Static and dynamic factors in an information-based multi-asset artificial stock market

Linda Ponta [a,*], Stefano Pastore [b], Silvano Cincotti [a]

[a] DIME—CINEF, Università degli studi di Genova, Via Opera Pia 15, 16145 Genova, Italy
[b] DIA, Università degli studi di Trieste, Via A. Valerio 10, 34127 Trieste, Italy

HIGHLIGHTS

- An information-based multi-asset artificial stock market is modeled and simulated.
- The artificial market is populated by heterogeneous agents.
- Agents are characterized by sentiments and organized in sparsely connected networks.
- Single stock price processes exhibit the principal stylized facts.
- Multivariate price process shows the presence of static factors and common trends.

Keywords:
Artificial financial market
Agent-based modeling

ABSTRACT

An information-based multi-asset artificial stock market characterized by different types of stocks and populated by heterogeneous agents is presented. In the market, agents trade risky assets in exchange for cash. Beside the amount of cash and of stocks owned, each agent is characterized by sentiments and agents share their sentiments by means of interactions that are determined by sparsely connected networks. A central market maker (clearing house mechanism) determines the price processes for each stock at the intersection of the demand and the supply curves. Single stock price processes exhibit volatility clustering and fat-tailed distribution of returns whereas multivariate price process exhibits both static and dynamic stylized facts, i.e., the presence of static factors and common trends. Static factors are studied making reference to the cross-correlation of returns of different stocks. The common trends are investigated considering the variance–covariance matrix of prices. Results point out that the probability distribution of eigenvalues of the cross-correlation matrix of returns shows the presence of sectors, similar to those observed on real empirical data. As regarding the dynamic factors, the variance–covariance matrix of prices point out a limited number of assets prices series that are independent integrated processes, in close agreement with the empirical evidence of asset price time series of real stock markets. These results remarks the crucial dependence of statistical properties of multi-assets stock market on the agents' interaction structure.

Introduction

The increasing interest towards complex systems characterized by a large number of simple interacting units has carried to the birth of co-operations between the fields of engineering, physics, mathematics and economics [1–10]. According to

* Corresponding author.
E-mail addresses: linda.ponta@unige.it (L. Ponta), pastore@units.it (S. Pastore), silvano.cincotti@unige.it (S. Cincotti).

the classical approach, simple analytically tractable models with a representative, perfectly rational agent have been the main corner stones and mathematics has been the main tool of analysis. Conversely, a complexity science approach, where markets are populated by boundedly rational, heterogeneous agents using rule of thumb strategies, fits much better with agent-based simulation models and computational and numerical methods have become an important tool of analysis.

The large availability of financial data has allowed to improve the knowledge about the price processes and many so-called stylized facts have been discovered, e.g., the fat tails of return distributions, the absence of autocorrelation of returns, the autocorrelation of volatility, the distribution of trading volumes and of intervals of trading, etc. [11–16].

In particular, focusing on the distribution of intertrade time between different financial transactions, previous works have demonstrated the presence of Weibull distribution [17,18]. Moreover, empirical study has demonstrated that the dynamics of price and volume of transactions, including the volatility over different time horizons, are influenced by the correlations and temporal patterns of the intertrade times. The rules that regulate the interactions among agents strongly depends on the regulatory mechanisms of each individual market [19].

Furthermore, it is worth noting that, as many biological and physical system, financial markets exhibit multifractal dynamics [20–22].

Generally speaking, these features cannot be reproduced within the theoretical framework of single representative agent.

Over the last 20 years, a number of computer-simulated, artificial financial markets have been put forward. Following the pioneering work done at the Santa Fe Institute [23–25], a large number of researchers have proposed model for artificial markets populated with heterogeneous agents endowed with learning and optimization capabilities [2,26].

This led to a great interest in developing of artificial financial markets based on interacting agents. Several examples of artificial stock markets have been proposed in the literature, e.g., Santa Fe Institute Artificial Stock Market [24] and Genoa Artificial Stock Market [27–38]. While early attempts at microscopic simulations of financial markets appeared unable to account for the ubiquitous scaling laws of returns (and were, in fact, not devised to explain them), the recent models seem to be able to explain some of the statistical properties of financial data, but in most cases the attention is focused only to a single stylized fact. Generally speaking, the objective of artificial markets is to reproduce the statistical features of the price process with minimal hypotheses about the intelligence of agents [39]. Several artificial markets populated with simple agents have been developed and have been able to reproduce some stylized facts, e.g., fat tails of returns and volatility autocorrelation [40,27,41–43,30]. Moreover, also the role of fraudulent agents and of corruption in financial markets has been investigated [38,44]. In particular, empirical analysis show that corruption influences the economic growth rate and foreign investment [45]. Furthermore, the artificial financial markets are a useful framework to study how the fraudulent agents impact on the markets [38]. For a detailed review on microscopic (''agent-based'') models of financial markets see [46,47]. Generally speaking, in the framework of artificial stock markets, attention has been focused on single asset artificial stock markets. This in order to understand and to reproduce the main stylized facts of an univariate price process. Computational experiments pointed out the possibility to reproduce some stylized facts in terms of the single price process but, results suggested a reduced capability in reproducing the well known unitary root stylized fact, as it was obtained only in the presence of exogenous cash inflow.

This limitation can be overcome employing recent results on a single-asset artificial stock market based on information propagation [32]. Starting from the model proposed in [32], generalizations for the multi-assets framework have been proposed, where some statistical properties of univariate and multivariate price processes have been reproduced, but the models resulted unable to reproduce the endogenous presence of common trends [33,34]. This paper presents an extension of the Genoa Artificial Stock Market (GASM) that addresses such topic by means of an information-based multi-asset artificial stock market model. The market is populated by heterogeneous agents that are seen as nodes of sparsely connected graphs. The market is characterized by different types of stocks and agents trade risky assets in exchange for cash. Beside the amount of cash and assets owned, each agent is characterized by sentiments. Moreover, agents share their sentiments by means of interactions that are determined by graphs. The allocation strategy is based on sentiments and wealth. A central market maker (clearing house mechanism) determines the price processes for each stock at the intersection of the demand and the supply curves. The validation method followed in this paper is the capability of the information-based artificial stock market to reproduce static and dynamic stylized facts for univariate and multivariate price processes. Concerning univariate processes, the three main stylized facts are taken as reference, i.e., unitary root of price processes, fat-tails distribution of returns and volatility clustering. The multi-assets environment offers a new set of stylized facts for validation, i.e., the statistical properties of cross-correlation matrices of returns [48–50] and of variance–covariance matrices of prices [51], that make reference to static and dynamic factors, respectively.

