperforming Strategic Agents in the Risk-Return Space

This set of indices is systematically evaluated for pairs of agents on each *window*. These pairs of agents are always a combination of one *Strategic Agent* and the *B&H* Agent. This procedure allows us to place *Strategic Agents* in a risk-return space for subsamples of observations. Assuming we have 10 *windows*, we thus will have to consider 10 risk-return spaces. In these spaces, outperforming *Strategic Agents* are placed in a part of the plan above the line crossing the origin and reaching one point representing the performance of the *B&H* Agent (see Fig. 4.5).

4.2.5 The Data Snooping Issue

Although the process presented in 4.2.3 might appear to be very harsh, it is clearly not sufficient to "prove", if at all it is possible, that any persistent, abnormal over-efficiency of some specific agents really occurs. Since we investigate the performance of a very large set of agents, we must consider the *data-snooping* problem.

To give an illustration of "data-snooping", let's consider the following example (derivated from Jensen and Cohen [4]): suppose you would like to hire someone having *abilities* to predict the next movements in a particular stock exchange. Obviously, the person to be hired would have to perform this task better than merely taking chances. To select a good candidate, you propose the following test: *"predict the next 14 fluctuations of the stock exchange in the following terms:* 0 *if the market closes up,* 1 *if it closes down"*. In other words, each candidate would have to propose a 14 characters-long string like 00101110010101. Suppose now you decide to hire someone providing at least a 75% rate of correct predictions (at least 11 good answers). The probability for someone to succeed here only by chance is very weak:

$$\sum_{i=11}^{14} C_{14}^{i} 0.5^{14} \simeq 0.02869$$

In other words, someone without any skill to predict these fluctuations would roughly have a 97% chance to fail. Suppose now 10 candidates face this challenge, none of them having any particular ability to predict the stock exchange, then the probability that *"at least one of them would succeed"* is sufficiently large to warrant careful examination of the successful candidate:

$$1 - (0.9713)^{10} \simeq 0.2525$$

Basically, if you increase the number of applicants to a certain point, you will probably hire someone passing the test[4] but nothing proves that this person has performed better than merely taking chances. This problem, known as the "data-snooping" bias, has been recognized very early in financial research, where data-mining has a long tradition [5][5]. One way to mitigate this issue is to apply a procedure called *Bootstrap Reality Check* (BRC) proposed by White (2000). In this research, BRC is intended to decide whether or not the selected agents, at the end of our experimental procedure, have positively out-performed the benchmark or not, that is, if they have out-performed the market exploiting weak-form inefficiencies or if this result must be attributed to chance.

Bootstrap Really Check (BRC) Procedure

BRC consists of testing the following null hypothesis: *"H_0: the best Strategic Agent does not outperform the B&H agent.*

Let's note θ_k one specific performance index for the k-th *Strategic-Agent* in a set of M agents,

$$H_0 : \max_{k=1\ldots M} E(\theta_k) \leq 0 \tag{4.3}$$

This performance is calculated over n subsamples (n=200) taken from RWU_2 or AU_2.
In this research, θ_k is:

$$\overline{\theta_k} = 1/n \sum_{T=1}^{n} (SI_{T,i} - SI_{T,B\&H}) \tag{4.4}$$

[4] with 100 candidates, the probability than none of them succeed is around 5%

[5] *"Let us [...] assume that we have access to a large computer and a body of security price data. Now, if we begin to test various mechanical trading rules with enough variants, we will eventually find one or more which would have yield profits [...] superior to a buy and hold policy [...] We cannot be certain that [...] results did not arise from mere chance"*

In equation 4.4, T is a specific *window*. In other words, we focus on an average over performance, appreciated with the *Synthetic Al Index* (see 4.2.4), over all the n considered *windows* for the selected agents. The best agent will therefore have an estimated performance as follows:

$$\overline{V} = \max_{i=1\ldots M} (\sqrt{M} \times \overline{\theta_i}) \tag{4.5}$$

