Agent-Based Models Simulations for High Frequency Trading

D. Dezsi, E.Scarlat and I. Măries
Department of Economic Cybernetics, Bucharest University of Economics, Romania

Abstract
High frequency computer-based trading (HFT) represents a challenging topic nowadays, mainly due to the controversy it creates among investors on the financial market. The hereto paper compares two types of agent-based models, one with zero-intelligence traders and the other with intelligent traders in order to simulate the tick-by-tick high frequency trades on the stock market for the selected U.S. stocks. The simulations of the agent-based models are done with the help of Adaptive Modeler software application which uses the interaction of 2,000 heterogeneous agents to create a virtual stock market for the selected stock with the scope of forecasting the price. Within the intelligent agent-based model the population of agents is continuously adapting and evolving by using genetic programming in forming new agents by using the trading strategies of the best performing agents and replacing the worst performing agents in a process called breeding, while the zero-intelligence agent based model does not evolve, agents do not breed, and they trade in a random manner. After comparing the fitting of the two models with the real data, the results show that in almost all the cases the intelligent agent-based model performed better when compared to the zero-intelligence agent-based model, which could be interpreted as lower market efficiency, allowing for predictions of the stock market price, or even stock market manipulation. Also, the zero-intelligent agent-based model generates more trades and lower wealth for the population, compared to the intelligent agent-based model. The high-frequency data turns out to be very hard to simulate and analyse due to its particularities which differentiate them from daily data, as price changes are discrete, being multiples of the minimum price increment, the price changes not being independent.
Keywords high frequency trading, agent-based modeling, zero-intelligence traders, double auction, financial market.

1 Introduction

The changes of the stock market structure due to the technological improvement which led to the high speed computer-based trading have switched the market from an investor-focused mechanism to a trader-focused mechanism, where the investors' trust and concerns are ignored. According to a study conducted by Beddington in 2012 on the impact of the computer trading on the financial markets based on which the report entitled The Future of Computer Trading in

Financial Markets-An International Perspective [1] was released, although the investors are worried regarding the possible abuse of the market generated by the high-frequency trading (HFT) and claims on market manipulation using HFT techniques are reported by institutional investors all over the world, it seems like so far economic research has provided no direct evidence that HFT has increased market abuse, but the authors of the report also underlined that the research regarding the measurement of market abuse is at an early stage and incomplete, not being the main focus of the report. Furthermore, Brogaard (2010) [2] examines the impact of HFT on the U.S. equity markets and the results show that HFT activities are not detrimental to non-HFT activities and that HFT tends to improve market quality.

Despite the lack of research on this subject, and even though the abuse is present or not in the market, regulators and policy makers should take in consideration the perception of the investors which show worries in this regard, because this is what determines trading behavior, portfolio management and investment decisions. Thus regulators should increase their ability to detect abuse and obtain significant empirical proof on either to confirm or deny the manipulation related to HFT in order to restore market confidence.

Market abuse through HFT is studied by the economists due to its implications on market liquidity, volume, pricing efficiency and even social welfare. According to the studies conducted by Cumming et al. (2012) [3] on the possible risks of market abuse generated by HFT, the authors find that HFT lead to lower incidence of manipulation, as also another study conducted by Aitken et al. (2012) [4] reports that HFT improves market efficiency without harming the market integrity. The papers mentioned above analyze market abuse at the end of the trading day, while empirical studies for the HFT market abuse during the continuous trading period have not been conducted yet. These results are in contrast with the surveys regarding the perception of the investors over the degree of manipulation in the market generated by HFT, which were conducted on the stock markets all over the world, and which underline the considerable concerns showed by the large investors regarding the market abuse and the lack of action and detection of manipulation by the regulators.

The statistical modeling facts, models and challenges for high frequency financial data have been studied by Cont [5] who outlined the empirical characteristics of high frequency financial time series, providing an overview of stochastic models for the continuous-time dynamics of the limit order book described as queuing systems, pointing that the gap between microstructure models and stochastic models should be filled in by inputting theoretical issues from microstructure models to design new stochastic models with a better economic interpretation. Furthermore, Ponta et al. [6] propose a non-homogeneous normal compound

Poisson process for describing non-stationary returns for high-frequency financial time series for the Italian stock exchange, also testing if the model can reproduce some stylized facts of high-frequency financial time series. In order to model the high-frequency Foreign Exchange market, Aloud et al. [7] constructed an agent-based model which is able to reproduce the stylized facts of the trading activity on the Foreign Exchange market, by using zero-intelligence directional-change event trading strategy.

According to the results obtained by Li and Krause (2009) [8] by comparing the market structures with near-zero-intelligence traders with the use of agent-based model simulations, they have observed that the properties of returns arising in double auction markets are not very sensitive to the trading rules employed. Also, Sunder (2004) [9] underlines that artificial and computer intelligence are very important in understanding the distinction between individual behaviour and market outcomes, outlining that computer simulations helped to discover that allocative efficiency is largely independent of variations in individual behaviour under classical conditions, while Herbert Simon was convinced that “the possibility of building a mathematical theory of a system or of simulating that system does not depend on having an adequate microtheory of the natural laws that govern the system components. Such a microtheory might indeed be simply irrelevant” [9].

The aim of our research is to identify the main distinctions between the high-frequency financial data simulation results of two agent-based models, one with intelligent agents and the other with zero-intelligence agents. In order to achieve our research aim, we use the Adaptive Modeler software to simulate the two types of agent-based market models-the zero-intelligence agent-based model and the intelligent agent-based model-for price forecasting of real world market-traded securities such as stocks. Thus, heterogeneous agents trade a stock floated on the stock exchange market, placing orders depending on their budget constraints and trading rules, where the virtual market is simulated as a double auction market. For the intelligent agent-based model, the software uses evolutionary computing such as the Strongly Typed Genetic Programming [10] in order to create adaptive, evolving and self-learning market modeling and forecasting solutions, while for the zero-intelligence agent-based model the agents do not breed nor evolve, as they trade in a random manner. The results obtained from the simulations are compared revealing that the high frequency tick-by-tick data is best modeled by the intelligent agent-based model, although the high-frequency data turns out to be very hard to simulate and analyse due to its particularities which differentiate them from daily data, such as discrete price changes, which are multiples of the minimum price increment, thus the price changes are not independent.

To the best of our knowledge, none of the works in the literature have attempted to compare the results obtained from simulating zero-intelligence agent-based models with the results obtained from simulating intelligent agent-based models for high frequency financial data.

The organization of this paper is as follows. Section 2 provides an overview of the high frequency financial data and its statistical properties, Section 3 describes the datasets used in this study, Section 4 presents the agent-based models used in the simulations, while the results of the simulations that have been performed are presented in section 5, the paper ending with the conclusions and avenues for future work.

2 High Frequency Financial Data and Their Statistical Properties

The high frequency financial data or high frequency trading (HFT) represents a subset of algorithmic trading or a trading strategy where a large number of small size orders are sent into the market at high speed, therefore the securities are bought and sold by a computer algorithm and held for a very short period, usually seconds or milliseconds. The automated trading applies on the order-driven electronic platforms which aggregate all outstanding orders in an order book and market orders are executed in a mechanical manner, usually a double auction mechanism on the stock market. For a better comprehension, Table 1 illustrates the hierarchy of time scales for financial data, as pointed out by Cont in [5], while thousand of buying or selling orders are submitted in a 10-second interval.

The HFT is one of the most significant market structure developments in

Table 1 Financial data series time scales

Regime	Time scale	Issues
Ultra-high frequency (UHF)	$10^{-3} - 0.1$s	Microstructure
High Frequency (HF)	1-100s	Trade execution
Daily	$10^3 - 10^4$s	Trading strategies

recent years, as stated by the Securities and Exchange Commission (SEC), but is also controversial due to its divers impact on investors' perception, which are worried about the degree of manipulation in the market generated by HFT, according to surveys conducted on the stock markets. Despite the investors worries described above, research studies on the impact of HFT over the market show that HFT tends to improve market quality [2], while Vuorenmaa (2012) [11] reached the conclusion that the benefits of the HFT have the most weight and dominate the negative aspects of HFT.

Recently, EU Parliament voted in favour of MiFID II, a draft law for strengthening the regulation of HFT and other financial instruments. The new rules will apply starting with 2014 and aims to reduce speculation without harming the real economy, by introducing a synchronised clock for trading shares, bonds, commodities and other instruments across the EU so regulators can spot abuses more easily in a market where many exchanges and platforms trade the same shares, and by extending the minimum time in which high-frequency orders have to remain in the market from 3 milliseconds currently to at least 500 milliseconds. That way, purely speculative business with high-frequency transactions will be discouraged.

Recall that high-frequency and algorithmic trading impact on the financial market has raised worries from behalf of the regulators and the investors since May 6, 2010, when about USD 862 billion was erased from stock values in 20 minutes before share prices recovered from the plunge, while on August 1, 2012 a trading malfunction at Knight Capital Group which poured in the market unintended orders lead to high traded volume and high price volatility, which finally led to a USD 440 million trading loss for the company.