The computational experiments discussed in this paper show that concerning univariate price processes, the proposed model is able to reproduce unit root, volatility cluster and fat tails of returns. Concerning the multivariate price process, they exhibits both static and dynamic stylized facts, i.e., the presence of static factors and common trends. Thus, results show that the main statistical properties of univariate and multivariate price processes are reproduced, thus remarking the crucial role covered by the interaction structure among the agents. The paper is organized as follows: Section 1 presents the model, Section 2 the empirical data, Section 3 shows the computational experiments and Section 4 the discussion of results. Finally, Section 5 provides the conclusion of the study.

1. The model

Heterogeneous and informed agents trade risky assets in exchange for cash depending on the interactions among agents. They are modeled as liquidity traders, i.e., decision making process is constrained by the finite amount of financial resources (cash and stocks) they own. Let N be the number of traders and K the number of assets. Let L be the number of sectors that characterize the economy, e.g., construction, information technology, manufacturing, etc. Let l denote the particular economic sector and the pair l, m the asset m in the sector l. Let M_l be the number of assets of sector l and $S_i^{l,m}$ be the sentiment about the asset l, m of agent i.

For each asset, the traders in the market are organized according to a directed random graph, where the agents are the nodes and the branches represent the interactions among agents. The graphs are directed, i.e., the interactions are assumed unidirectional (i.e., agent j influences agent i but not necessarily vice versa) and characterized by a strength $g_{ji}^{l,m}$, assumed a positive real number. Generally speaking, due to the presence of a directed graph, both an output node degree, related to the output branches of a given node, and an input node degree, related to the input branches, should be defined.

The agents are ranked according to a Zipf law, i.e., the importance of each agents is approximately proportional to its rank. That model the evidence that the most important agents have larger amount of cash owned, larger number of assets owned. All the parameters of the agents are calculated according to such a ranking. Moreover, for each stock an agent is randomly connected to a set of other agents whose number and strength $g_{ij}^{l,m}$ are inversely proportional to his rank, i.e., richer agents influences a larger number of agents with a higher strength. Consequently, the output degree distributions over the nodes are set to power laws and the input degree distributions result power laws too.

Agent i is characterized by a sentiment $S_i^{l,m}$ (i.e., real number in the interval $[-1,1]$) that represent a propensity to invest in asset m of sector l. The graphs are responsible of the changes in agent's sentiments. At each time step h, information is propagated through the market and sentiments of agent i is updated. For each assets, belonging to the chosen sectors, a positive average sentiment denotes a propensity to buy, whereas a negative average sentiment corresponds a propensity to sell.

Let $\Im_i^{l,m}$ the set of agents that influence the behavior of trader i for the asset m belonging to the sector l and $p^{l,m}$ the market price of the risky asset l, m. The new sentiment $S_i^{l,m}$ of agent i at time $h+1$ is functions of her previous sentiment at time t, of the influence of neighbor agents, of market feedback and of the average sentiments of the agent, i.e.,

$$S_i^{l,m}(h+1) = tanh\left[\alpha_{P,i}S_i^{l,m}(h) + \alpha_{N,i}\hat{S}_i^{l,m}(h) + \alpha_{M,i}r^{l,m}(h) + \alpha_{G,i}\tilde{S}_i(h) + \alpha_{S,i}S_i^{l}(h)\right] \tag{1}$$

where

$$\hat{S}_i^{l,m}(h) = \frac{\sum_{j\in\Im_i^{l,m}} g_{ji}^{l,m} S_j^{l,m}(h)}{\sum_{j\in\Im_i^{l,m}} g_{ji}^{l,m}} \tag{2}$$

represents the influence of neighbor agents,

$$r^{l,m}(h) = \log\left[p^{l,m}(h)\right] - \log\left[p^{l,m}(h-1)\right] \tag{3}$$

represents the market feedback,

$$\tilde{S}_i(h) = \frac{\sum_{l,m} S_i^{l,m}}{K} \tag{4}$$

models the vision of agent i on the global market, and

$$S_i^{l}(h) = \frac{\sum_{m\in l} S_i^{l,m}}{M_l} \tag{5}$$

is the average sentiment of agent i on sector l. Eq. (1) generalizes the model proposed in [33], by introducing the agent vision for each sector (see Eq. (5)). It is worth remarking that the $\alpha_{P,i}$ coefficient in Eq. (1) are inversely proportional to agent's rank,i.e., richer agents have stronger believes. Moreover a constraint on graph interaction is considered

$$|\alpha_{N,i}| = (\eta - |\alpha_{P,i}|) \tag{6}$$

i.e., self-interaction is a counterpart of graph interactions, with randomly (i.e., based on an uniform distribution) changes in sign at each time step. Eq. (6) models a specific behavior of agents, i.e., the fact that sometimes an agent changes idea about the sentiments of neighbor, and so she changes her reaction. In fact, Eq. (6) points out that agent that are strongly influenced by their previous sentiment (e.g. big traders, banks, mutual funds, etc.) are poorly influenced by the neighboring agents' sentiments (e.g., small single investors) and η represents the self-neighboring sentiment balance coefficient [32,33].

The amplitude of market feedback depends on rank, so that the coefficients $\alpha_{M,i}$ are inversely proportional to agent ranks, that is agents with higher ranks are less sensitive to the single asset trends. Moreover, the $\tilde{S}_i(h-1)$ term is a stabilizing element for the sentiment, so that the coefficient $\alpha_{G,i}$ in Eq. (1) is always negative. Finally, the $S_i^l(h)$ term emphasize the sector l sentiment with respect to the global market vision and α_S in Eq. (1) is always positive.

