Reverse Engineering Financial Markets with Majority and Minority Games Using Genetic Algorithms

J. Wiesinger · D. Sornette · J. Satinover

Abstract   Using virtual stock markets with artificial interacting software investors, aka agent-based models, we present a method to reverse engineer real-world financial time series. We model financial markets as made of a large number of interacting boundedly rational agents. By optimizing the similarity between the actual data and that generated by the reconstructed virtual stock market, we obtain parameters and strategies, which reveal some of the inner workings of the target stock market. We validate our approach by out-of-sample predictions of directional moves of the Nasdaq Composite Index.

Keywords   Reverse-engineering · Financial markets · Agent-based models · Genetic algorithms · Forecast · Trading strategies · Market regimes

Abbreviations

EMH	Efficient Market Hypothesis
ABM	Agent-based model
GA	Genetic algorithm
SGA	Simple Genetic Algorithm
3PG	Third Party Game
MG	Minority Game
GCMG	"Grand Canonical" Minority Game
GCMjG	"Grand Canonical" Majority Game
delGCMG	Delayed "Grand Canonical" Minority Game

J. Wiesinger (✉) · D. Sornette · J. Satinover
ETH Zurich, Zurich, Switzerland
e-mail: judith.wiesinger@gmail.com

delGCMjG   Delayed "Grand Canonical" Majority Game
MixG   Mixed Game
ISD   Initial strategy distribution

"What I cannot create, I cannot understand": On physicist Richard Feynman's blackboard at time of death in 1988; as quoted in The Universe in a Nutshell by Stephen Hawking.

1 Introduction

The prediction of financial markets has long been the object of keen interest among both financial professionals and academics. The widely—if not universally—accepted Efficient Market Hypothesis (EMH) (Fama 1970, 1991) provides a powerful argument that markets are inherently unpredictable, in particular on the basis of prior price data: Because all information about the future is incorporated into the current price (for all practical purposes immediately), price changes must follow a random walk (Malkiel 2003). There is considerable evidence however that prices do not perfectly follow a random walk and that some price inefficiency is present, varying over time, perhaps enough at times to be exploitable (Dahlquist and Bauer 1998). However, recent assessments of the performance of hedge-funds (Barras et al. 2008) and of mutual funds (Fama and French 2009) cast doubt on the reality of the gains resulting from the practical implementation of these inefficiencies, if they exist. As illustrated in the approaches of Barras et al. (2008) and Fama and French (2009), deviations from the EMH are searched in the form of anomalous performance, beyond what can be explained by risk premia associated with exposures to a few dominating risk factors.

The near-absence of predictability in financial markets, or more precisely of risk-adjusted arbitrage opportunities, is truly remarkable. A rich academic literature has clarified the zen-like nature of the EMH in the sense that, the more intelligent are the investors and the harder are their efforts to gather information to make the best possible investment decisions, the fewer trading opportunities there are, and the more efficient is the market. The fact that markets are close to efficient can thus be understood as a macroscopic organization that result from the collective actions of the active investors. Borrowing from the jargon of complex system theory, market efficiency is an emergent phenomenon. Emergence, the existence of qualitatively new properties exhibited by collections of interacting individuals, is often taken to be the defining characteristic of complex adaptive systems.

Reciprocally, we ask here how the observation of the large scale behavior of a macroscopic system can (i) uncover the internal properties of a system and the organization among its constituents and (ii) be used for its prediction. Following Richard Feynman, we argue that, in order to really understand a system, we need to be able to strip things down, then rebuild them in order to play with the reconstructed simplified system and analyze variants, from which understanding can emerge.

We address this question of "reverse engineering" in the context of one-dimensional financial (market) time-series. The challenge consists in building a virtual stock market with artificial interacting software investors. The method presumes that real-world discrete market price changes may be in principle modeled as the aggregated output of a large number of interacting boundedly rational agents. These agents have limited

knowledge of the detailed properties of the markets they participate in and create, have access to a finite set of strategies to take only a small number of actions at each time-step and have restricted adaptation abilities. Given the time series data, our method of reverse engineering determines what set of agents, with which parameters and strategies, optimizes (in the sense of various robust metrics) the similarity between the actual data and that generated by an ensemble of virtual stock markets peopled by software investors.

We provide a validation step by testing the performance of the reverse engineered artificial market in predicting out-of-sample directional moves of the real-world time series. Using only some of the simplest strategies and agents, the p value for the statistical significance of the prediction of the directional moves for more than 600 trading days of the Nasdaq Composite Index is smaller than 0.02. The results are robust with changes of the styles of agents' strategies and for different market regimes.

Our work uses the extensive literature on agent-based models that has been developing at least since the 1960s [see LeBaron (2000) and references therein]. In agent-based models (ABMs), a system is modeled as a collection of autonomous decision-making entities, called agents. Repetitive competitive interactions among agents generate complex behavioral patterns. Due to the evolutionary switching among strategies, ABMs are highly nonlinear. The aggregation of simple interactions at the micro level may generate sophisticated structures at the macro level which provide valuable information about the dynamics of the real-world system which the ABM emulates. The main benefits of ABMs are that they (i) capture emergent phenomena; (ii) provide a natural description of a system; (iii) are flexible. ABMs have already been successfully applied in real-world problems, such as, flow simulation, organizational simulation, diffusion simulation and market simulation (Bonabeau 2002). In this article we focus on financial market simulation.

Hommes (2002, 2006) shows that ABMs can explain the main statistical regularities observed in financial time series—their so-called "stylized facts"—such as excess volatility and volatility clustering, high trading volume, temporary bubbles and trend following, sudden crashes and mean reversion, and fat tails in the distribution of returns. Toy models such as the Minority Game (MG), described in detail in Challet and Zhang (1997), capture key features of one generic market mechanism (competition for a scarce resource). The basic interaction between agents and public information is described in Challet et al. (2001); Marsili (2001). Details of the ABMs we employ will be introduced as we describe the implementation of our reverse engineering process. In brief, we concentrate on the so-called MG and its key variants and on the so-called $-game and related majority games.