According to a study conducted by Beddington in 2012 [1] on the impact of the computer trading on the financial markets based on which the report entitled The Future of Computer Trading in Financial Markets-An International Perspective was released, although the investors are very worried regarding the possible abuse of the market generated by the HFT, so far economic research has provided no direct evidence that HFT has increased market abuse. Furthermore, according to the studies conducted by Cumming et al. (2012) [3] on the possible risks of market abuse generated by HFT, the authors find that HFT lead to lower incidence of manipulation, while also another study conducted by Aitken et al. (2012) [4] reports that HFT improves market efficiency without harming the market integrity. The papers mentioned above analyze market abuse at the end of the trading day, while empirical studies for the HFT market abuse during the continuous trading period have not been conducted yet.

On the other hand, the surveys conducted regarding the perception of the investors over the degree of manipulation in the market generated by HFT underline the considerable concerns showed by the large investors regarding the market abuse and the lack of action and detection of manipulation by the regulators. Also, HFT allows the program trader to “see” ahead of others the major incoming orders and to act benefitting from its speed advantage, in order to gain profits, as HFT has been shown to have a potential Sharpe ratio, which is a measure of reward per unit of risk, thousands of times higher than the classical buy-and-hold strategies [12]. A major disadvantage is that it can generate trade crashes, such as the one on May 6, 2010, and the one on August 1, 2012, thus eroding the trust of investors in the market.

As regards to statistical properties of the tick-by-tick high frequency financial data compared to daily frequency data, the main differences are the fact that price changes are discrete, being multiples of the minimum price increment as it can be observed in the histograms from Fig.1 below.

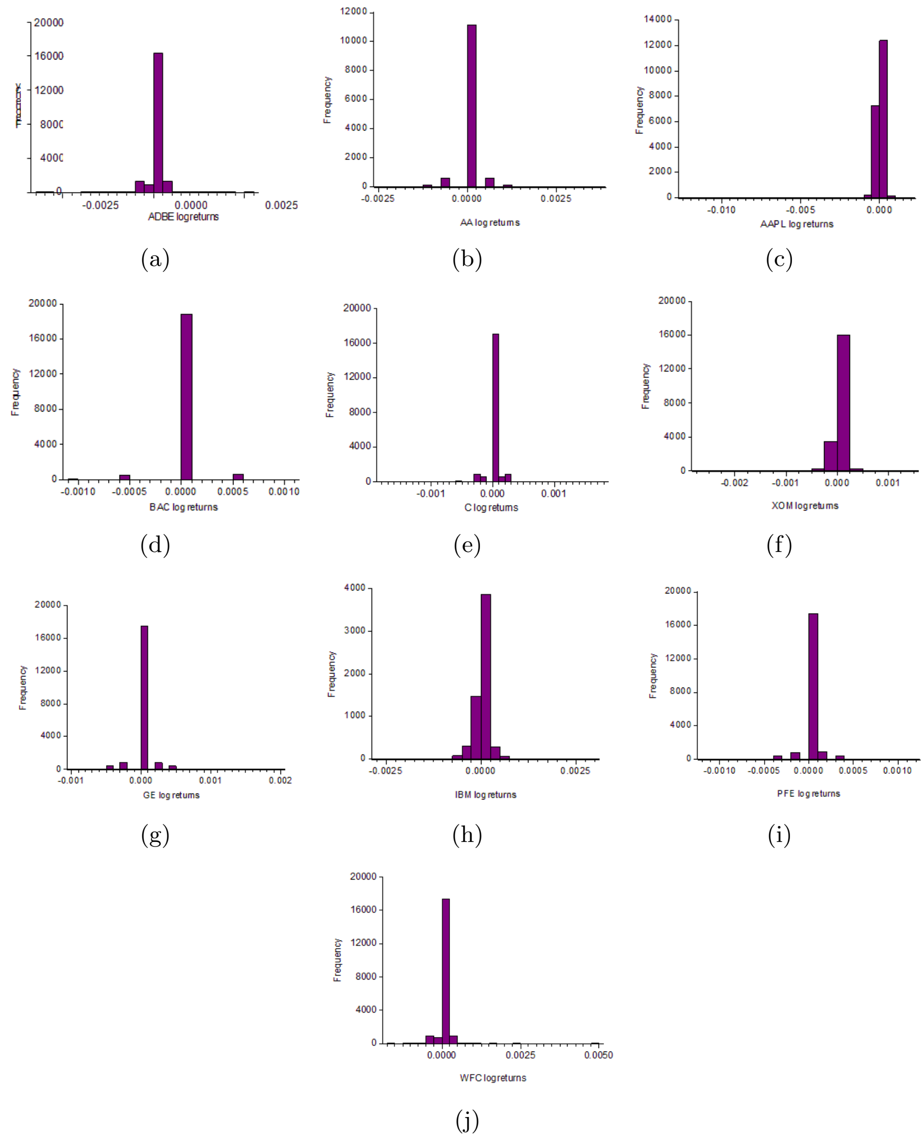

Fig.1 Histogram of log returns for ADBE(a), AA(b), AAPL(c), BAC(d), C(e), XOM(f), GE(g), IBM(h), PFE(i), WFC(j)

The autocorrelation function of high frequency price returns is significantly negative at the first lag and for higher lags it rapidly decreases to zero, meaning that price changes are not independent [13]. Another particularity is that trades occur at irregular intervals, they are random and endogenous, linked to the behaviour of the price and to previous trading history, and also seasonality effects is observed, as the trading activity is more intense at the market opening and closing [5].

3 Data

The high frequency financial data in this paper was retrieved from the website http://www.finam.ru/analysis/profile041CA00007/default.asp and contains tick-by-tick data from BATS stock exchange. The tick-by-tick data was retrieved from this website, for the trades that occurred on the BATS (Better Alternative Trading System) stock exchange during regular trading hours on December 3rd, 4th, 5th and 6th of 2012 limited to 20,000 quotes for each stock, and includes the date, time and closing price information for each stock. The unique dataset used in this paper, namely the tick-by-tick high-frequency financial dataset contains real trading data for the following stocks floated on the BATS, a computerized U.S. stock exchange market: Addobe Systems Inc. (symbol: ADBE), Alcoa Inc. (symbol: AA), Apple Inc. (symbol: AAPL), Bank of America (symbol: BAC), Citigroup Inc. (symbol: C) Exxon Mobil (symbol: XOM), General Electric (symbol: GE), IBM (symbol: IBM), Pfizer Inc. (symbol: PFE), Wells Fargo (symbol: WFC). The BATS stock exchange has been chosen for the data retrieval because it is one of the leading U.S. venues for HFT, covering 12-13% of all U.S. equity trading on a daily basis according to the stock exchange's website http://www.batstrading.com/about/. The distinction between the HFT and non-HFT was not defined in this paper, although taking in consideration the nature of the BATS stock market, we consider the trades to be mainly HFTs.

Table 2 Symbols and numbers of observations for the 10 analysed stocks

Stock name	Symbol	Sector	No. of obs. (bars)	Period of Time
Addobe Systems	ADBE	Technology	20,000	$3^{rd} - 6^{th}$ of December 2012
Alcoa Inc.	AA	Basic Materials	12,612	$3^{rd} - 6^{th}$ of December 2012
Apple Inc.	AAPL	Technology	19,999	$3^{rd} - 5^{th}$ of December 2012
Bank of America	BAC	Financial	20,000	$3^{rd} - 4^{th}$ of December 2012
Citigroup Inc.	C	Financial	20,000	$3^{rd} - 4^{th}$ of December 2012
Exxon Mobil	XOM	Basic Materials	20,000	$3^{rd} - 6^{th}$ of December 2012
General Electric	GE	Industrial Goods	20,000	$3^{rd} - 5^{th}$ of December 2012
IBM	IBM	Technology	6,144	$3^{rd} - 6^{th}$ of December 2012
Pfizer Inc.	PFE	Healthcare	20,000	$3^{rd} - 5^{th}$ of December 2012
Wells Fargo	WFC	Financial	20,0000	$3^{rd} - 5^{th}$ of December 2012

The time period of trades is expressed in milliseconds, the number of observations and period of time for each stock is summarized in Table 2, while the descriptive statistics for the tick-by-tick log returns are presented in Table 3, in which the excess kurtosis can be observed. Also, Fig.1 shows the histogram for the log returns of the analysed stocks, in order to show how the returns are distributed, being possible to see the discrete character of returns even after the logarithmic transformation.

Table 3 Descriptive statistics for the tick-by-tick log returns

Stock name	Symbol	Mean×10^6	Std.Dev.×10^4	Skewness	Kurtosis
Addobe Systems	ADBE	0.486	1.46	-0.360791	43.17925
Alcoa Inc.	AA	1.120	2.64	0.086025	19.32359
Apple Inc.	AAPL	-3.700	1.90	-16.53734	979.9840
Bank of America	BAC	-5.060	1.37	-0.334374	23.19635
Citigroup Inc.	C	-0.909	1.10	-0.400060	32.14271
Exxon Mobil	XOM	-0.443	1.10	-1.226893	41.91198
General Electric	GE	0.0472	1.24	0.436346	18.27284
IBM	IBM	-0.890	1.94	0.449297	32.84855
Pfizer Inc.	PFE	1.230	1.05	0.267775	16.65013
Wells Fargo	WFC	-0.653	1.21	3.666686	156.9951

4 The Agent-Based Models Specifications

An agent-based model is a computational model for simulating the actions and interactions of multiple agents in order to analyze the effects on a complex system as a whole, and represents a powerful tool in the understanding of markets and trading behavior. An agent-based model of a financial market consists of a population of agents (representing investors) and a price discovery and clearing mechanism (representing a virtual market). As regards to the financial markets, agent-based models can successfully replicate time series features like fat-tailed distributions and volatility clustering, on which standard financial models offer few explanations. Conventionally, financial markets have been studied using analytical mathematics based on a generalization of market participants and other simplifications and idealizations. However, the behavior of financial markets as observed in reality can't be fully described by such mathematical models. In reality, market prices are established by a large diversity of investors with different decision making methods and different investment goals. The complex dynamics of these heterogeneous investors and the resulting price formation process require a simulation model of multiple heterogeneous agents and a virtual market [14]. Research has shown that complex behavior can emerge from simulations of agents with relatively simple decision rules. Furthermore, commonly observed stylized facts of financial time series (price and order flow) have been reproduced by several agent-based market models [7,15-17].