Agent's trading decision is based on cash and stocks owned and on sentiment. In particular, the stock price processes depend on the propagation of information among the interacting agents, on budget constraints and on market feedbacks. In this respect, also the $\alpha_{P,i}$ coefficient in Eq. (1) is proportional to agent's rank, i.e., richer agents have stronger beliefs.

Let $c_i(h)$ the amount of cash, $q_i^{l,m}(h)$ the amount of asset l, m owned by the trader i at time h.

The risky wealth $W_i^r(h)$ owned by trader i at time step h is:

$$W_i^r(h) = \sum_{l,m} q_i^{l,m}(h) p^{l,m}(h) \tag{7}$$

whereas $W_i(h) = c_i(h) + W_i^r(h)$ represents the total wealth of agent i.

At each simulation step, trader i issues orders in a subset of sectors. Let us assume that trader i chooses sector l, then she invests in the assets belonging to this sector. Moreover, trader i tries to allocate in risky assets a fraction γ_r of his total wealth related to his vision of the market trend, i.e.,

$$\widehat{W}_i^r(h+1) = \gamma_r(h) W_i(h), \tag{8}$$

where $\gamma_r = \frac{1+\tilde{S}_i(h)}{2}$.

$S_i\tilde{(h)}$ is the average sentiments on all assets at time t described by Eq. (4). The symbol $\widehat{\cdot}$ denotes that $\widehat{W}_i^r(h+1)$ is the amount that agent $i-th$ desires to allocate in the risky investment, whereas the real amount $W_i(h+1)$ effectively allocated in stocks will depend on the market result. It is worth remarking that markets are assumed imperfect, i.e., rationing appears for both demand and supply. In this model only long positions are allowed. Thus, if the agent i is characterized by a positive average sentiment on the risky asset l, m, the desired quantity is given by:

$$\widehat{q}_i^{l,m}(h+1) = \left\lfloor \frac{\gamma_a^{l,m} \widehat{W}_i^r(h+1)}{p^{l,m}(h)} \right\rfloor \tag{9}$$

where γ_a is:

$$\gamma_a^{l,m} = \frac{S_i^{l,m}}{\sum_{l,m \in A_i} S_i^{l,m}}, \tag{10}$$

and A_i is the set of assets with positive sentiments in sector l. The symbol $\lfloor .. \rfloor$ in Eq. (9) denotes the integer part. Conversely, if the sentiment relatives to asset l, m is negative, the agent i is characterized by a desired quantity $\widehat{q}_i^{l,m}(h+1) = 0$.

The quantity $\Delta_i^{l,m}(h+1)$

$$\Delta_i^{l,m}(h+1) = \hat{q}_i^{l,m}(h+1) - q_i^{l,m}(h). \tag{11}$$

is the difference between the desired quantity of stock l, m at time step $h+1$ and the quantity of stock l, m in the portfolio by agent i at time step h. If $\Delta_i^{l,m} > 0$ the order is a buy order. Conversely, if $\Delta_i^{l,m} < 0$ the agent issues a sell order. Finally, every order is associated with a limit price as discussed in the following subsection.

1.1. Clearinghouse mechanism

According to previous models [27–30,32–34], we stipulate that buy (sell) orders cannot be executed at prices above (below) their limit price $d_i^{l,m}$, i.e.,

$$d_i^{l,m}(h+1) = p^{l,m}(h) \cdot N_i(\mu_i^{l,m}, \sigma_i^{l,m}), \tag{12}$$

$N_i(\mu_i^{l,m}, \sigma_i^{l,m})$ is a random draw from a Gaussian distribution with average

$$\mu_i^{l,m} = \left(1 + sgn(\Delta_i^{l,m}) |S_i^{l,m}|\right). \tag{13}$$

It is worth noting that for a buy order (i.e., $\Delta_i^{l,m} > 0$) in average $d_i^{l,m}(h+1) > p^{l,m}(h)$, whereas for a sell order (i.e., $\Delta_i^{l,m} < 0$) in average $d_i^{l,m}(h+1) < p^{l,m}(h)$. Furthermore, the standard deviation $\sigma_i^{l,m}$ is proportional to the historical volatility $\sigma^{l,m}(T_i)$ of the price $p^{l,m}(h)$ of stock l, m through the equation $\sigma_i^{l,m} = \xi \sigma^{l,m}(T_i)$. Linking limit orders to volatility takes into account a realistic aspect of trading psychology: when volatility is high, uncertainty on the "true" price of a stock grows and traders place orders with a broader distribution of limit prices. In our model, ξ is a constant for all agents, whereas $\sigma^{l,m}(T_i)$ is the standard deviation of log-price returns of asset l, m, computed in a time window T_i proper for agent i. In particular, T_i is randomly drawn from a uniform distribution of integers in the range from 10 to 100 for each trader at the beginning of the simulation [27]. All buy and sell orders issued at time step $h+1$ are collected and the demand and supply curves are consequently computed. The intersection of the two curves determines the new price (clearing price) $p^{l,m}(h+1)$ of stock l, m (see [27] for more details on market clearing).

Buy and sell orders with limit prices compatible with $p^{l,m}(h+1)$ are executed. After any transactions, traders' cash, portfolio and sentiments are updated. Orders that do not match the clearing price are discarded.

Table 1
Economic scenarios considered in the computational experiments.