A major thrust of the literature of ABMs dealing with finance is aimed at developing artificial stock markets and then analyzing the conditions which yield the stylized facts of real markets. Changes of parameters or of the model proper affect the collective behavior of the model and thus provide potential insight into the underlying structure of the real-world market. We take this one step further and focus on reverse engineering specific financial markets with the help of ABMs. Reverse engineering means that we are trying to find a generating process of a real financial time series based on the time series itself. We provide a first validation step, not by quantifying how well

the reconstructed synthetic market explains stylized facts but, rather by testing simple predictability.

In Jefferies et al. (2000); Johnson et al. (2001), the authors developed a first reverse engineering approach, using a "Grand Canonical" Minority Game (GCMG), whose detailed description is found in Johnson et al. (2000). The GCMG is an extension to the basic MG, in which the total number of actively participating agents fluctuates. The authors did not report results using real financial time series, but a time series generated by a known ensemble of such agents they pretend to know nothing about this ensemble apart from its output. Hence, they called it a "black box" ensemble. They then began with an ensemble of agents with randomized parameters [so-called Third Party Games (3PGs)] and, by iteration, "evolve" in parameter space this ensemble of 3PGs until its output matches that of the known, black box, ensemble. Here, matching meant maximizing the cross-correlation of the black box and 3PG time series. One may then open the black box to determine how well this procedure has approximated the structure of the unknown black box. When the evolutionary process is successful, this can be applied to a real world series. In the sequel, we will follow this general procedure and treat heuristically the resulting 3PG ensemble as a model of the truly unknown real world market structure of traders.

The main challenge in this procedure is finding an adequately optimized set of parameters for the 3PG, as parameter space is large and grows extremely rapidly with every increasing level of sophistication. Furthermore, the landscape of the solution space is extremely rugged, reflecting the underlying degree of frustration among competing agents in the model (and presumably, in the market being modeled). For this search, we use a genetic algorithm (GA), which is a methodology that adopts evolution used in nature to optimize the adaptation of life to the environment (Goldberg 1989; Holland 1992).

Repsilber (2002) has applied reverse engineering based on a GA to a biological problem. But, at this stage, it is not obvious what are the lessons learned that could be useful in our context, since Repsilber (2002) used statistical data that are arguably much less noisy that the financial time series investigated here. Another difference is that the rate of evolution and adaptation in financial time series is very fast so as to introduce an addition question, namely that of the stationarity of the time series.

Another use of GAs is found in the work of Arifovic (1996); Chen et al. (2008); Lettau (1997); Palmer et al. (1994), who equip agents with learning behavior for acting more profitably. In contrast, we chose the agents' learning process to be fixed via a simple adaptation rule using a virtual counter tracking the performance for each strategy because we want to map predefined ABMs (described in Sect. 2.3) to the financial data. We apply a GA for finding an ensemble of agents and their strategies best able to reproduce the time series that we hope to predict, referred to as the external time series throughout the article. In Andersen and Sornette (2005), a prototype is developed which identified a new mechanism for short-term predictability in ABMs. Including individual learning would amount to generalize our search to any possible ABM. Our procedure is thus rigid in the sense that we search in a reduced set of games.

In order to test the validity of this approach, i.e., to test how well the generating process of the time series can be captured by the reverse engineered ABM 3PG,

we analyze the predictions we obtain from the identified 3PG when it must predict out-of-sample real financial data.

2 Model/Methodology

In the following, the reverse engineering method, its input, the applied types of agents, the GA and the validation process are described in more detail. Let us mention in addition that the program is written in C++ and uses the standard C++ library boost.

2.1 General Set-Up of the Reverse Engineering Method

Figure 1 illustrates the whole process from the input to the prediction which will be explained stepwise in the following. In a nutshell, given a financial time series over some time interval and for a fixed ABM, using a GA (specified by its structure and parameters governing its search), we select a set of "best" ABMs i.e., their output best matches the financial time series. By "best" match we mean a minimization of "distances" between the financial time and the 3PG series, based on correlations (Lamper et al. 2001) and different standard norms. The results are found robust with respect to the choice of these norms.

2.2 The Nasdaq Composite Index as the Input of the Reverse Engineering Process

As input to our model, we use daily adjusted closing prices of the Nasdaq Composite Index. We study the Nasdaq because its focus on technology and growth stocks make
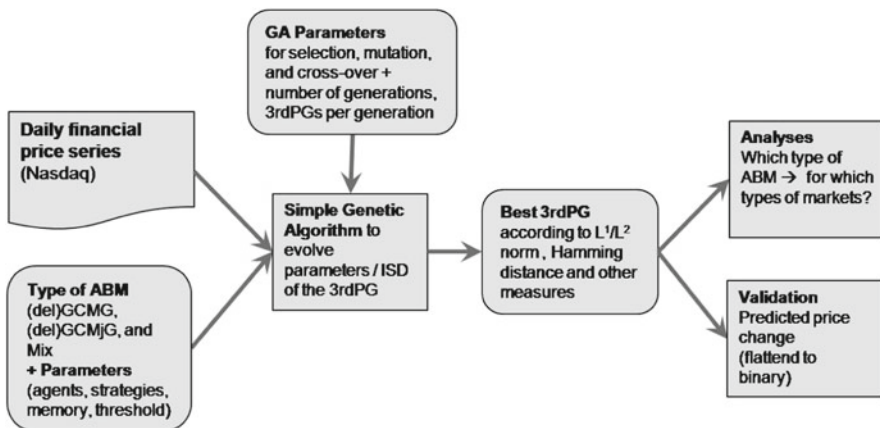
Fig. 1 Demonstrates a process overview of how a time series of daily adjusted closing prices is fed to a GA along with different parameters determining (i) the type of ABM for the 3PG which is used for generating a similar time series to the real one (during the in-sample period) and (ii) the convergence behavior of the GA for the search of the best—most similar—3PG according to measurements like L^1, L^2 and the Hamming distance. This result is then analyzed with respect to the types of ABMs present in which types of markets. The generating process is validated by the accuracy of one-step predictions

it prone to speculation and arbitrage, and thus to our diagnostic of interesting effects. Trading a composite index is nowadays straightforward via diverse derivative financial instruments, such as, synthetic assets, ETF (exchange traded funds) and futures.