Although agent-based research has been introduced over 30 years ago in economics, it is still considered a niche field because of its detailed and complex platforms. Another impeachment in using computational agent-based modeling is represented by the fact that it uses source codes which usually are not disseminated along with the paper, thus being hard to be evaluated, tested and developed for further research by the rest of the community, as underlined by Barr et all (2008) [12].

Furthermore, agent-based modeling can generate important facts when used for institutional design. At a large scale they are already used for computer simulation by government agencies, an important example being the traffic simulations, while as a small to intermediate size model it can be used by the government agencies as an advisory and prediction instrument.

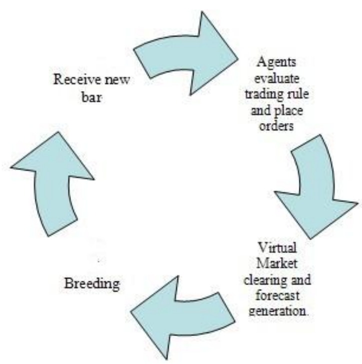

Fig.2 Agent-based model cycle

The agent-based models referred to in this paper are simulated in Adaptive Modeler software, which supports up to 2,000 agents and 20,000 observations for each simulation. To the best of our knowledge, none of the work in the literature has attempted to compare the simulation results generated from the simulations of zero-intelligence agent-based models with the results obtained from simulating intelligent agent-based models for high frequency financial data. Each model consists of a population of agents and a Virtual Market on which the agents trade the envisaged security. The agents are autonomous entities representing the traders of the stock market, each having their own wealth (cash and shares) and their own trading strategy.

The general cycle of the agent-based models used in this paper, which repeats at each of the 20,000 analyzed quotes, is illustrated in Fig.2 and Fig.3 in more details. It can be described as follows:

1. Import real stock market data: Real stock market data is imported for the simulated stock.

2. Initialization of the model: The cycle starts after the model initialization and after the parameters are set. At the initialization, the agents receive an

initial wealth of 100,000 and no shares. Each agent receives a trading rule which is called the genome, which is randomly created by taking in account the selected genes (which represent functions) using genetic programming described in [10].

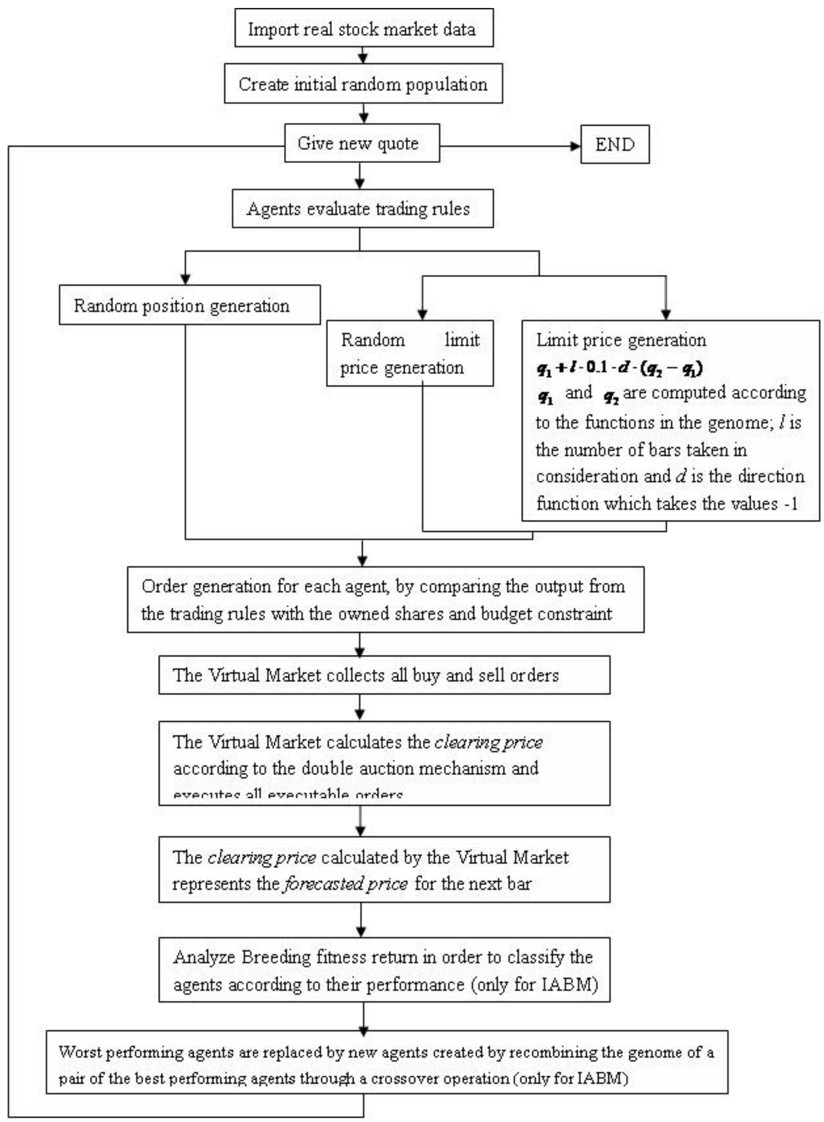

Fig.3 Agent-based model flowchart

There are no broker fees in these simulations, and no market maker. All the parameters for each of the two models and their values are described in Table 4.

Table 4 General settings of the models. Market and agents' parameters configuration in the simulations

Model	Parameter Type	Parameter Name	Parameter Value
ZI Agent-Based Model	Market Parameters	No. of trading periods (bars)	Max. 20,000; see Table 1.
		No. of agents	2,000
		Spread	20.005%
		Variable Broker fee	0%
	Agent Parameters	Wealth Distribution	Equal for all agents: 100,000
		Position Distribution	Equal for all agents% : initial position 0%
		Min. position unit	20%
		Min. initial genome depth	2
		Max. initial genome depth	5
		Genes	RndPos, RndLimit, Advice
		Breeding Cycle Frequency	1,000,000 bars (so that the agents never breed, don't evolve)
Intelligent Agent-Based Model	Market Parameters	No. of trading periods	Max. 20,000; see Table 1.
		No. of agents	2,000
		Spread	20.005%
		Variable Broker fee	0%
	Agent Parameters	Wealth Distribution	Equal for all agents: 100,000
		Position Distribution	Equal for all agents% : initial position 0
		Min. position unit	20%
		Max. genome size	1000
		Max. genome depth	20
		Min. initial genome depth	2
		Max. initial genome depth	5
		Genes	CurPos, LevUnit, Rmarket, Vmarket, Long, Short, Cash, Bar, InvPos, RndPos, close, bid, ask, average, min, max, volume, >, change, +, dir, isupbar, upbars, pos,
		Breeding Cycle Frequency	1 bar (so that the agents breed at each new bar, conditioned of the fact that they are of minimum breeding age)
		Minimum breeding age	80
		Initial selection: randomly select	100% of agents of minimum breeding age or older
		Parent selection	5% agents of initial selection will breed
		Mutation probability	10% per offspring

3. Receive new quote bar: A new quote bar from the imported real stock market data series is received by the Virtual Market.

4. Agents evaluate trading rules and place orders: Agents get access to historical prices and evaluate their trading rules according to the genomes allocated in the initialization process, resulting in a desired position as a percentage of wealth limited by the budget constraints, and a limit price. For the zero-intelligence agent-based model, a desired position and a limit price order are generated in a random manner. For the intelligent agent-based model, the position is also gen-

erated in a random manner, while the limit price is generated after a technical analysis has been performed, according to the genome structure which represent trading functions.

5. Virtual Market clearing and forecast generation: The Virtual Market determines the clearing price, executes all executable orders, and forecasts the price for the next bar, using a double auction mechanism.

6. Breeding (only for the Intelligent Agent-Based Model): During the breeding process, new agents are created from best performing agents in order to replace the worst performing agents, creating new genomes by recombining the parent genomes through a crossover operation. The breeding process repeats at each bar, with the condition that the agents must have a minimum breeding age of 80 bars, in order to be able to assess the agents performance.

7. The model waits for a new quote: If the model receives new quotes, it will repeat the process described at points 4-6. If there are no more quotes to be processed, the simulation ends.