Scenario	Sector Context	Allocation Universe	$\alpha_{S,i}$
i	No exogenous sectors	The whole market	$= 0$
ii	Exogenous sectors	The whole market	$= 0$
iii	Exogenous sectors	The sector with largest average sentiment $S_i^l(h)$	$= 0$
iv	Exogenous sectors	The sector with largest average sentiment $S_i^l(h)$	> 0

2. Empirical data

The data set used for this paper consists of daily close prices taken from a subset of the assets belonging to the S&P500, i.e., the most capitalized US assets traded at the NYSE and NASDAQ markets. We considered the time period from 2, Jan, 2009 to 2, Jan, 2012 and the resulting database was composed by 500 time series, 800 points long each. The subset of S&P500 employed in our computational experiments is composed by 100 assets randomly chosen among the 500 composing the S&P500. In order to find the number of sectors in which the 100 assets are distributed, the Inverse Participation Ratio (IPR), as explained in [50], has been used. In particular, the IPR quantifies the reciprocal of the number of eigenvector components that contribute significantly. In our case the significant components represent the assets belonging to the same sectors. We found the presence of four sectors, whereas each asset belongs to at least one sector.

3. Computational experiments

Generally speaking, the main objective of an artificial market is to reproduce the statistical features of the price processes with minimal hypotheses about the intelligence of agents. In this paper we adopted this approach in order to validate the model. The computational experiments have been performed on a market characterized by 100 different stocks, each related to a specific firm. At the beginning of each simulation, cash and stocks are distributed randomly among agents. Starting from the results obtained by the IPR analysis on real data, we define the number of market sectors presented in our model and which firms belong to each sector. According to empirical data, all stocks are divided in four sectors and each asset may belong to one or more sectors. Two different economic situations are considered:

(a) absence of exogenously-defined sectors
(b) exogenously-defined sectors with randomly assigned assets.

In the case of exogenous sectors, the allocation universe is a single sector including either the whole market (i.e., all 100 assets) or the sector characterized by the largest average sentiment $S_i^{l,m}(h)$ (see Eq. (5)). Conversely, in the case of absence of exogenous sectors, the allocation universe is the whole market (i.e., all 100 assets). Finally, two cases are considered for coefficient $\alpha_{S,i}$ (see Eq. (1)), i.e., (1) $\alpha_{S,i} = 0$ and (2) $\alpha_{S,i} > 0$. It is worth remarking that $\alpha_{S,i} > 0$ results in a distinctive presence of the sentiment update rule on the average sentiment $S_i^l(h)$ of sector l with respect to the other sectors, as discussed in Section 1.

Table 1 summarizes the main characteristics of the four scenarios considered in the computational experiments.

4. Results

In all simulations, the market is characterized by 100 different assets and the number of iteration is equal to 1920 time step (i.e., 8 years). Using the results of the IPR analysis on 100 assets chosen randomly among the S&P500 index, the number of market sectors considered is 4. The number of agents N is set equal to 2,278 initially characterized by a random distribution of cash and number of stocks. Each agent is also characterized by the following values: α_P ranging between 0.22 and 0.50, α_N between -0.38 and 0.38, α_M between -6.00 and 6.00, α_G between -0.10 and 0.04 and α_S between 0 and 0.3. The sentiment $S_i^{l,m}$ (see Eq. (1)) ranges between -1 and 1. Furthermore, time window T_i for the calculation of the historical standard deviation is randomly chosen by a uniform distribution in the range $(10, 100)$.

Fig. 1 shows the prices processes for the $k = 100$ assets for the scenarios described in Table 1. The price processes exhibit relevant differences, depending on the specific nature of the asset. Indeed, after a transient (not shown in Fig. 1 for the sake of compactness) price levels result significantly different. This suggest a possible herding behavior induced by the graphs that drives the agent propensities to buy/sell the assets.

A statistical analysis on single asset is performed so to verify the univariate stylized facts. Fig. 2 shows the prices and the returns process of asset number 27 and the corresponding autocorrelation function of raw returns, of absolute value of returns and of the square returns for scenario (i). Volatility clusters and long memory effect in the autocorrelation function of absolute value of returns and square returns are pointed out. As clearly shown in Fig. 2(c) autocorrelation of raw returns shows immediate decay within noise level of the correlation after just one lag, whereas absolute value of returns and square returns exhibits slow decay of the autocorrelation. These properties are robust features of the proposed artificial market, as remarked by the same results obtained for the other scenarios in Table 1 not included for the sake of compactness.