We assess the performance of the reverse engineering approach from its ability to predict the signs of out-of-sample returns on the same Nasdaq time series. Results are obtained for 606 predictions. We present both aggregated metrics as well as results sorted according to different market regimes (upward trend, downward trend and no-trend), and compare with standard benchmarks i.e., with buy-and-hold (winning in upward trends), sell-and-hold (winning in downward trends), and random strategies. By distinguishing the three market regimes, we can infer from the performance of different type of ABMs which population of investors were dominant. For instance, it is intuitive and we confirm that trend-following strategies are dominant during upward trending markets. More surprising is the evidence we find for contrarian (or minority-type) strategies also performing well during such market phases, as we describe below.

The size of our statistical tests over 606 predictions constitutes a significant improvement with respect to prior effort of Andersen and Sornette (2005), which dealt with only a few tens of predictions. We were able to improve on this previous work using more efficient coding and the access to more computer resources available at ETH Zurich through the Brutus super-cluster. While much larger, our sample size remains limited by the high computational processing costs associated with the search of the GA exploring a large parameter space.

2.3 Description of the Different Types of ABMs

While we use different ABMs to be described shortly, the following properties are common to all of them. For a given ABM with N agents, each agent has to repeatedly choose among buying, selling or staying out of the market, according to their strategies, whereby the agents make their decision in parallel. They base their decision on (i) the previous performance of their strategies indicated by the virtual point counters, (ii) their threshold—is it profitable to trade with their strategies?—and (iii) their memory of prior returns—in its binary representation (up/down)—of the external time series.

The following types of ABMs are used, which differ in the incentives provided to the agents.

1. *"Grand Canonical" Minority Game (GCMG)* In the GCMG, an agent is rewarded for being in the minority (Johnson et al. 2000), whereby the extension of the classical MG consists therein that an agent has the possibility not to trade and hence, allowing for a fluctuating number of agents invested in the stock market.
2. *"Grand Canonical" Majority Game (GCMjG)* In the GCMjG, an agent is rewarded for being in the majority instead of in the minority (Marsili 2001).
3. *Delayed "Grand Canonical" Majority Game (delGCMjG)* In the delGCMjG, an agent is rewarded similarly to an agent in the GCMjG but for the fact that the return following the decision is delayed by one time step, in order to reflect the more realistic market property that returns are accrued after some time following

an investment decision. The grand canonical version is derived from the so-called $-game introduced by Andersen and Sornette (2003).

4. *Delayed "Grand Canonical" Minority Game (delGCMG)* This game is the analog of the delGCMjG, except for the minority payoff, whereby each agent is rewarded according to how the return at the next time step is compared with her decision taken at the previous time step. In other words, the delGCMG is a delayed GCMG.

5. *Mixed Game (MixG)* In the version of the MixG used here, we consider a mix of agents, with 50% of the agents obeying the rules of the GCMG and the other 50% obeying the rules of the GCMjG.

2.4 Description of the Genetic Algorithm

The 3PG which best reproduces the external time series provides the solution to our reverse engineering problem. This 3PG is determined from a search in the space of parameters of the ABM using a Simple Genetic Algorithm (SGA) as shown in Algorithm 1. First a population of 3PGs is initialized, whereby the number of agents, the number of strategies an agent obtains, the size of her memory, and her threshold are constant in the current version of the SGA. The only aspect in which the 3PGs differ is the initial strategy distribution (ISD) which is the crucial parameter set over which we optimize the fit to an external time series.

Algorithm 1 Simple Genetic Algorithm.

```
function SGA(extReturns, fitness(·))
    t ← 0                                          ▷ Time in nbr of generations
    p ← p_0                                        ▷ Initialization of 3PGs
    while (not terminal condition) do              ▷ Evolution
        t ← t+1
        fitness(p_{t-1}, extReturns)               ▷ Calculate the fitness
        p_t ← crossOver(selection(p_{t-1}))        ▷ Create a new generation
        mutation(p_t)                              ▷ Mutate randomly
    end while
    return bestOf(p_t)                             ▷ Return best 3PG
end function
```

For the first generation, the ISDs are initialized randomly. Then for every 3PG, its fitness—reflecting how well the time series generated by the 3PG matches the external time series—is determined. This value is computed via a fitness function using different metrics, such as the L^1 and L^2-norms, the Hamming distances (with binary and ternary coding) between the two time series.

3PGs are selected to produce offspring according to their fitness, with the fittest yielding more offspring. Each new generation of 3PGs is obtained as a mixture of the agents and of the strategies of the previous parent generation. Many generations evolve until a convergence criterion is reached, which leads us to finally identify a 3PG which best represents the external time series within the in-sample period. The search is performed ten times to obtain ten 3PGs. The differences between these ten solutions

provide a measure of the quality of the reverse engineering method. The ten 3PGs are also used to quantify the uncertainty in the next-day out-of-sample prediction.

Figure 2 shows the excess demand obtained from the aggregate decisions of all the agents of one selected best 3PG for a given time window of the Nasdaq Composite Index. The data point to the right of the vertical line is the next-step, out-of-sample, prediction, whereas all points to the left of it belong to the in-sample period during which the 3PG has been trained on the external time series and has been optimized in terms of its ISD. Every run of the GA results in one best—according to its fitness–3PG which then can be used to predict the next day return.

For each time window of the Nasdaq Composite Index, we obtain the best 3PG for each of the five types of ABMs defined in Subsection 2.3, using the above GA. This provides us with five different "lenses" to examine the Nasdaq Composite, that reveal its different characteristics.