In order to obtain random seed, the Adaptive Modeler software uses the Mersenne Twister algorithm [18] to generate pseudo random number sequences for the initial creation of trading rules or genomes and for the crossover and mutation operators of the breeding process. More information regarding the Adaptive Modeler software and how it works can be found at http://altreva.com/Adaptive_Modeler_Users_Guide.htm.

4.1 Zero-Intelligence Traders Model

The concept of Zero Intelligence (ZI) traders has been introduced by Gode and Sunder (1993) [19] in order to study the lower limit of rationality required to participate in a double auction market, which implies that agents have no strategy and they behave in a random manner subject to a budget constraint, so that the market mechanism can be observed. In case the ZI traders model generate good results, this means that the market mechanism used is incentive-compatible and most probably is not influenced by the trading strategies used by the investors. This represents a very important result for the market mechanism design, as a market mechanism which performs well despite the trader's irrationality is preferred over the mechanism that performs well only with perfectly rational traders, as underlined in Walia (2003) [20].

In Ladley's (2012) [21] point of view, any effects that are not observed in the ZI model cannot solely be due to the market mechanism and requires the interaction of the investors strategies, therefore it can be possible to separate the effects of the market mechanism and trader strategy and to determine the driving forces within markets. In this context, a model that assumes stock market traders have zero intelligence has been found to mimic the behavior of the London Stock Exchange very closely, the research being conducted by Farmer et al. (2004) [22] at the Santa Fe Institute, which suggest that the movement of the markets depend less on the strategic behavior of the traders and more on the design of the trading system.

A double auction is a trading mechanism in which buyers and sellers can enter bid or ask limit orders and accept asks or bids entered by other traders [23]. This trading mechanism was chosen to be used for the virtual market simulation in the Adaptive Modeler models because most of the stock markets are organized as double auctions, which is considered to be a continuous-time game of incomplete information [24]. Double auction stock markets converge to the equilibrium derived by assuming that the traders are profit-maximizing Bayesian, being an example of a microeconomic system, as described in Hurwicz (1986) [25] and Smith (1982) [26]. In the double auction markets, agents introduce bid or ask orders, each order consisting of a price and quantity. The bids and asks orders received are put in the order book and an attempt is made to match them. The price of the trades arranged must lie in the bid-ask spread (interval between bid price and ask price). Furthermore, the use of double auction trading mechanism generates stochastic waiting times between two trades, as stated by Scalas [27]. An example of the order book resulted from buy and sell orders collected by the Virtual Market can be viewed in the Fig.4, 5, 6 and 7, where blue bars represent buy orders, red bars represent sell orders and yellow bars represent buy and sell orders at the same price before the market clearing. Fig.4 and 6 illustrate the order book before the clearing, while Fig.5 and 7 illustrate the order book after clearing, so only the unexecuted buy (blue) and sell (red) orders remain visible in these two figures. The clearing price can be observed in the top of the figures.

In the hereto paper, the ZI traders model is designed within the context of the double auction mechanism, which sets the order placement, clearing price, the supply and demand. The model is implemented in the Adaptive Modeler software and will be used in this paper to simulate tick-by-tick high-frequency data. The Adaptive Modeler software in the ZI agent-based model uses only three types of genes or trading strategies, as mentioned in Table 4, namely RndPos, RndLim and Advice. The RndPos gene returns a position value ranging from -100% to 100% which is randomly generated from a uniform distribution, the RndLim gene returns a random limit price that is generated as follows: the last closing price is taken from either the Virtual Market or the Real Market which is chosen randomly, then the price is multiplied with a normally distributed random value with $\mu = 1$ and $\sigma = 3.5 \cdot \sigma_m$ where σ_m is the standard deviation of the log returns of the last 20 bars of the chosen market. The Advice gene combines the position generated by RndPos function and the limit price value generated by the RndLim function into a buy or sell order advice. The buy or sell order is introduced in the market after comparing the desired position with the agent's

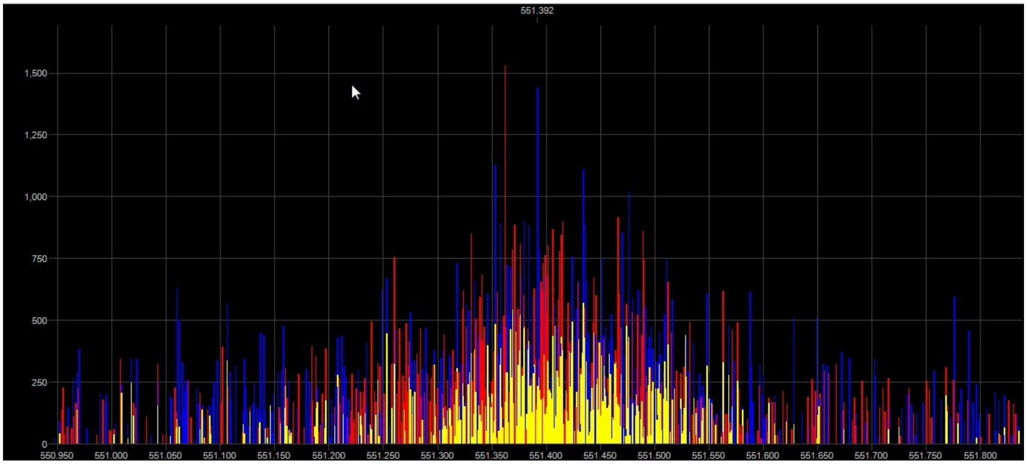

Fig.4 Order book example from the ZI-ABM simulation for AAPL stock. Blue bars represent buy volume, red bars represent sell volume, yellow bars represent buy and sell volume at the same price before clearing; clearing price can be seen above the chart

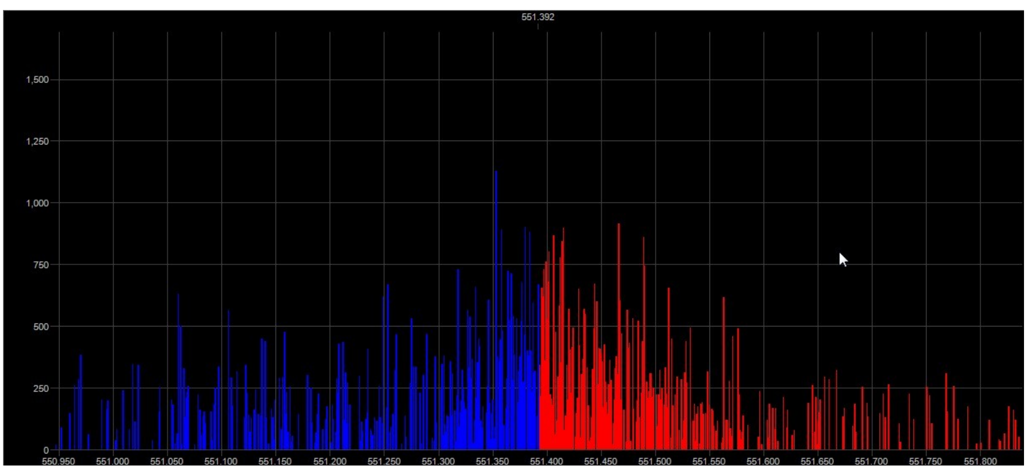

Fig.5 Order book example from the ZI-ABM simulation for AAPL stock. Blue bars represent buy volume, red bars represent sell volume after clearing; clearing price can be seen above the chart

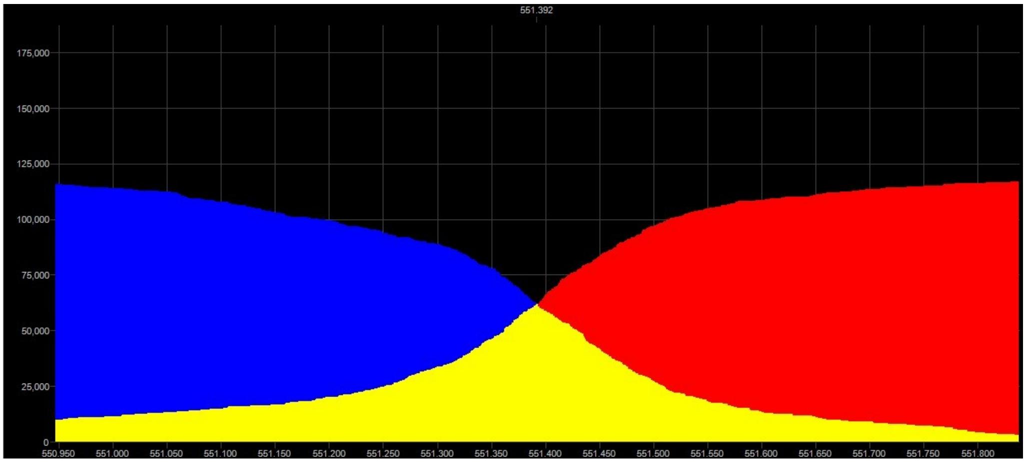

Fig.6 Order book example from the ZI-ABM simulation for AAPL stock. Blue bars represent cumulative buy volume, red bars represent cumulative sell volume, yellow bars represent buy and sell volume at the same price before clearing; clearing price can be seen above the chart