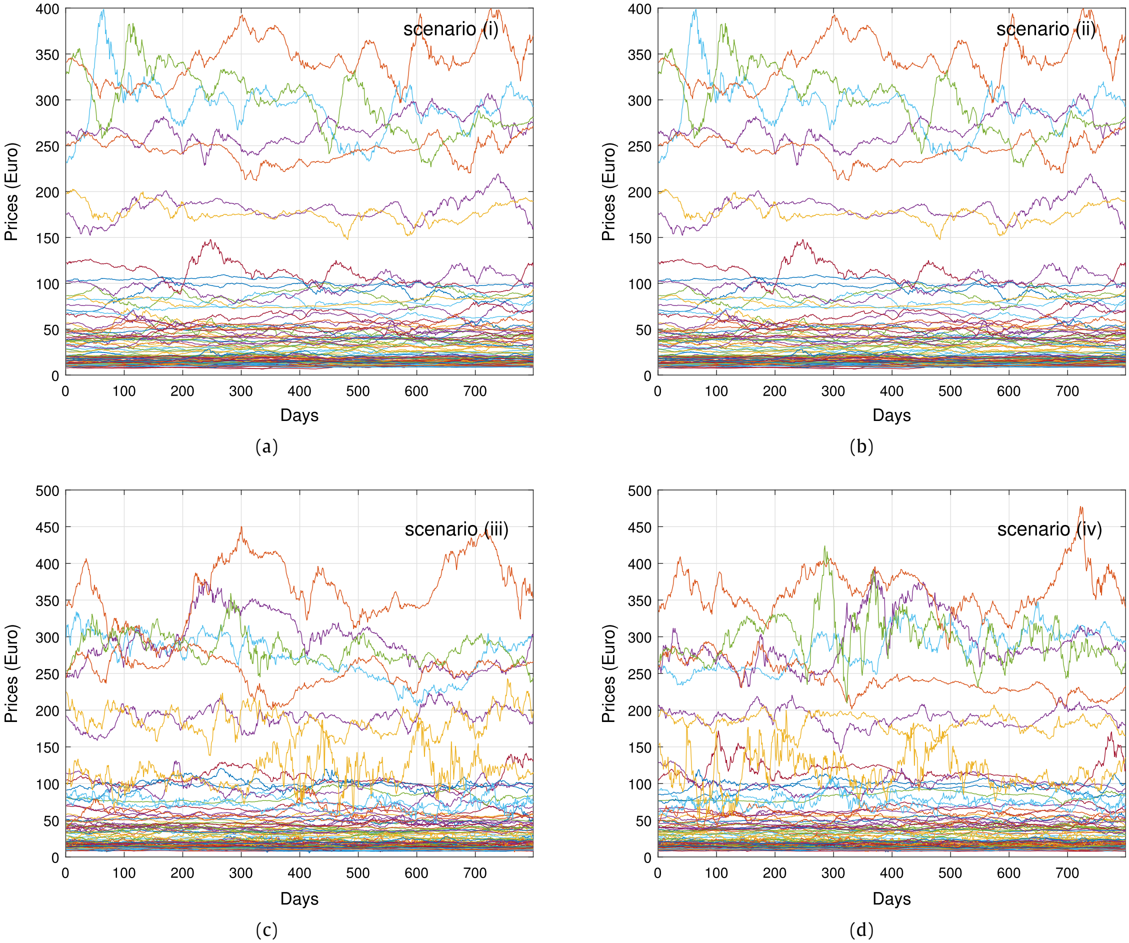

Fig. 1. Price processes of $k = 100$ assets for all scenarios.

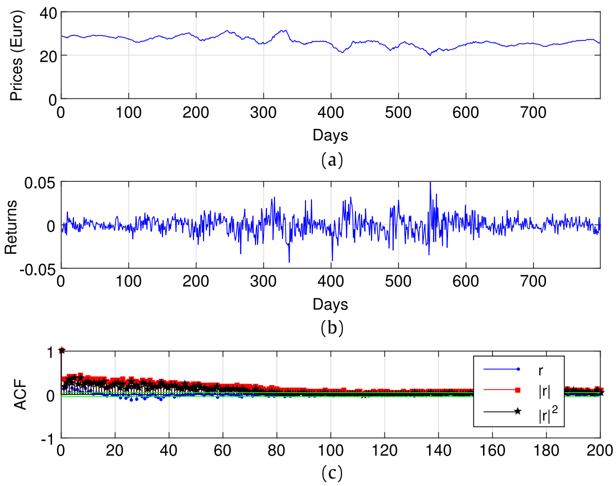

Fig. 2. Price process, returns and autocorrelation function of returns process of stock number 27 for scenario (i).

Focusing attention on statistical tests, the normal distribution, the unitary root of returns and the presence of heteroscedastic effect have been checked by Jarque–Bera, Augmented Dickey–Fuller and ARCH tests, respectively, at the

Table 2
Augmented Dickey–Fuller test, Jarque–Bera test and ARCH test for GASM and real (RND_{100}(S&P500)) data.

Data	ADF test not rejected	J–B test rejected	Arch test rejected
Scenario (i)	98	100	93
Scenario (ii)	98	100	93
Scenario (iii)	74	99	92
Scenario (iv)	70	100	65
RND_{100}(S&P500)	80	100	71

Table 3
Number of static and dynamic factors determined using RMT and co-integration respectively (see text).

	Static factors	Dynamic factors
Scenario (i)	3	20
Scenario (ii)	4	20
Scenario (iii)	3	5
Scenario (iv)	5	2
RND_{100}(S&P500)	5	11

significance level of 5%. Table 2 summarizes the results of the computational experiments together with those obtained from the S&P500 (that are included for the sake of comparison). Table 2 reports the number of assets that do not reject the hypothesis of unitary root (ADF test), the number of assets whose returns process does not follow a normal distribution (J–B test) and the number of asset that present heteroscedastic effects. It is worth noting that the value of GASM data and the real data are in very good agreement i.e., the GASM reproduced the main stylized facts of univariate processes in most of the cases. Moreover, these results are in good agreement with the observations of market dynamics in historical and archaeological data [52]. This confirm the quality of the information-based artificial stock market allowing us to conclude that the interactions networks are able to reproduce the univariate stylized facts also in the presence of a large number of assets.

Stated these results, the attention has been focused on the statistical properties of the multivariate process of prices and returns. Generally speaking, the analysis of multivariate stylized facts leads to the definition of factor models. Furthermore, in the context of factor models, two main classes can be identified, i.e., static and dynamic factors. Concerning the former class, attention is paid to returns as the return processes result (in the first approximation) quasi-stationary. In particular, the risk of a security can be described as superposition of different source of risks (also described by stationary processes) and this general formulation is the basic for classical portfolio theory and risk management, e.g., CAPM, multifactors CAPM, APT, etc. [53–56].

Conversely, in the case of dynamic factors attention is paid to asset prices and the main employed concept is co-integration. In particular, statistical analysis on empirical data points out that in a large market it is not possible to reject the hypothesis of integrated univariate price processes, but at the aggregate level the price processes are not independent. Indeed, only few independent integrated processes can be identified, whereas all the others price processes are co-integrated with them, i.e., it is possible to identify linear combinations of $I(1)$ price processes that result stationary $I(0)$ processes (so called co-integration equations) [53–56].

In this paper, the static factors are studied according to the cross-correlation matrix of returns. In particular, following the approach introduced in the econophysics literature by [48–50], the cross-correlations of returns have been studied by means of the random matrix theory (RMT). Fig. 3 shows the probability density function (PDF) of eigenvalues of the cross-correlation matrix for the different scenarios in Table 1. Furthermore, for the sake of comparison, the theoretical PDF of a random matrix (represented by the continuous line) as well as the PDF of eigenvalues for the 100 stocks of the S&P 500 index (i.e., those used to evaluate the IPR and number of sectors) are also shown. Table 3 summarizes the presence of outliers in the computational experiments well above the bounds determined according to RMT (i.e., eigenvalues larger than the largest eigenvalue determined by the RMT), in close agreement to the empirical evidence shown by S&P500 data.