In general, it can be argued whether our agents exhibit overconfidence and to which degree overconfidence has an impact on the market dynamics. The literature is of two minds concerning this issue. Many papers, such as, Bondt and Thaler (1994); Odean (1998a,b,c); Gervais and Odean (2001); Barberis and Thaler (2003) and Statman et al. (2008) lean towards overconfidence as being the explanation for excess volume. Moreover, Daniel et al. (2008) for example relate the pricing anomalies to the overestimation of the investors' private information. On the other hand, Salma and Ezzeddine (2008) have analyzed to what extend overconfidence correlates with the trading volume in
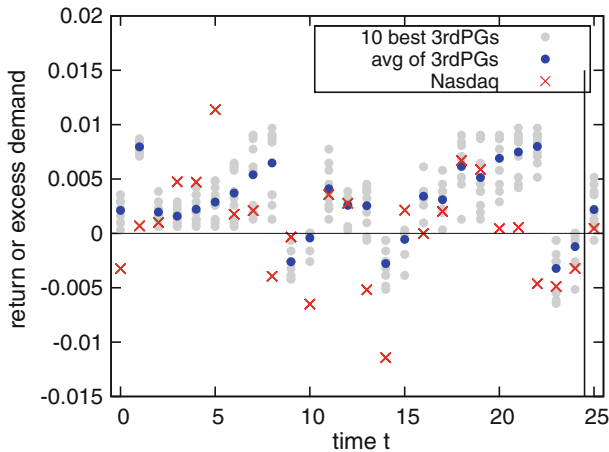
Fig. 2 This is an illustration of the procedure which is repeated for each day out of the analyzed period of 606 days and for each of the 5 types of ABMs described in Subsection 2.3. In the figure, the actual data (Nasdaq return illustrated as *red crosses*) and that generated by the reconstructed virtual stock market (the best 10 3PGs—in this sample consisting of GCMjG agents—illustrated as *gray dots* and their average in *blue*) are plotted, whereby the *vertical line* separates the in-sample period during which their similarity is optimized (here 25 days) from the out-of-sample period (one-step prediction). Since in this sample plot GCMjG agents have been used, it is no surprise that the agents' decision tends to overestimate the returns of Nasdaq. What actually is of relevance to us is the binary direction (= sign) of the agents' decision versus the returns of Nasdaq; whereby it is good (bad) if the sign of the agents' decision and the sign of the return of Nasdaq are the same (different). (Color figure online)

the Tunisian stock market and find little supporting evidence for this overconfidence hypothesis.

## 3 Validation by the Statistical Significance of the Success Rate of Next-Day Prediction

In order to test the predictive value of the reverse engineered 3PG for each of the five ABMs, we report the success rate, that is, the fraction of days out-of-sample for which the predicted and realized returns have the same sign.

In order to assess the statistical significance of the obtained success rates, we compare them to those of 1,000 *random strategies*, obtained by predicting with equal probability 1/2 the rise or decline of each next-day market price. Using random strategies has been shown to provide the most robust estimations of the statistical significance of strategies in the presence of biases and trends (Daniel et al. 2009). In Table 1, we report the p value of the 3PGs for each ABM, calculated as the fraction of random strategies that perform better.

Table 1 Success rates (average, minimum, and maximum) and their p values (stated in *parentheses*) for each type of ABM cumulated over (i) all days, (ii) the trending periods, and (iii) the non-trending periods

Agent type	(p val) avg	Min	Max
All periods			
GCMG	(0.01) 0.55	0.51	0.60
GCMjG	(0.00) 0.57	0.54	0.60
delGCMjG	(0.00) 0.57	0.54	0.59
delGCMG	(0.02) 0.54	0.51	0.57
MixG	(0.00) 0.56	0.53	0.58
Trending periods			
GCMG	(0.02) 0.57	0.54	0.63
GCMjG	(0.00) 0.66	0.63	0.68
delGCMjG	(0.00) 0.67	0.64	0.70
delGCMG	(0.01) 0.58	0.55	0.61
MixG	(0.00) 0.67	0.62	0.68
Non-trending periods			
GCMG	(0.07) 0.53	0.49	0.58
GCMjG	(0.13) 0.53	0.49	0.55
delGCMjG	(0.24) 0.52	0.49	0.54
delGCMG	(0.15) 0.52	0.50	0.55
MixG	(0.33) 0.51	0.48	0.54

The trending periods cover 202 days from 1985-10-25 until 1986-03-20, and from 1984-01-05 until 1984-05-29. The non-trending periods cover 404 days from 1976-05-10 until 1976-09-30, from 1984-04-05 until 1984-08-28, from 2002-06-20 until 2002-11-11, and from 2008-10-21 until 2009-03-17
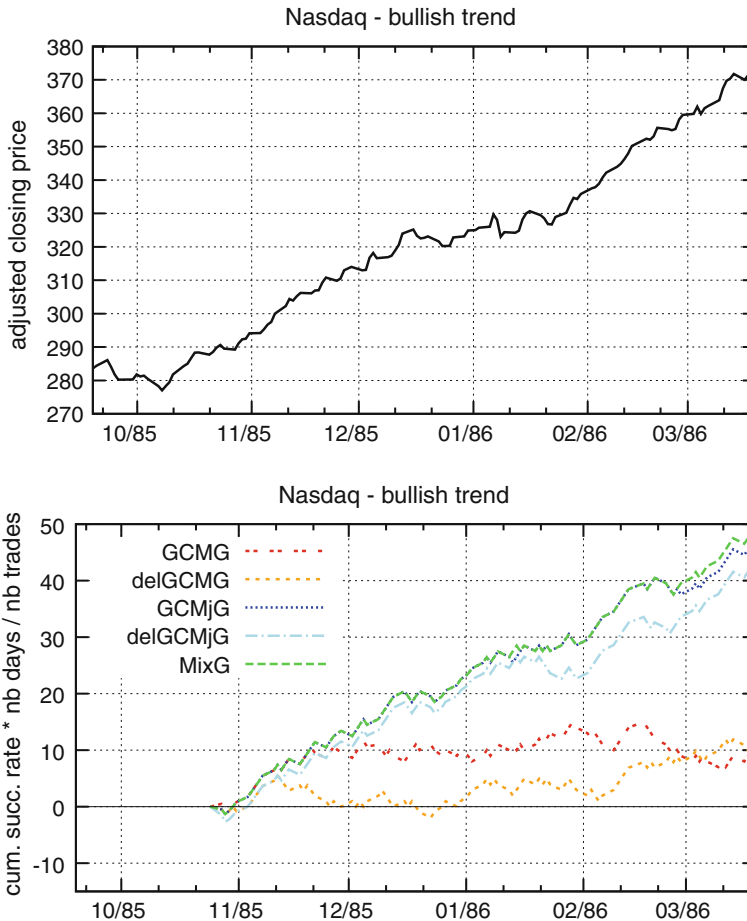
Fig. 3 *Upper panel*: price time series of Nasdaq (including the insample period of the first prediction and the 101 consecutive predicted trading days), it lower panel: cumulative success rates of the five ABMs, starting from zero and being increased (decreased) by one for every correct (incorrect) binary prediction of the future return of Nasdaq

Table 1 reports the success rates and their corresponding p-values for each type of ABM averaged over (i) all 606 days, (ii) for the trending[1] periods (202 days) and (iii) for the non-trending[2] periods (404 days). In the second column, the success rates are averaged over all parameter sets of the GA. The third and fourth column report the minimum and maximum success rates over the parameter sets of the GA.