We then generate B bootstrapped series using a process described by Politis and Romano [9]. Over those bootstrap series, we estimate again the whole set of performance indices θ_i. To distinguish these indices from those coming from the initial set of data, we note them $\theta_{b,i}$ (b being the *b-th* bootstrap series). The *p-value* for the Null is therefore:

$$p = \sum_{b=1}^{B} \frac{Z_b}{B} \tag{4.6}$$

with

$$Z_b = \left\{ \begin{array}{l} 1 \text{ if } \overline{V_b^*} > \overline{V} \\ 0 \text{ else} \end{array} \right| \tag{4.7}$$

4.3 Results

The simulations have been conducted over two different universes, as explained previously (see 4.2.3). Our results are presented successively for each universe.

4.3.1 Do Strategic Agents Behave Well in the Artificial Universe?

By construction, the Artificial Universe does not "hide" any useful information on date t allowing to predict what will probably happen on date $t+1$. Thus, these data perfectly replicate the behavior of an efficient market index. No *Strategic Agent* should be able, in this specific, virtual context to pass the filters we have programmed. Only chance could explain such an improbable success. Table 4.2 presents the best agents after each simulation step.

After Step 2, *none* of the Strategic Agents could be selected with the required 75% success rate. Table 4.2 shows the number of *Strategic Agents* outperforming the *B&H* agent in at-least 50% of the windows.

This leads us to consider the following explanation for this series of simulations:

- Our simulation process is sufficiently harsh and proves its efficiency in selecting good candidates: when no structure is hidden in a time series, no agent can outperform a basic *B&H* behavior.
- Our Artificial World does not reflect properly the real-world data (mainly because we have designed it as an *i.i.d.* process), and more complex dynamics in the Artificial World might have given a different result (ARCH process as example).

4.3.2 Does *RWU* Exhibit Inefficiencies?

The whole set of results concerning the two-step selection process exposed in 4.2.3 is reported in Table 4.3.

In-sample Selection: RWU_1.

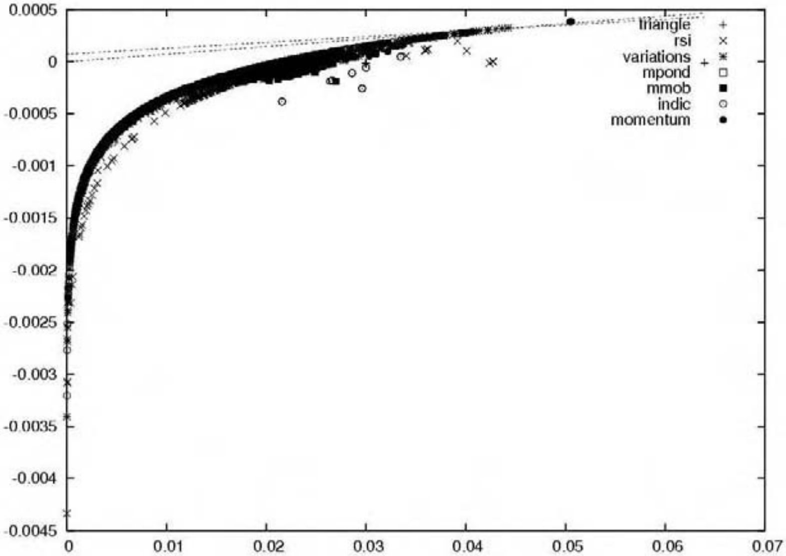

Fig. 4.6. *Strategic Agents* in the risk-return space for one window of RWU_1

Table 4.2. Agents "outperforming" the Artificial Index

Family Agent	Num. *after Step 1*	Num. *after Step 2* *50% selection rate*
Periodical	13368	3
Indicator	13	–
Rectangle	286	2
Triangle	20	–
Variation	28	1
Momentum	13	–
Moving Average	14	–
Weighted Moving Average	3	–
RSI	76	2
Total	13.821	8