It is worth noting that the largest outlier (i.e., the eigenvalue representing the market) is always present within the proposed scenarios. This suggest that the presence of the market factor is a strong feature of the information-based artificial stock market, mostly due to finiteness of the context (i.e., limited cash amount and share numbers). Conversely, the allocation universe and strategy of the agents result critical so to obtain the outliers representing the other economic sectors. In particular, only the introduction of constraints in the sentiment dependencies are able to reproduce the static factor stylized fact, thus confirming the crucial role of the interaction networks. Furthermore, it is worth remarking that the term $\alpha_{S,i}$ in Eq. (1) play a critical trade-off role as the larger the $\alpha_{S,i}$ the larger the number of economic sectors (i.e., the number of outliers), but the larger the $\alpha_{S,i}$ the smaller the market eigenvalue (i.e., the largest eigenvalue). It is worth remarking that in Fig. 3 the PDF of eigenvalues of GASM data in scenarios from (i) to (iv) are presented in blue, electric blue, green, brown and yellow respectively and the theoretical PDF for random matrices is represented by the black continuous line. In this case, i.e., 100 series of returns and 800 time steps, the largest eigenvalue results equal 1.83. For the sake of comparison, the yellow colored histogram in Fig. 3 shows the PDF of eigenvalues for the random sample of 100 stocks included in the S&P500 index

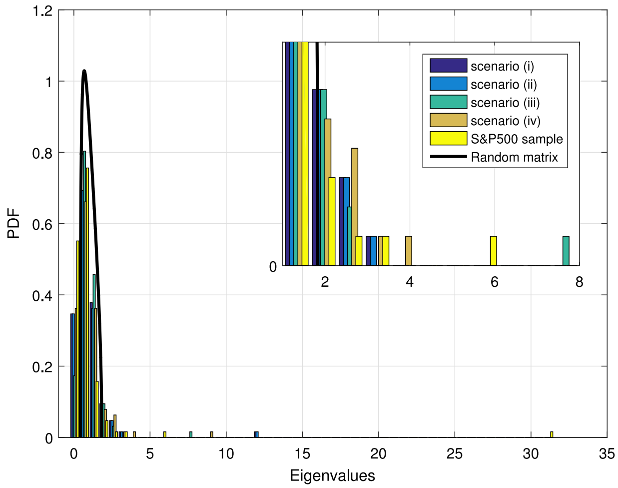

Fig. 3. Probability density function (PDF) for eigenvalues cross-correlation matrix of returns.

in a time window of 800 business days (i.e., closed prices from the year 2009 to the year 2012 are considered). As shown in Table 3, also Fig. 3 confirms the presence of outliers representing the market and the business sectors.

As regarding the dynamic factors, they have been studied by means of the variance–covariance matrix of prices. According to empirical analysis, only a reduced number of assets prices series in a large market are independent integrated processes [51]. In fact, the analysis of prices processes shows that financial assets are random walk, i.e., $I(1)$ processes, but aggregate of financial assets exhibits co-integration. The analysis of this property has been performed following the procedure described by Stock and Watson [51]. In particular, the PCA analysis on the variance–covariance matrix of prices allows one to identify portfolios with minimum variance. Conversely to price processes, these portfolios, i.e., linear combination of prices, generally accept the hypothesis of stationarity [51] that can be verified by the ADF test at significance level of 5%. Fig. 4 shows the results of the ADF test for the GASM computational experiments and for the series of the 100 randomly selected assets of S&P500 index. As clearly stated in Fig. 4, in the case of the assets of S&P500 data only a reduced number of portfolios (i.e., equal to 11) reject the hypothesis of stationarity. These time series are the only independent $I(1)$ processes, i.e., the common trends of the aggregate, whereas there exist 89 cointegration equations (i.e., the $I(0)$ portfolios). Table 3 summarizes the results of the co-integration analysis and point out that also in the case of GASM data, only a reduced number of portfolios reject the hypothesis of stationarity. These series are the only independent $I(1)$ processes, i.e., the common trends of the aggregate. It is worth remarking that the evidence of dynamic factors directly originate by the interaction of agent, i.e., the information-based decision process taking place by means of the network interconnections. Furthermore, the presence of the common trends of the aggregate is a strong features that is present in all considered scenarios, thus remarking the crucial role played by information-based interactions.

These results allow us to conclude that the proposed information-based artificial stock market is able to reproduce the statistical properties of single-asset environment as well as the stylized facts of multi-assets. It is worth noting that both the stylized facts on returns and on prices (i.e., static and dynamic factors) are in close agreement with empirical evidences for real data. This points out that interactions process governed by the networks is the main driving and explanation mechanism.

5. Conclusion

An artificial stock market characterized by heterogeneous and informed agents has been studied. In this complex system, agents are characterized by cash, stocks and sentiments. Sentiments denote propensities to buy or to sell. Agents are seen as nodes of sparsely connected graph, so that each agent is influenced by a subset of other agent, the only ones that are "near" to him. The statistical properties (i.e., stylized facts) of the univariate and the multivariate process of prices and returns have been investigated. In particular, concerning univariate price processes, the proposed approach was able to reproduce unit root, volatility cluster and fat tail distribution of returns. Furthermore, concerning the multivariate price process, the cross-correlations between returns of different stocks have been studied using methods of Random Matrix Theory (RMT) and the variance–covariance matrix of price using the Principal Component Analysis (PCA). The computational experiments pointed out the ability to reproduce both static and dynamic stylized facts of uni- and multi-variate processes. In particular, results confirmed the crucial role played by the interaction networks for the univariate properties. Furthermore, finiteness of the market was shown to be responsible for the evidence of the largest static factors whereas the presence of sectors arise only as consequence of sentiment dependencies in the allocation universe of the traders. Finally, the network based