Over all days independently of the presence or absence of trends, the success rates of the reverse engineered 3PG are superior to all random strategies ($p < 0.001$) for

---

[1] Bullish markets consisting of at least double the amount of days having a positive return than days having a negative return. In other words, on 2 out of 3 days the market goes up; vice-versa for bearish markets.

[2] Not trending markets are composed of an equal amount of days on which the market is going up as going down.
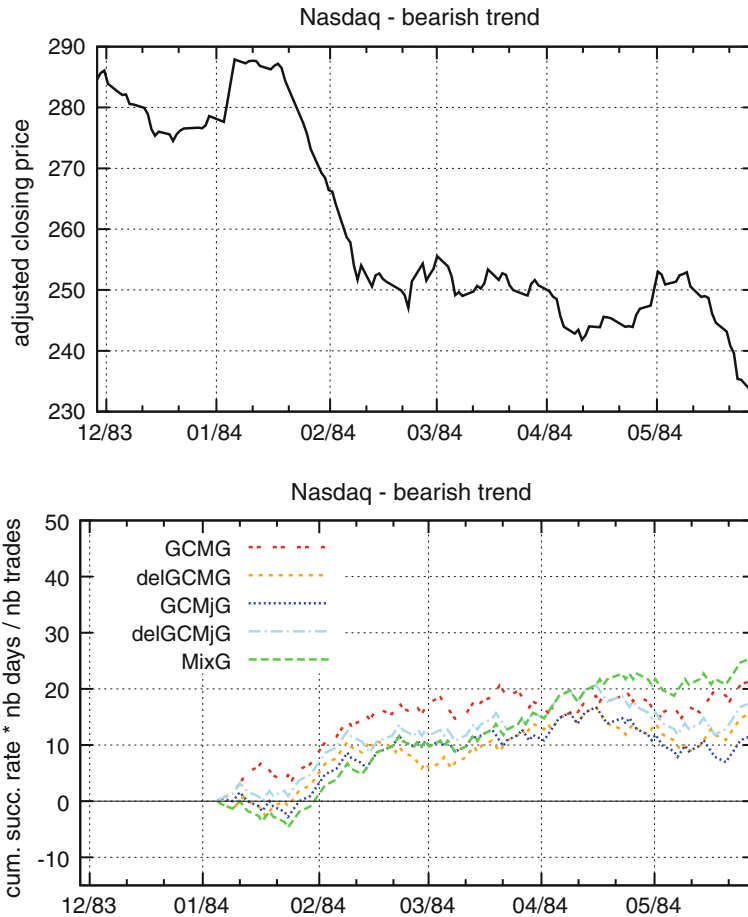
Fig. 4 Same as Fig. 3 for a bearish trend

the GCMjG, delGCMjG, and MixG. For the GCMG and delGCMG, the results are still very significant with p-values given respectively by 0.01 and 0.02.

Decomposing the 606 test days into trending and non-trending periods, we find that the success rates are very significant for the former periods (see Figs. 3, 4) and less so for the later periods (see Figs. 5, 6, 7, 8). The reverse engineering procedure is thus a good trend detection method.

While it is expected that the GCMjG, delGCMjG, and MixG would perform in trending period due to their majority incentive, it is a priori quite surprising that the GCMG and delGCMG also perform very significantly. We interpret this result as follows. First, the reverse engineering process applied with the GCMjG, delGCMjG, and MixG selects the trend-following strategies which, when used by a majority of agents, allow a good fit to the trend. Second, the fact that GCMG and delGCMG also perform well in trending periods implies that these trending periods are not just simple
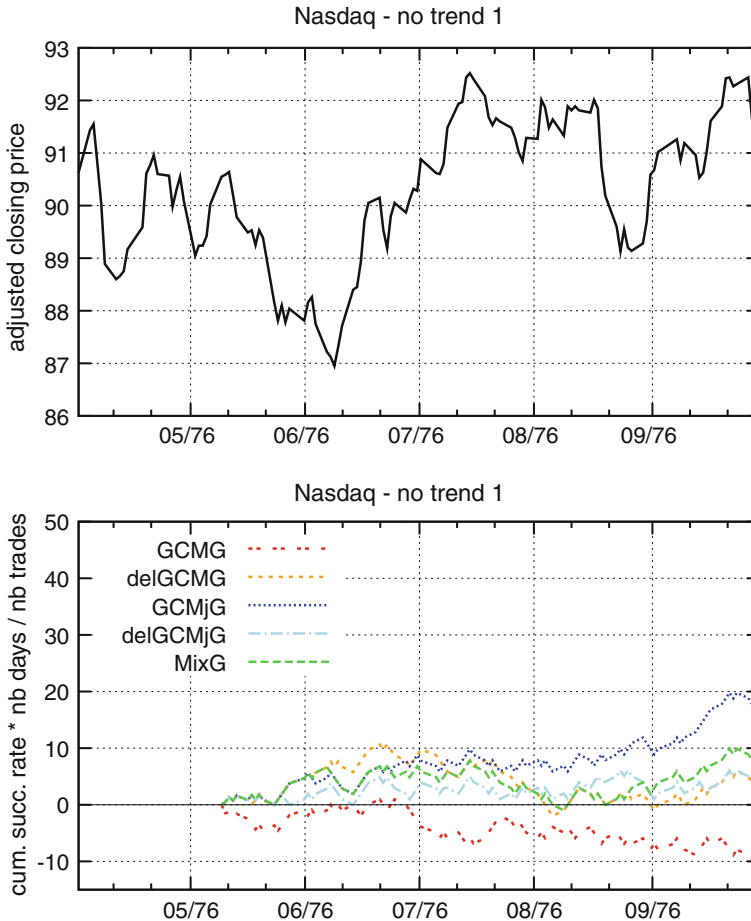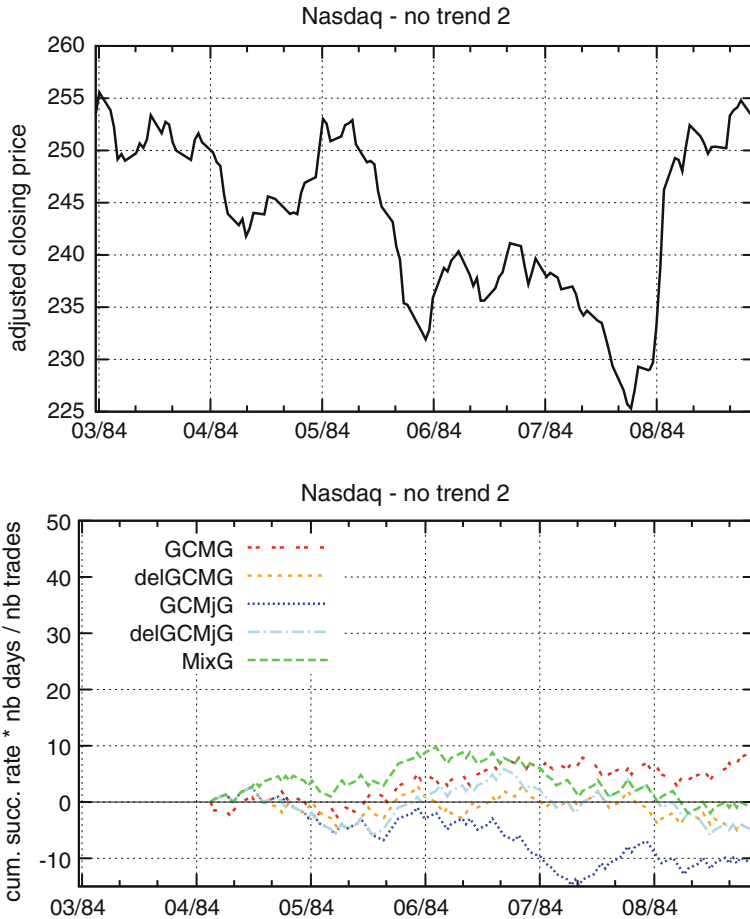
Fig. 5 Same as Fig. 3 for a first period of the market without trend