Here, many *Strategic Agents* (6.057) outperform the benchmark agent in more than 50% of the cases. Each *Family Agent* has at least one of its children selected at the end of this simulation step. The major part comes from the *Periodical Family Agent* which is *"as expected"* but not significant since this family does not rely on classical technical signals "revealing" a "pattern". Fig. 4.6 shows, for one window in RWU_1, *Strategic Agents* from various *Families* in the risk-return space. The characteristic concave shape of the plot can be explained by the weigh of transaction costs that penalize the most active agents.

The selection rate over the initial population is between 0.34% and 11.5%, which is low, but "as expected". One has to keep in mind that our procedure involves a very large number of agents; it is perfectly normal that some of them seem to perform well at this initial stage. Fig. 4.7 shows the behavior of some interesting agents in terms of level of portfolio.

Out-of Sample Tests: RWU_2 and BRC

After the second selection process, only 19 *Strategic Agents* have out-performed the benchmark agent. They come from just two *Family Agents*: *Rectangle Trader* and *Variation Trader*. Some of them have outperformed the benchmark agent in each of the 200 simulations.

Clearly, the proportion of "good candidates" at the end of this out-of-sample test is very low. This is not surprising since modern stock markets are obviously not inefficient.

Fig. 4.8 shows the behavior of two very interesting *Strategic Agents*, *variation 7, 10, 3* and *variation 7, 10, 13*. The first one obtains a 100% score over 200 random windows in RWU_2 while the second one obtains a 76% score. Lines show the portfolio of these agents and the *B&H*'s for the specific window (01/2003-04/2005).

The next step in the analysis is to verify if this result can provide a kind of basis to reject the weak-form EMH. Thus, we have applied carefully White's *Reality Check* over 500 bootstrap series to control potential spurious results. The procedure leads to consider again the whole set of 6.057 *Strategic Agents* passing the first selection step.

Although the simulations seem to be very harsh in terms of selectivity for "good candidates", we cannot reject the null hypothesis: *"The best Strategic Agents cannot out-perform the Buy & Hold Agent"* at ordinary p-values (*p-value=28.2%*). Therefore, we cannot reject the initial weak-form EMH and cannot report evident market inefficiencies for data *with basic Strategic Agents* using simple trading rules. This result seems, therefore, to be a strong support for the weak-form efficiency of the French Market.