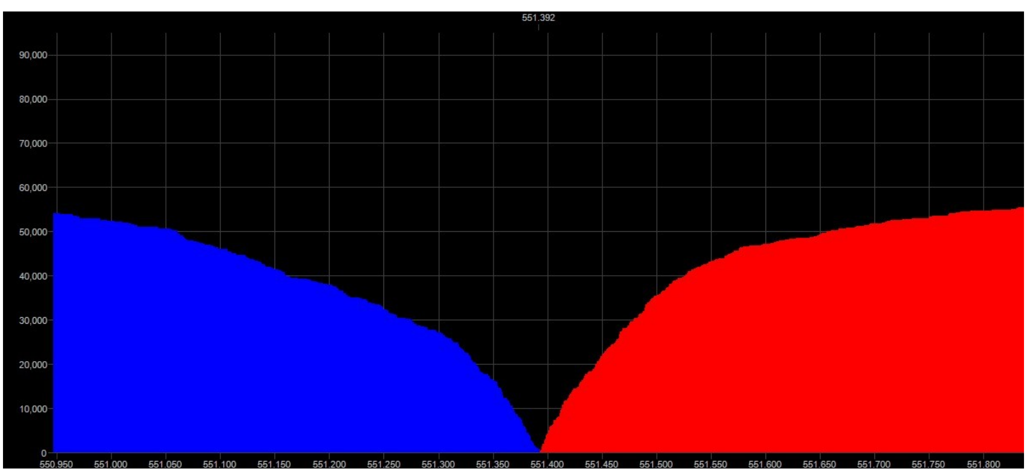

Fig.7 Order book example from the ZI-ABM simulation for AAPL stock. Blue bars represent cumulative buy volume, red bars represent cumulative sell volume after clearing; clearing price can be seen above the chart

current position and calculating the number of shares that need to be bought or sold, taking also in consideration the available cash. This order generation process is explained in detail in the paper Agent-Based Simulation of a Financial Market by Raberto (2001) [17], together with the ability of this model to exhibit the stylized facts of financial time series, such as fat tails and volatility clustering, using simple trading rules, budget restriction of the agents, order limit prices, and the creation and matching of demand and supply curves. The zero-intelligence agent-based model always skips the breeding step by setting the breeding cycle frequency parameter at a very high value (1,000,000) as mentioned in Table 4, therefore agents are not replaced with better ones, the population does not evolve, nor adapt to the market conditions, agents trading in a random manner.

4.2 The Intelligent Agent-Based Model

The intelligent agent-based model cycle starts by receiving a new quote bar, so that agents can place a new order or remain inactive according to their trading strategy. After all agents have evaluated their trading strategy, the Virtual Market determines the clearing price, executes all executable orders and releases the price forecast for the next bar. After that, breeding of new agents and replacement by evolutionary operations such as crossover and mutation can take place, a process which repeats itself for each bar.

The trading rules of the model uses historical price data as input, either from the Virtual Market either from the Real Market, and return an advice consisting of a desired position, as a percentage of wealth, and an order limit price for buying or selling the security. The trading rules are implemented by genetic programming technology explained bellow. Through evolution the trading rules are set to use the input data and functions (trading strategies) that have the most predictive value.

The agents' trading rules development is implemented in the software by using the Strongly Typed Genetic Programming (STGP) approach, and use the input data and functions that have the most predictive value in order for the agents with poor performance to be replaced by new agents whose trading rules are created by recombining and mutating the trading rules of the agents with good performance. The STGP was introduced by Montana (2002) [10], with the scope of improving the genetic programming technique by introducing data types constraints for all the procedures, functions, variables and constants, thus decreasing the search time and improving the generalization performance of the solution found. Therefore, the genomes (programs) represent the agents' trading rules and they contain genes (functions), thus agents trade the security on the Virtual Market based on their analysis of historical quotes. In order to obtain random sequences for the initial creation of trading rules or genomes and for the crossover and mutation operators of the breeding process, the Adaptive Modeler software uses the Mersenne Twister algorithm [18] to generate pseudo random number sequences. More information regarding the Adaptive Modeler software and how it works can be found at http://altreva.com/Adaptive_Modeler_Users_Guide.htm. Also, all the genes as functions are described in the Guide, along with the types of the arguments used.

During the breeding process which is specific only to the intelligent agent-based model, new offspring agents are created from some of the best performing agents to replace some of the worst performing agents. In order to achieve this, at every bar, agents with the highest Breeding Fitness Return are selected as parents, and the genomes (trading rules) of pairs of these parents are then recombined through genetic crossover to create new genomes that are given to new offspring agents. These new agents replace agents with the lowest Replacement Fitness Return. The fitness functions are a measurement of the agents investment return over a certain period, therefore the Breeding fitness return is computed as a short term trailing return measure of the wealth over the last 80 analyzed quotes and represents the selection criterion for breeding (best agents), while the Replacement fitness return is computed as the average return per bar and represents the selection criterion for replacement (worst agents).

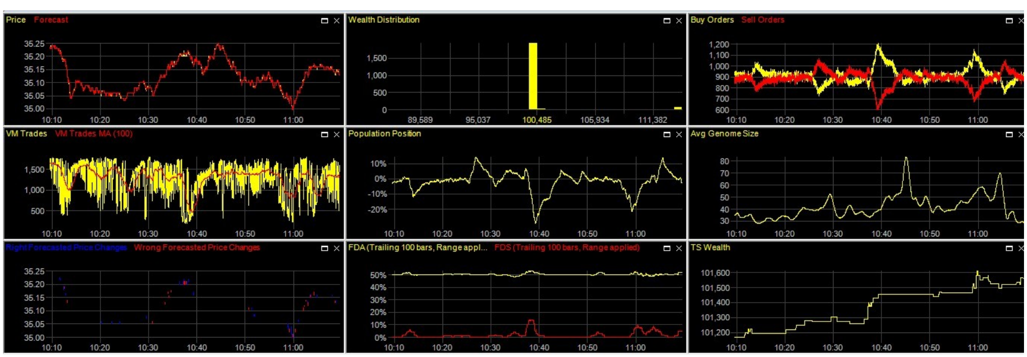

Fig.8 OAddobe stock simulation-Intelligent Agent-Based Model

5 Simulation Results

The results of the simulations are illustrated in Fig.8-25, and were conducted over each of the ten analyzed stocks, for both types of agent-based models, with intelligent agents and zero intelligence agents. The results of the simulations mainly illustrate a better performance of the intelligent agent-based model in terms of forecast accuracy, except for the stocks Addobe and Citigroup. The results must be further tested for how stylized facts are reproduced with these models, compared to real stock quotes.