trends, but are decorated with cycles or alternating correction phases that the minority mechanism is able to pick up.

In contrast, all ABMs show a strong drop in performance in the non-trending periods, with the best performing game being the GCMG. This later result can be rationalized by the minority incentive of this game, which is known to lead to oscillation prices resulting from the frustration inherent to the minority payoff (Marsili 2001).

Figures 3, 4, 5, 6, 7, 8 show (i) on the upper plot the price time series of Nasdaq (including the insample period of the first prediction and the 101 consecutive predicted trading days) and (ii) on the lower plot the cumulative success rates of the five games, starting from zero and being increased (decreased) by one for every correct (incorrect) binary prediction of the future return of Nasdaq.

Figure 3 shows that GCMjG, delGCMjG and MixG are by far out-performing GCMG and delGCMG during the bullish trend. For the bearish trend shown in Fig. 4, the performances are quite similar among the five games. Figs. 5, 6, 7, 8 show that, in

Fig. 6 Same as Fig. 3 for a second period of the market without trend

absence of a clear trend, a given game can dominate or be dominated. This depends on structures that are more subtle than just the trend. It is in fact one of the goals of using ABMs to help characterize market regimes beyond the standard classification in terms of trends, by using strategies as "spectroscopy" of financial markets (**?**).

When analyzing the impact with agents themselves using one of our developed strategies (equivalent to the different ABMs used for prediction), we find the following: If one agent uses one of our strategies, she does not have a significant impact so that the computer estimation of her gain / loss is reliable, with our finding showing that she would be able to predict the binary move of the price significantly better than random. In contrast, if all agents in the market have the same behavior, their aggregate impact becomes enormous and this leads to various idiosyncratic behaviors, depending on the nature of the ABM: herding and bubbles for the GCMjG and anti-persistence for the GCMG.
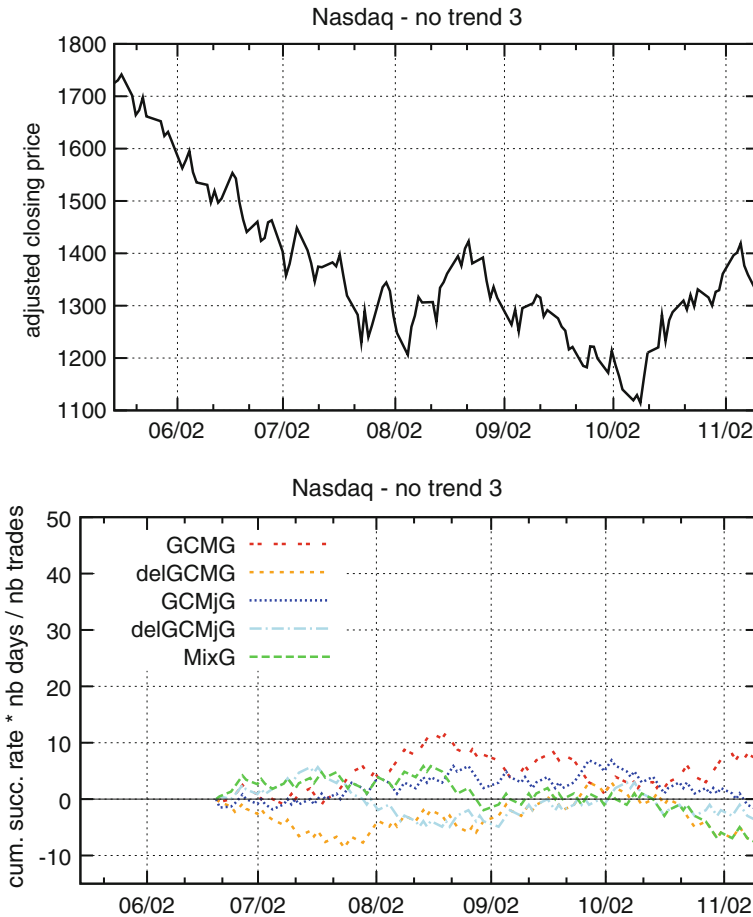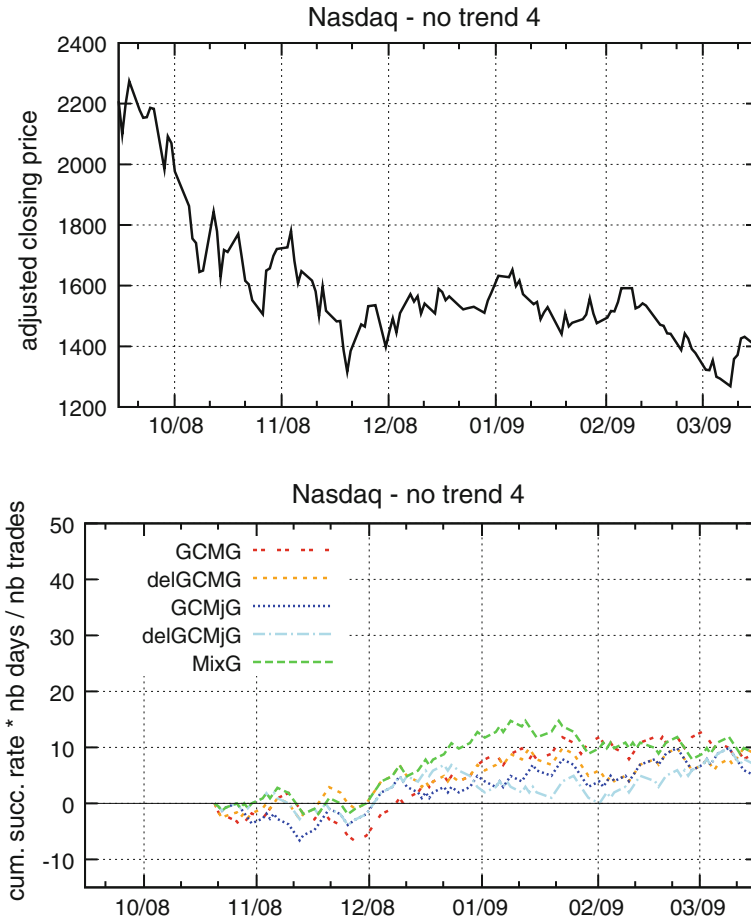
Fig. 7 Same as Fig. 3 for a third period of the market without trend