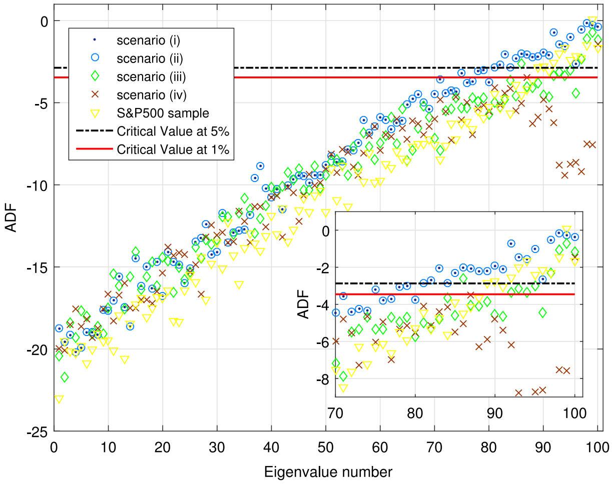

Fig. 4. ADF test statistics of the cointegration portfolios in the case of RND_{100} (S&P500) and GASM data.

information approach was able to endogenously reproduce the presence of dynamic factors in the artificial stock market. It is worth remarking that for the first time all static and dynamic features at uni- and multi-variate level presented in the paper has been reproduced in a single framework.

Acknowledgment

This work has been partially supported by the University of Genoa, under grant no. 100025-2015-LP-FRA_001.

References

[1] G.-H. Mu, W.-X. Zhou, W. Chen, J. Kertsz, Order flow dynamics around extreme price changes on an emerging stock market, New J. Phys. 12 (2010) 075037.
[2] F. Westerhoff, An agent-based macroeconomic model with interacting firms, socio-economic opinion formation and optimistic/pessimistic sales expectations, New J. Phys. 12 (2010) 075035.
[3] R. Coelho, C.G. Gilmore, B. Lucey, P. Richmond, S. Hutzlera, The evolution of interdependence in world equity markets — evidence from minimum spanning trees, Physica A 376 (2007) 455–466.
[4] P.C. Ivanov, M.G. Rosenblum, C.K. Peng, J. Mietus, S. Havlin, H.E. Stanley, A.L. Goldberger, Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis, Nature 383 (1996) 323–327.
[5] M. Riccaboni, F. Pammolli, S.V. Buldyrev, L. Ponta, H.E. Stanley, The size variance relationship of business firm growth rates, Proc. Natl. Acad. Sci. USA 105 (2008) 19595–19600.
[6] R.N. Mantegna, J. Kertész, Focus on statistical physics modeling in economics and finance, New J. Phys. 13 (2011) 025011.
[7] M. Tumminello, F. Lillo, J. Piilo, R.N. Mantegna, Identification of clusters of investors from their real trading activity in a financial market, New J. Phys. 14 (2012) 013041.
[8] A.C. Tsoumanis, C.I. Siettos, G.V. Bafas, I.G. Kevrekidis, Equation-free multiscale computations in social networks: From agent-based modeling to coarse-grained stability and bifurcation analysis, Int. J. Bifurc. Chaos 20 (11) (2010) 3673–3688.
[9] S. Li, J.-Z. Li, A multi-agent-based hybrid framework for international marketing planning under uncertainty, Intell. Syst. Account. Finance Manage. 16 (2009) 231–254.
[10] G.M. Caporale, A. Serguieva, H. Wu, Financial contagion: evolutionary optimization of a multinational agent-based model, Intell. Syst. Account. Finance Manage. 16 (2009) 111–125.
[11] B. Mandelbrot, The variation of certain speculative prices, J. Business 36 (1963) 394–419.
[12] B. Mandelbrot, Fractals and Scaling in Finance, Springer, Berlin, 1997.
[13] U.A. Mueller, M. Dacorogna, R.B. Olsen, O. Pictet, M. Schwarz, C. Morgenegg, Statistical study of foreign exchange rates, Empir. J. Bank. Finance 14 (1990) 1189–1208.
[14] R.N. Mantegna, H.E. Stanley, Scaling behavior in the dynamics of an economic index, Nature 376 (6535) (1995) 46–49.
[15] P. Gopikrishnan, V. Plerou, X. Gabaix, H.E. Stanley, Statistical properties of share volume traded in financial markets, Phys. Rev. E 62 (2000) 4493–4496.
[16] T. Aste, W. Shaw, T. Di Matteo, Correlation structure and dynamics in volatile markets, New J. Phys. 12 (2010) 085009.
[17] P.C. Ivanov, A. Yuen, B. Podobnik, Y. Lee, Common scaling patterns in intertrade times of US stocks, Phys. Rev. E 69 (2004) 056107.
[18] S. Cincotti, L. Ponta, M. Raberto, E. Scalas, Poisson-process generalization for the trading waiting-time distribution in a double-auction mechanism, in: D. Abbott, J.P. Bouchard, X. Gabaix, J.L. McCaulty (Eds.), Noise and Fluctuations in Econophysics and Finance, in: Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 5848, Spie-Int Soc Optical Engineering, 2005, pp. 215–224, Conference on Noise and Fluctuations in Econophysics and Finance, Austin, TX, MAY 25–26, 2005.
[19] P.C. Ivanov, A. Yuen, P. Perakakis, Impact of stock market structure on intertrade time and price dynamics, Plos One 9 (2014) e92885.
[20] P.C. Ivanov, L.A. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z.R. Struzik, H.E. Stanley, Multifractality in human heartbeat dynamics, Nature 399 (6735) (1999) 461–465.
[21] P.C. Ivanov, L.A. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, H.E. Stanley, Z.R. Struzik, From $1/f$ noise to multifractal cascades in heartbeat dynamics, Chaos 11 (3) (2001) 641–652.