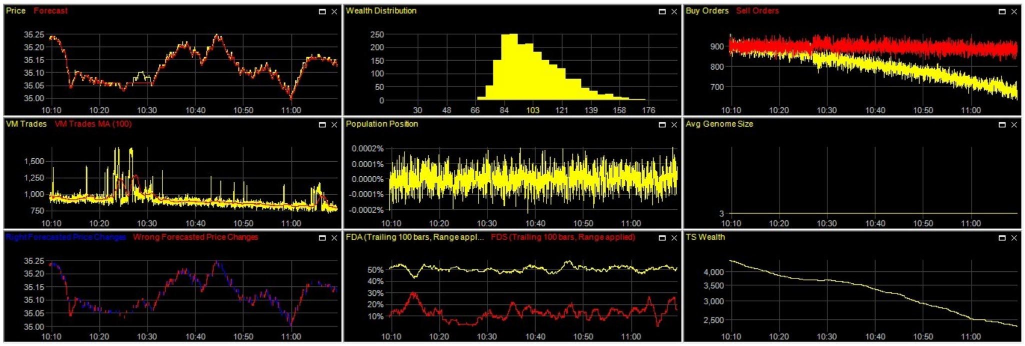

Fig.9 Addobe stock simulation-Zero-Intelligence Agent-Based Model

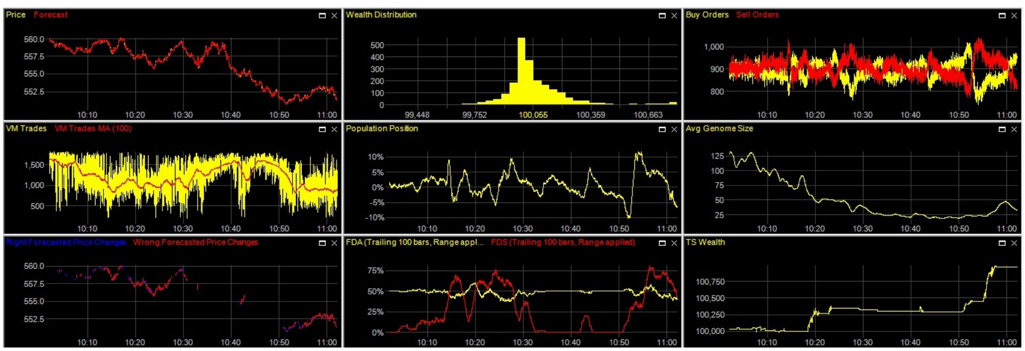

Fig.10 Apple stock simulation-Intelligent Agent-Based Model

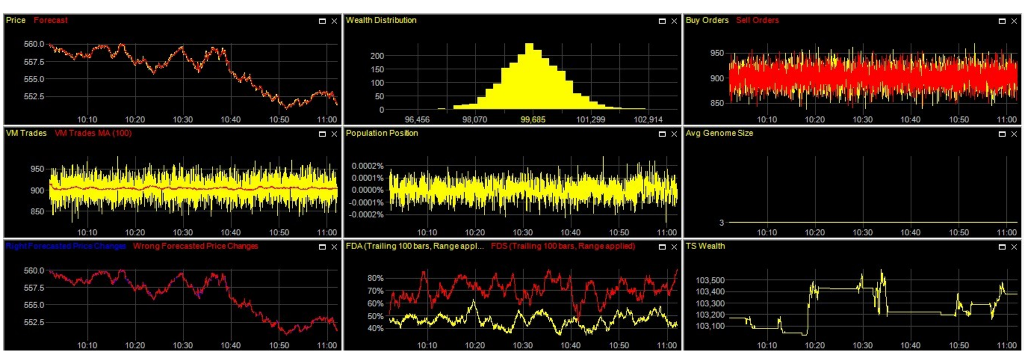

Fig.11 Apple stock simulation-Zero-Intelligence Agent-Based Model

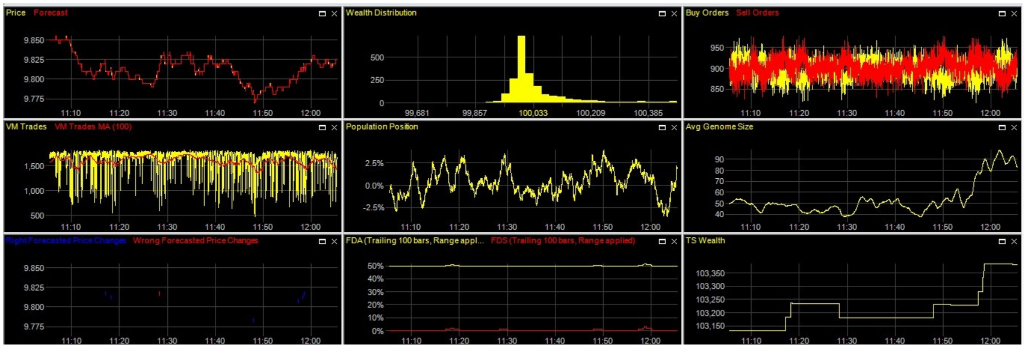

Fig.12 Bank of America stock simulation-Intelligent Agent-Based Model

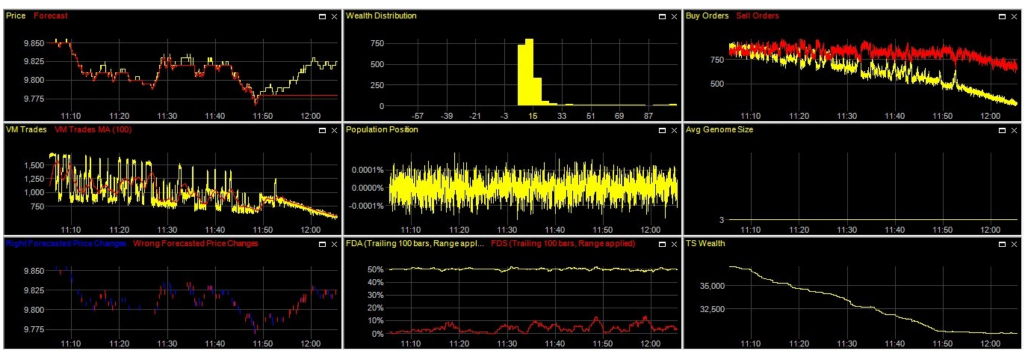

Fig.13 Bank of America stock simulation-Zero-Intelligence Agent-Based Model

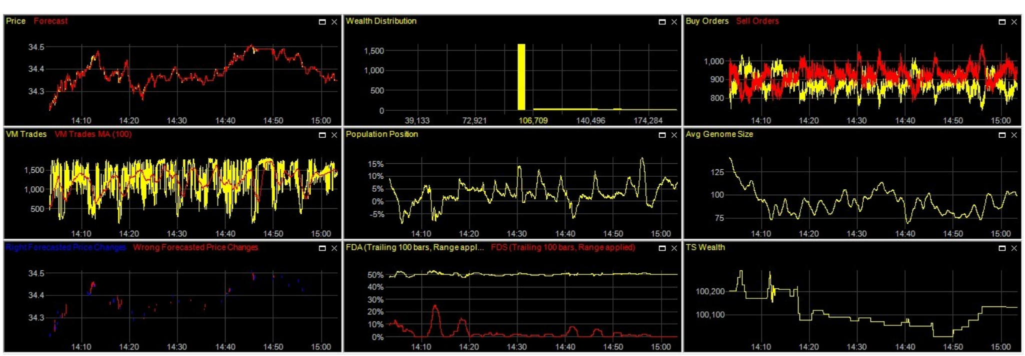

Fig.14 Citigroup stock simulation-Intelligent Agent-Based Model

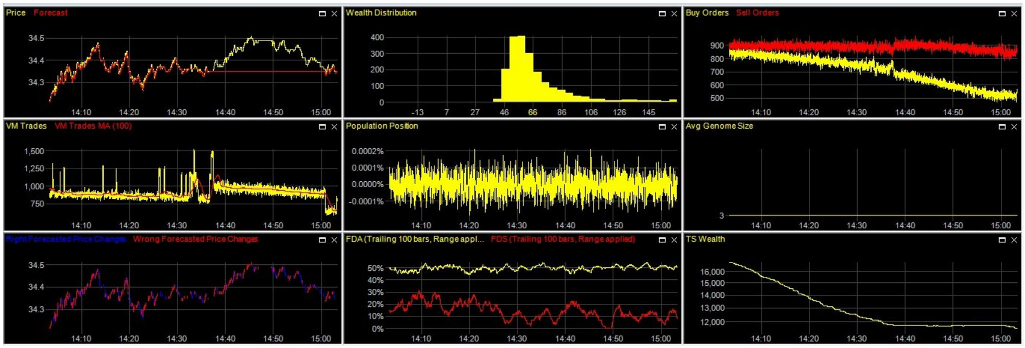

Fig.15 Citigroup stock simulation-Zero-Intelligence Agent-Based Model

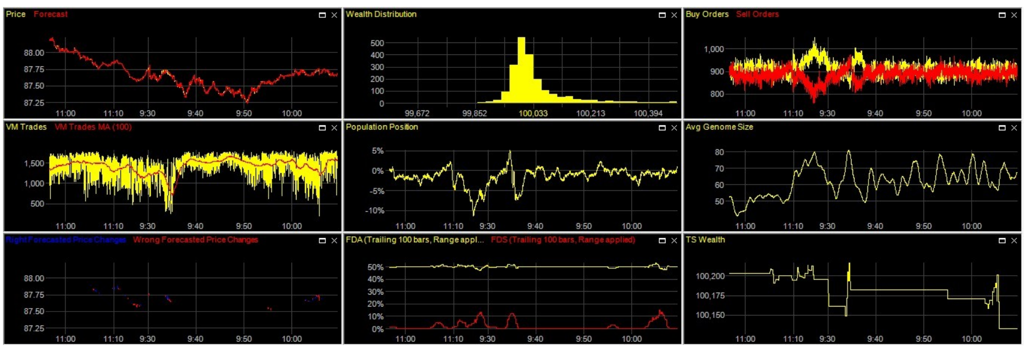

Fig.16 Exxon Mobil stock simulation-Intelligent Agent-Based Model

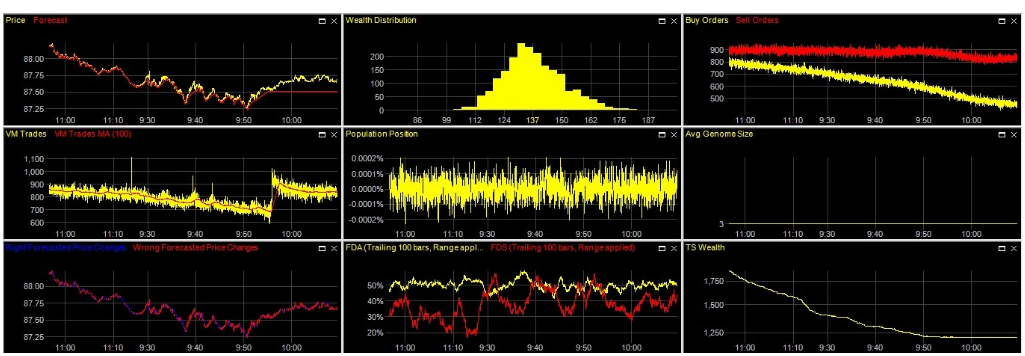

Fig.17 Exxon Mobil stock simulation-Zero-Intelligence Agent-Based Model

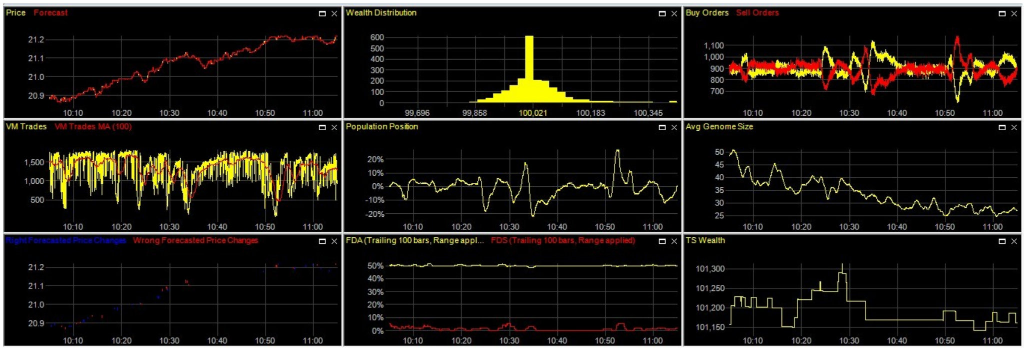

Fig.18 General Electric stock simulation-Intelligent Agent-Based Model

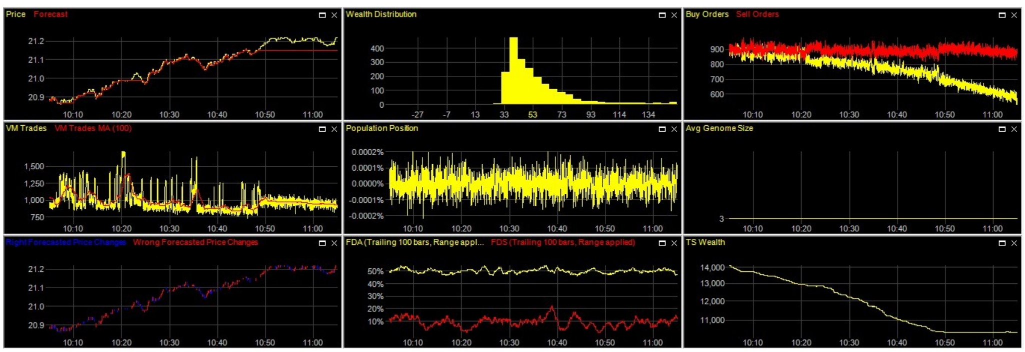

Fig.19 General Electric stock simulation-Zero-Intelligence Agent-Based Model

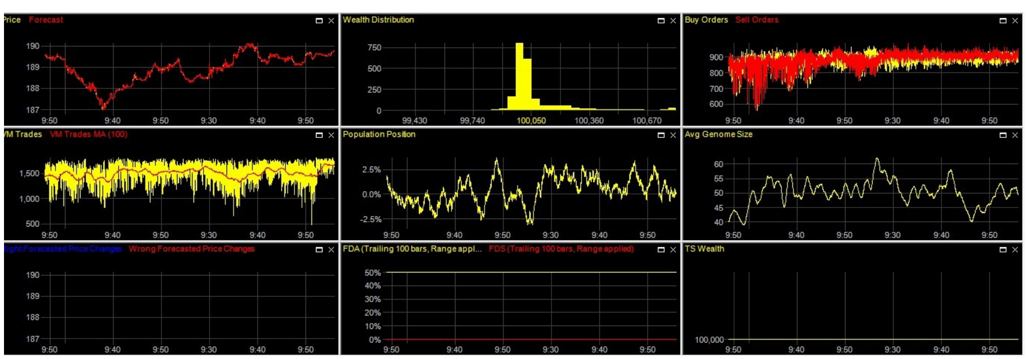

Fig.20 IBM stock simulation-Intelligent Agent-Based Model

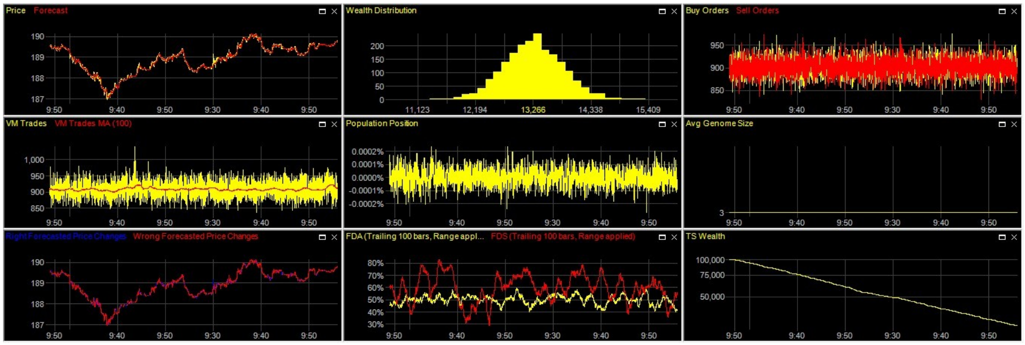