4 Conclusion

In conclusion, we have shown that reverse engineering a real financial time series with simple ABMs selected by using a genetic algorithm might be possible and provide novel insight in the properties of financial time series. Notwithstanding the simplicity (some would say "naivety") of the used ABMs, the aggregation of simple interactions at the micro level is sufficient to generate sophisticated structures at the macro level, which is probably the explanation for the good performance obtained in the validation step.

Finally, the method developed here is more generally applicable to the prediction of complex systems with an underlying multi-agent structure.

Nasdaq - no trend 4

Fig. 8 Same as Fig. 3 for a fourth period of the market without trend

Appendices

A Efficient Market Hypothesis

Traditionally in economics, all market participants are considered to be completely rational. Furthermore, markets are supposed to constantly be in equilibrium, which means that actual prices in financial markets reflect fundamental values. In other words, a stock price is supposed to be conform to everyone's expectations of future cash flows. Markets are efficient, which means that they digest all information infinitely fast, so it

becomes immediately and perfectly represented in prices. Prices would only change when their fundamental value changes, which could only be due to news arriving which change the expected return and the risk. Thus, it should be impossible to make any profit from trading.

This widely accepted hypothesis is called the EMH (Fama 1970, 1991) for which there are three different interpretations:

- The *weak-form efficiency* assumes that no excess returns can be earned by using investment strategies based on historical stock prices or other financial data, which implies that there are no statistically observable patterns.
- The *semistrong-form efficiency* states that stock prices adjust within an arbitrarily small amount of time to public information, so that no excess returns can be earned by trading conditioned to that information. This implies that it is impossible to earn any profit due to fundamental analysis. The only possibility to earn profit is due to private information.
- The *strong-form efficiency* claims that stock prices reflect all information, public and private, and no one can achieve excessive returns. The only way to outperform the market would be by luck, whereas prices would change only with the arrival of new information.

B P Value Test

The p value test is used in order to compare the performance of our strategies with the performance achieved by random strategies (randomly buying or selling). We run 1000 random strategies (as described in Sect. 3) and order their performances from the best performing strategy (maximum theoretically possible value = 1) to the worst performing strategy (minimum theoretically possible value = 0). E.g. the p value would be 0.01, obtained as 10/1,000 (rank of our strategy divided by the total number of random strategies in our statistical sample). The p value is thus the fraction of random strategies that perform better than our strategy. A small p value implies that our strategy has genuine skill, since it is unlikely that the obtained performance can be obtained by chance. Typical levels for rejecting the hypothesis that our strategy performance is just chance are $p = 0.05$ (corresponding to the 95% confidence level) and $p = 0.01$ (corresponding to the 99% confidence level).

C Third Party Game

The characteristic of a 3PG is that one feeds some time series and tries to find some parametrization of an ABM in order to be able to reproduce the fed time series. As input any time series can be used; a time series of an ABM whose parameters are pretended not to be known or a real financial time series. The types of agents we use are described in Sect. 2.3. whereby a relatively small number of agents per 3PG is sufficient for our simulations.

References

Andersen, J. V., & Sornette, D. (2003). The $-game. *The European Physical Journal B - Condensed Matter and Complex Systems 31*(1), 141–145. doi:10.1140/epjb/e2003-00017-7.

Andersen, J. V., & Sornette, D. (2005). A mechanism for pockets of predictability in complex adaptive systems. *EPL (Europhysics Letters) 70*(5), 697–703. doi:10.1209/epl/i2005-10033-3.

Arifovic, J. (1996). The behavior of the exchange rate in the genetic algorithm and experimental economies. *The Journal of Political Economy 104*(3), 510–541. doi:10.2307/2138862.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance (Chapter 18). *Handbook of the Economics of Finance 1*, 1053–1128. doi:10.1016/S1574-0102(03)01027-6.

Barras, L., Scaillet, O., & Wermers, R. R. (2008). False discoveries in mutual fund performance: measuring luck in estimated alphas. Social Science Research Network Working Paper Series. http://ssrn.com/abstract=869748.