[22] J.P. Bouchaud, M. Potters, M. Meyer, Apparent multifractality in financial time series, Eur. Phys. J. B 13 (3) (2000) 595–599.
[23] R.G. Palmer, W.B. Arthur, J.H. Holland, B.D. LeBaron, P. Tayler, Artificial economic life: A simple model of a stock market, Physica D 75 (1994) 264–274.
[24] W.B. Arthur, J.H. Holland, B.D. LeBaron, R. Palmer, P. Tayler, Asset pricing under endogeneous expectations in an artificial stock market, in: W. Arthur, S. Durlauf, D. Lane (Eds.), The Economy as an Evolving Complex System II, in: SFI Studies in the Sciences of Complexity, Addison Wesley Longman, Santa Fe Institute, 1997.
[25] B.D. LeBaron, Time series properties of artificial stock market, J. Econ. Dyn. Control 23 (9–10) (1999) 1487–1516.
[26] M. Goykhman, Wealth dynamics in a sentiment-driven market, Physica A 488 (2017) 132–148.
[27] M. Raberto, S. Cincotti, S.M. Focardi, M. Marchesi, Agent-based simulation of a financial market, Physica A 219 (2001) 319–327.
[28] M. Raberto, S. Cincotti, S.M. Focardi, M. Marchesi, Traders' long-run wealth in an artificial financial market, Comput. Econ. 22 (2–3) (2003) 255–272.
[29] M. Marchesi, S. Cincotti, S.M. Focardi, M. Raberto, in: R. Cowan, N. Jonard (Eds.), The Genoa Artificial Stock Market: Microstructure and Simulation, in: Lecture Notes in Economics and Mathematical Systems, vol. 521, Springer, 2003, pp. 277–289.
[30] S. Cincotti, S.M. Focardi, M. Marchesi, M. Raberto, Who wins? study of long-run trader survival in an artificial stock market, Physica A 324 (1–2) (2003) 227–233.
[31] M. Raberto, S. Cincotti, Modeling and simulation of a double auction artificial financial market, Physica A 355 (2005) 35–45.
[32] S. Pastore, L. Ponta, S. Cincotti, Heterogeneous information-based artificial stock market, New J. Phys. 12 (2010) 053035.
[33] L. Ponta, S. Pastore, S. Cincotti, Information-based multi-assets artificial stock market with heterogeneous agents, Nonlinear Anal. RWA 12 (2011) 1235.
[34] L. Ponta, M. Raberto, S. Cincotti, A multi-assets artificial stock market with zero-intelligence traders, Europhys. Lett. 93 (2011) 29002.
[35] L. Ponta, E. Scalas, M. Raberto, S. Cincotti, Statistical analysis and agent-based microstructuremodeling of high frequency financial trading, IEEE J. Sel. Top. Sign. Proces. 6 (2012) 381–387.
[36] S. Cincotti, S.M. Focardi, L. Ponta, M. Raberto, E. Scalas, The waiting-time distribution of trading activity in a double auction artificial financial market, in: The Complex Network of Economic Interactions: Essay in Agent-Based Economics and Econophysics, Vol. 567, 2006, pp. 239–247.
[37] M. Raberto, S. Cincotti, C. Dose, S.M. Focardi, M. Marchesi, Price formation in an artificial market: Limit Order book versus matching of supply and demand, in: Nonlinear Dynamics and Heteroenous Interacting Agents, Vol. 550, 2005, pp. 305–315.
[38] E. Scalas, S. Cincotti, C. Dose, M. Raberto, Fraudolent agents in a artificial financial market, in: Nonlinear Dynamics and Heteroenous Interacting Agents, Vol. 550, 2005, pp. 317–326.
[39] M. Levy, H. Levy, S. Solomon, Microscopic Simulation of Financial Markets, Academic Press, New York, 2000.
[40] B.D. LeBaron, Agent-based computational finance: Suggested readings and early research, J. Econom. Dynam. Control 24 (5–7) (2000) 679–702.
[41] L. Matassini, On financial market trading, Physica A 289 (2001) 526–542.
[42] C. Chiarella, G. Iori, A simulation analysis of the microstructure of double auction markets, Quant. Finance 2 (5) (2001) 346–353.
[43] M. Raberto, S. Cincotti, S.M. Focardi, M. Marchesi, Traders' long-run wealth in an artificial financial market, Comput. Econ. 22 (2–3) (2003) 255–272.
[44] B. Podobnik, J. Shao, D. Njavro, P.C. Ivanov, H.E. Stanley, Influence of corruption on economic growth rate and foreign investment, Eur. Phys. J. B 63 (4) (2008) 547–550.
[45] J. Shao, P.C. Ivanov, B. Podobnik, H.E. Stanley, Quantitative relations between corruption and economic factors, Eur. Phys. J. B 56 (2) (2007) 157–166.
[46] C.H. Hommes, Heterogenous agent models in economics and finance, in: Leigh Tesfatsion, Kenneth L. Judd (Eds.), Handbook of Computational Economics, 2006.
[47] E. Samanidou, E. Zschischang, D. Stauffer, T. Lux, Agent-based models of financial markets, Rep. Progr. Phys. 70 (2007) 409–450.
[48] L. Laloux, P. Cizeau, J.P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, Phys. Rev. Lett. 83 (1999) 1467–1470.
[49] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, Universal and nonuniversal properties of cross correlations in financial time series, Phys. Rev. Lett. 83 (7) (1999) 1471–1474.
[50] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, Phys. Rev. E 65 (6) (2002) 066126-1:18.
[51] J.H. Stock, M.W. Watson, Tesing for common trends, J. Amer. Statist. Assoc. 83 (404) (1988) 1097–1107.
[52] N.E. Romero, Q.D.Y. Ma, L.S. Liebovitch, C.T. Brown, P.C. Ivanov, Correlated walks down the Babylonian markets, Europhys. Lett. 90 (2010) 18004.
[53] W.F. Sharpe, Capital Asset Prices: a theory of market equilibrium under conditions of risk, J. Finance 19 (1964) 425–442.
[54] R.C. Merton, An intertemporal capital asset pricing model, Econometrica 41 (1973) 867–887.
[55] S.A. Ross, The arbitrage theory of capital asset pricing, J. Econom. Theory (1976) 343–362.
[56] M. Forni, M. Hallin, M. Lippi, L. Reichlin, The generalized dynamic factor model consistency and rates, J. Econometrics 119 (2004) 231–255.