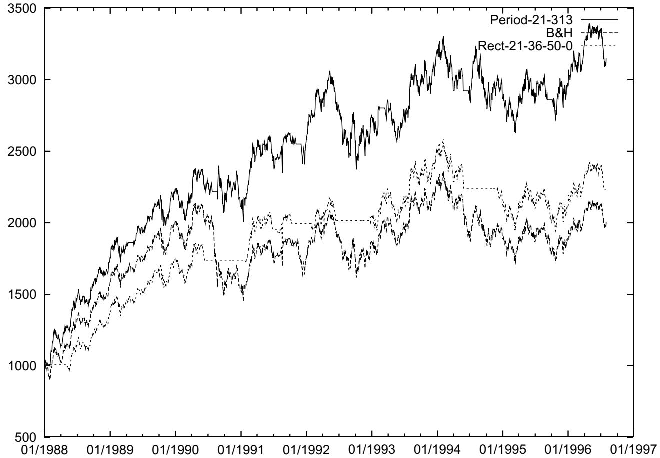

Fig. 4.7. Examples of "good *Strategic Agents*" on RWU_1

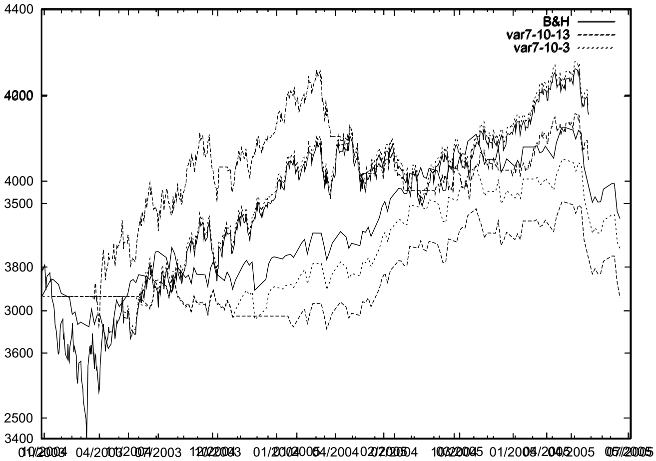

Fig. 4.8. Portfolios of two *Strategic Agents* and *B&H Agent*

4.4 Concluding Remarks

In this research, we show that technical traders cannot outperform a simple *Buy and Hold* Agent on Paris Euronext Stock Exchange. This result is derived from the following observations: our MAS can select some (apparently) very robust agents, producing very good risk-return scores after a harsh filtering process.

Table 4.3. Simulations: Real-World Universe

	Stage 1		*Stage 2*	
Family Agent	Number	*% / Initial Population*	Number	*% Remaining Population*
Periodical	5.484	2.19%	–	–
Indicator	5	0.34%	–	–
Rectangle	367	5.03%	12	3.27%
Triangle	74	4.78%	–	–
Variation	28	1.40%	7	25%
Momentum	16	7.27%	–	–
Moving Average	20	10%	–	–
W. Moving Average	23	11.50%	–	–
RSI T.	40	0.96%	–	–
Total	6.057	2.268%	19	0.314%

Nevertheless, these agents do not prove clearly their ability to obtain this performance in exploiting some kind of inefficiencies. Said differently, once a *bootstrap reality check* procedure has been performed, we cannot provide evidence that their performance is not due to mere chance.

In this research, we focus on the simplest level of the implemented MAS, that is, *Family Agents* and *Strategic Agents*. This first step was necessary to investigate the weak-form EMH with a broad-spectrum design. Although we cannot provide here evidence of market inefficiencies, these results suggest that more complex agents, behaving like real-world technical traders, combining various indicators to shape their strategies, might obtain a very different result. This last part of our work, based on *Meta Agents*, has still to be perfected to capture, if at all possible, some anomalies in financial data. Nevertheless, this is a necessary next step if one wants to invalidate the usual objection coming from many chartists or technical traders when quantitative analysis refute their so-called ability to outperform the market: their "knowledge" is often presented as a combination of complex receipts, which make a scientific verification difficult. MAS and Artificial Intelligence may here be very useful to design strict and robust tests.

References

[1] Brock, W., Lakonishock, J., and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance.*

[2] Detry, P. and Gregoire, P. (2001). Other evidences of the predictive power of technical analysis: the moving-average rules on european indices. In *Proceedings of the Lugano European Financial Management Association Conference.*

[3] Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance.*
[4] Jensen, D. and Cohen, P. (2000). Multiple comparisons in induction algorithms. *Machine Learning.*
[5] Jensen, M. (1967). Random walks, reality or myth - comment. *Financial Analyst Journal.*
[6] Jensen, M. and Benington, G. (1969). Random walks and technical theories: Some additional evidence. *Journal of Finance.*
[7] Malkiel, B. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives.*
[8] Park, C. and Irwin, S. (2004). The profitability of technical analysis. *AgMAS project Research Report 2004-4.*
[9] Politis, D. and Romano, J. (1994). The stationary bootstrap. *Journal of the American Sta- tistical Association.*
[10] Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review.*
[11] Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives.*
[12] White, H. (2000). A reality check for data snooping. *Econometrica.*
[13] Wooldridge, M. (2002). *An Introduction to MultiAgent Systems.* John Wiley and sons.