Fig.21 General Electric stock simulation-Zero-Intelligence Agent-Based Model

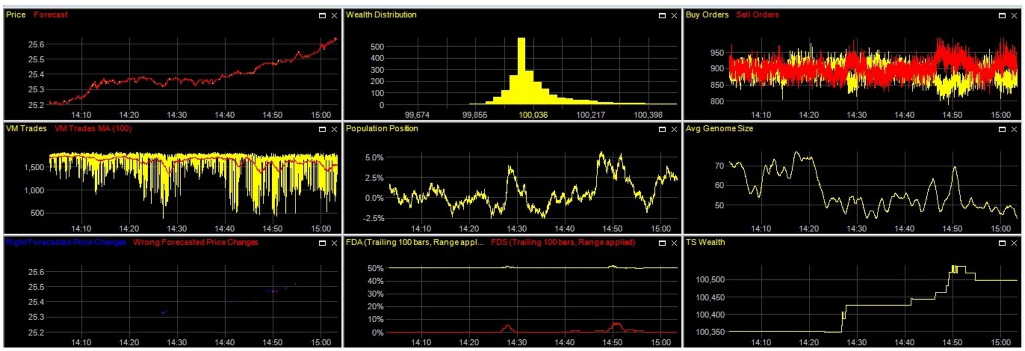

Fig.22 Pfizer stock simulation-Intelligent Agent-Based Model

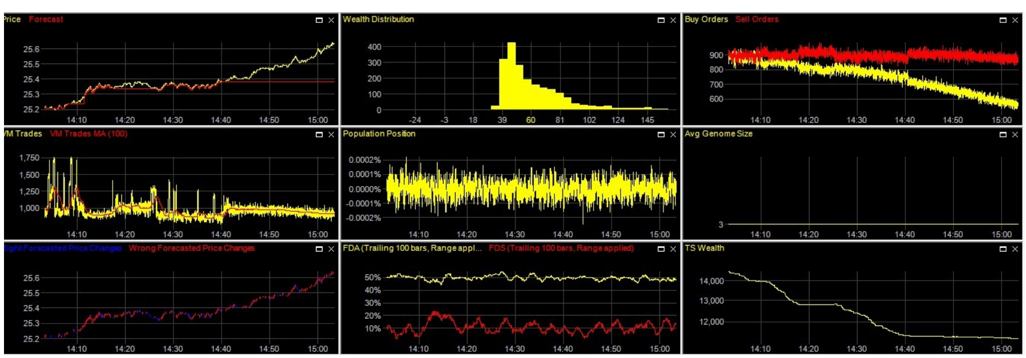

Fig.23 Pfizer stock simulation-Zero-Intelligence Agent-Based Model

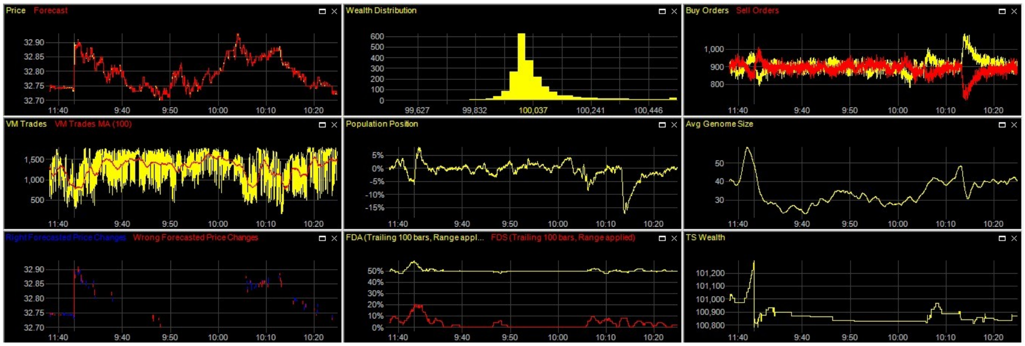

Fig.24 Wells Fargo stock simulation-Intelligent Agent-Based Model

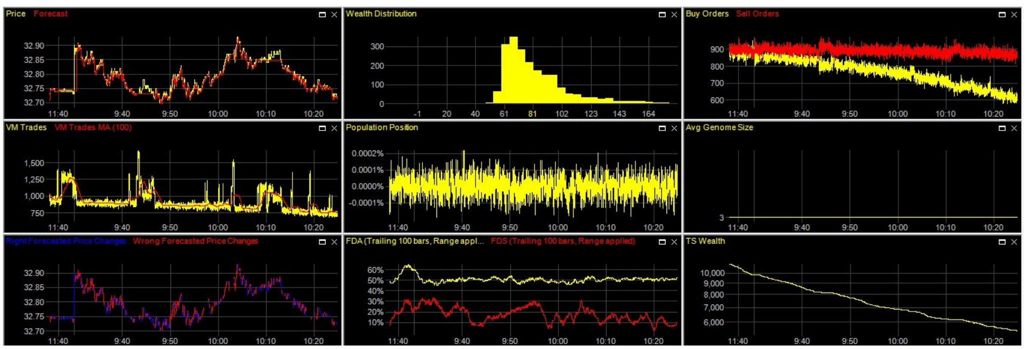

Fig.25 Wells Fargo stock simulation-Zero-Intelligence Agent-Based Model

It is worth mentioning that wealth is much better and uniformly distributed among agents in the case of zero-intelligence agent-based model, as it can be observed in the Fig.8-25, which is of great importance for market design, while intelligent agent-based models generate discrepancies among agents' wealth distribution. On the other hand, the wealth in the intelligent agent-based model simulations is much better preserved compared to the zero intelligence agent-based models, due to the fact that intelligent agent-based model is not a closed economy due to the replacement of the agents which bring new capital in the Virtual Market, while the zero-intelligence agent-based model uses the same traders during all the simulation period. This generates a divergence between the buy orders and sell orders due to lack of cash, thus buy orders.