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America 99 Suppl 3* (Suppl 3), 7280–7287. doi:10.1073/pnas.082080899.

Bondt, W. F., & Thaler, R. H. (1994). Financial decision-making in markets and firms: A behavioral perspective. National Bureau of Economic Research Working Paper Series, 4777+. http://www.nber.org/papers/w4777.

Challet, D., Chessa, A., Marsili, M., & Zhang, Y. C. (2001). From minority games to real markets. *Quantitative Finance 1*(1), 168–176. doi:10.1080/713665543.

Challet, D., & Zhang, Y. (Dec. 1997). Emergence of cooperation and organization in an evolutionary game. *Physica A: Statistical and Theoretical Physics 246*(3-4), 407–418. doi:10.1016/S0378-4371(97)00419-6.

Chen, F., Gou, C., Guo, X., & Gao, J. (2008). Prediction of stock markets by the evolutionary mix-game model. *Physica A: Statistical Mechanics and its Applications 387*(14), 3594–3604. doi:10.1016/j.physa.2008.02.023.

Dahlquist, J. R., & Bauer, R. J. (1998). *Technical Markets Indicators: Analysis & Performance*. New York, NY: John Wiley & Sons Inc. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471197211.html.

Daniel, G., Sornette, D., & Woehrmann, P. (2009). Look-ahead benchmark bias in portfolio performance evaluation. *The Journal of Portfolio Management 36*, 121–130. doi:10.3905/JPM.2009.36.1.121.

Daniel, K. D., Hirshleifer, D. A., & Subrahmanyam, A. (2008). Investor psychology and security market under- and over-reactions. Social Science Research Network Working Paper Series. http://ssrn.com/abstract=1288967.

Fama, E. F. (1970). Efficient Capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417. doi:10.2307/2325486.

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance 46*(5), 1575–1617. doi:10.2307/2328565.

Fama, E. F., & French, K. R. (2009). Luck versus skill in the cross section of mutual fund returns. Social Science Research Network Working Paper Series. http://ssrn.com/abstract=1356021.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *The Review of Financial Studies 14*(1), 1–27. http://www.jstor.org/stable/2696755.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning* (1st ed.). Boston, MA: Addison-Wesley Professional. http://www.worldcat.org/isbn/0201157675.

Holland, J. H. (1992). *Adaptation in natural and artificial systems*. Cambridge, MA: MIT Press. http://portal.acm.org/citation.cfm?id=129194.

Hommes, C. (2006). Heterogeneous agent models in economics and finance (Chapter 23) *Handbook of Computational Economics 2*, 1109–1186. doi:10.1016/S1574-0021(05)02023-X.

Hommes, C. H. (2002). Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proceedings of the National Academy of Sciences of the United States of America 99*(10), 7221–7228. doi:10.2307/3057845.

Jefferies, P., Hart, M. L., Hui, P. M., & Johnson, N. F. (2000). From market games to real-world markets. *The European Physical Journal B—Condensed Matter and Complex Systems 20*(4), 493–501. doi:10.1007/s100510170228.

Johnson, N. F., Hart, M., Hui, P. M., & Zheng, D. (2000). Trader dynamics in a model market. *Journal of Theoretical and Applied Finance 3*, 443–450. http://www.worldscinet.com/ijtaf/03/0303/S0219024900000358.html.

Johnson, N. F., Lamper, D., Jefferies, P., Hart, M. L., & Howison, S. (2001). Application of multi-agent games to the prediction of financial time-series. *Physica A Statistical Mechanics and its Applications 299*, 222–227. doi:10.1016/S0378-4371(01)00299-0.

Lamper, D., Howison, S., & Johnson, N. F. (2001). Predictability of large future changes in a competitive evolving population. *Physical Review Letters 88*(1), 017902+. doi:10.1103/PhysRevLett.88.017902.

LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control 24*(5-7), 679–702. doi:10.1016/S0165-1889(99)00022-6.

Lettau, M. (1997). Explaining the facts with adaptive agents: The case of mutual fund flows. *Journal of Economic Dynamics and Control 21*(7), 1117–1147. doi:10.1016/S0165-1889(97)00046-8.

Malkiel, B. G. (2003). *A Random Walk Down Wall Street: Completely Revised and Updated Edition*, completely revised and updated Edition. New York, NY: W.W. Norton & Co. http://www.worldcat.org/isbn/0393325350.

Marsili, M. (2001). Market mechanism and expectations in minority and majority games. *Physica A: Statistical Mechanics and its Applications 299*(1–2), 93–103. doi:10.1016/S0378-4371(01)00285-0.

Mu, G.-H., Zhou, W.-X., Chen, W., Sornette, D. (2011). Strategies used as spectroscopy of financial markets reveal new stylized facts. http://arxiv.org/abs/1104.3616.

Odean, T. (1998a). Are investors reluctant to realize their losses? Social Science Research Network Working Paper Series. http://ssrn.com/abstract=94142.

Odean, T. (1998b). Do investors trade too much? Social Science Research Network Working Paper Series. http://ssrn.com/abstract=94143.

Odean, T. (1998c). Volume, volatility, price, and profit when all traders are above average. Social Science Research Network Working Paper Series. http://ssrn.com/abstract=108028.

Palmer, R. G., Arthur, W. B., Holland, J. H., Lebaron, B., & Tayler, P. (1994). Artificial economic life: a simple model of a stockmarket. *Physica D Nonlinear Phenomena 75*, 264–274. doi:10.1016/0167-2789(94)90287-9.

Repsilber, D. (2002). Reverse engineering of regulatory networks: simulation studies on a genetic algorithm approach for ranking hypotheses. *Biosystems 66*(1–2), 31–41. doi:10.1016/S0303-2647(02)00019-9.

Salma, Z., & Ezzeddine, A. (2008). Overconfidence And trading volume: Evidence from an emergent market. *Annales Universitatis Apulensis Series Oeconomica 1*(10). http://ideas.repec.org/a/alu/journl/v1y2008i10p41.html.

Statman, M., Thorley, S., & Vorkink, K. (2008). Investor overconfidence and trading volume. Social Science Research Network Working Paper Series. http://ssrn.com/abstract=1097947.