As illustrated in the Average Genome Size window in the Fig.9-25 for the zero-intelligence agent-based models the value remains constant at 3, as there are only 3 genes (functions) that generate the trades of the population, while for the intelligent agent-based model the value ranges between 30 and 80, showing how the agents strategies constantly adapt and evolve during the simulations.

In order to compare the fit of the models to the real data and to compare the efficiency of the models, the Root Mean Squared Error and Mean Absolute Error have been computed and listed in Table 5. The results show that the intelligent agent-based model performed better compared to the zero-intelligent agent-based model in terms of price forecast.

Table 5 Root Mean Squared Error and Mean Absolute Error for the simulations

Stocks	Model simulated	Root Mean Squared Error	Mean absolute error
Addobe (ADBE)	ZI-ABM	0.010	0.006
	I-ABM	0.032	0.003
Alcoa (AA)	ZI-ABM	0.004	0.003
	I-ABM	0.002	0.001
Apple (AAPL)	ZI-ABM	0.279	0.124
	I-ABM	0.110	0.056
Bank of America (BAC)	ZI-ABM	0.007	0.003
	I-ABM	0.002	0.000
Citigroup Inc. (C)	ZI-ABM	0.027	0.011
	I-ABM	0.110	0.003
Exxon Mobile (XOM)	ZI-ABM	0.040	0.020
	I-ABM	0.010	0.005
General Electric (GE)	ZI-ABM	0.013	0.006
	I-ABM	0.003	0.001
IBM (IBM)	ZI-ABM	0.062	0.041
	I-ABM	0.037	0.020
Pfitzer Inc. (PFE)	ZI-ABM	0.035	0.011
	I-ABM	0.003	0.001
Wells Fargo (WFC)	ZI-ABM	0.016	0.007
	I-ABM	0.009	0.002

6 Conclusions

Taking into account the previous studies on the impact of high frequency trading on the financial markets, the results have shown that this algorithmic trading has led to higher liquidity, improved market efficiency without harming market integrity and lower incidence of market manipulation. Although the results of the research in this field is rather optimist thus far, regarding the impact of HFT over the financial market, the studies are few and do not comprise all the intra-day price impact, nor the impact of the unexecuted orders which are introduced into the market by these trading algorithms. Thus, investors are worried regarding the possible manipulation of the market following these high frequency trades, and currently regulators focus more and more on a better policy for regaining the trust of the investors in the market by limiting the HFT in the markets by imposing that the trading systems are configured such as not to cause market disturbances, a higher control of the market strategies is intended to be implemented, and a fee is also to be charged in the case of excessive use of the system and a limit introduced on the ratio of unexecuted orders.

The results show that in almost all the cases the intelligent agent-based model performed better when compared to the zero-intelligence agent-based model, which could be interpreted as lower market efficiency, allowing for predictions of the stock market price, or even stock market manipulation. This has to be further studied and Efficient Market Hypothesis should be tested. The hereto paper brings to light the importance and accuracy of the agent-based models when it comes to stock market price forecast and modeling, and more important it applies these types of models on high frequency trading data, obtaining high accuracy in forecasting.

Further research should focus on improving the fit of these models with the real stock market by adding brokerage fee and allowing zero-intelligence agents to breed (but not to adapt), thus renewing the population to obtain a zero-intelligence agent-based model closer to reality.

Acknowledgements

This work was co-financed from the European Social Fund through Sectorial Operational Programme Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213 "Ph.D. for a career in interdisciplinary economic research at the European standards".

References

[1] J. Beddington. (2012), "The future of computer trading in financial markets", *The Govornment Office for Science*, London.

[2] J. Brogaard. (2010), "High frequency trading and its impact on market quality", *Social Science Research Network*.

[3] D. Cumming, F. Zhan and M. Aitken. (2012), "High frequency trading and end-of-day manipulation", *Government Office for Science*, London.

[4] M. Aitken, F. Harris, T. Mclnish and A. Aspris. (2012), "High frequency trading - Assessing the impact on market efficiency and integrity", *Government Office for Science*, London.

[5] R. Cont. (2011), "Statistical modeling of high frequency financial data: Facts, Models and Challenges", *IEEE Signal Processing Magazine*.

[6] L. Ponta, M. Raberto, S. Cincotti and E. Scalas. (2012), "Modeling non-stationarities in high-frequency financial time series", *In CORD conference Proceedings*.

[7] M. Aloud, M. Fasli, E. Tsang, A. Dupuis and R. Olsen. (2013), "Modelling the high-hrequency FX market: an Agent-Based Approach", *University of Essex, Technical Reports, Colchester.*

[8] X. Li and A. Krause. (2009), "A Comparison of market structures with near-Zero-intelligence traders", *Intelligent Data Engineering and Automated Learning-IDEAL 2009*, *Lecture Notes in Computer Science*, Vol.5788, pp.703-710.

[9] S. Sunder. (2004), "Markets as artifacts: Aggregate efficiency form zero-intelligence traders", in M. E. Augier and J. G. March, editors, *Models of a Man: Essays In Memory of Herbert A. Simon*, The MIT Press, Cambridge.

[10] D. Montana. (2002), "Strongly typed genetic programming", *Evolutionary Computation*, Vol.3, No.2, pp.199-230, .

[11] T. Vuorenmaa. (2012), "The good, the bad, and the ugly of Automated High-Frequency Trading", *Social Science Research Network.*

[12] J. Barr, T. Tassier and L. Ussher. (2008), "A future of agent-based models in economics: A panel discussion, eastern economic association annual meeting, Boston", *Eastern Economic Journal*, Vol.34, pp.550-565.

[13] R. Cont. (2001), "Empirical properties of asset returns: stylized facts and statistical issues", *Quantitative Finance*, Vol.1, pp.223-236.

[14] "www.altreva.com", Altreva. (2003), Available: http://www.altreva.com /technology.html.

[15] V. Alfi, M. Cristelli, L. Pietronero and A. Zaccaria. (2009), "Minimal agent based model for financial markets I: origin and self-organization of stulized facts", *The European Physical Journal*, Vol.67, No.3, pp.385-397.

[16] G. Daniel. (2006), *Asynchronous Simulations of a Limit Order Book*, University of Manchester, Ph.D. Thesis.

[17] M. Raberto, S. Cincotti, S. Focardi and M. Marchesi. (2001), "Agent-based simulation of a Financial market", *Physica*, Vol.299, pp.319-327.

[18] M. Matsumoto and T. Nishimura. (1998), "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator", *ACM Trans. On Modeling and Computer Simulation*, Vol.8, No.1, pp.3-30.

[19] D. Gode and S. Sunder. (1993), "Lower bounds for efficiency of surplus extraction in double auctions, In the double auction market: Institutions,

theories, and evidence", *In the Double Auction Market, Eds. D. Friedman and J. Rust, SFI Studies in the Science of Complexity*, Proc.Vol. XIV.

[20] V. Walia, A. Byde and D. Cliff. (2003), "Evolving market design in zero-intelligence trader markets", *In IEEE International Conference on E-Commerce (IEEE-CEC03)*, Newport Beach, CA., USA.

[21] D. Ladley. (2012), "Zero intelligence in economics and finance", *The Knowledge Engineering Review*, Cambridge University Press, Vol.27, pp.273-286.

[22] D. Farmer, P. Patelli and I. Zovko. (2004), "The predictive power of zero intelligence in financial markets", *Proceedings of the National Academy of Sciences of the United States of America*, Vol.102, No.6, pp.2254-2259.

[23] D. Gode and S. Sunder. (1993), "Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality", *Journal of Political Economy*, Vol.101, pp.119-137.

[24] J. Rust, R. Palmer and J. Miller. (1989), "A double auction market for computerized traders", Santa Fe Institute, New Mexico.

[25] L. Hurwitz. (1986), "On informationally decentralized systems", *In Decision and Organization, Minnesota*, Univ. of Minnesota Press, pp.297-336.

[26] V. L. Smith. (1982), "Microeconomic systems as an experimental science", *The American Economic Review*, Vol.72, No.5, pp.923-955.

[27] E. Scalas. (2007), "Mixtures of compound poisson processes as models of tick-by-tick financial data", *Chaos, Solitons & Fractals*, Vol.34, pp.33-40.

Corresponding author
D. Dezsi can be contracted at: dianadezsi@yahoo.com