Agent-based computational model for crude oil futures market

João Miguel Abrantes
Under supervision of Prof. Tiago Domingos, Prof. Carlos Silva and Eng. João Magalhães
Environment and Energy Scientific Area, IN+, Instituto Superior Técnico (IST), Lisboa

The sharp rise in crude oil prices over the last decade has reinforced the interest of the scientific community in understanding the major causes of price movements. In this sense, the increasing "financialization" of commodity markets as well as the existence of speculators in the crude oil market can be two possible reasons for high price volatility and financial instability. The purpose of this work was to develop and test an agent-based computational model for the crude oil futures market, which simulates the crude oil price evolution through speculative behaviour of the market participants. This speculative component was modelled by the interaction of heterogeneous agents (fundamentalists, chartists, contrarians and noise traders) with learning ability, which adapt their beliefs and change their investment strategies over time according to their performance, measured by the number of winning trades. The results revealed that the intermittent behaviour that characterizes the oscillations of agents' strategies as well as the non stationarity of market activity are crucial to the emergence of stylized facts, namely the absence of autocorrelations in returns, heavy-tailed distribution of returns and volatility clustering. This model is presented as a starting point to further research about the key factors that replicate the oil price as well as the importance of speculation on its formation.

Keywords: Agent-based model, Behavioural Finance, Computational Finance, Crude oil price, Heterogeneous agents.

I. INTRODUCTION

Due to the importance of oil in the World Economy, its price fluctuations may have important macroeconomic origins and consequences[1,2]. On the one hand, we can distinguish price effects on both the supply and the demand side. First, changes in the supply are associated, for example, to the behaviour of OPEC and Non-OPEC countries, and to production costs involved in exploration, development and technological innovation[3,4]. Secondly, the past years have shown the combination of strong per capita income growth, rapid industrialization and population growth in countries like China and India. In fact, besides the large amount of oil that the Western World is still consuming, this emergent behaviour is responsible for the recent strong growth in oil demand[5–8].

On the other hand, many financial layers have emerged around crude oil benchmarks and the speculative demand represents one of the factors that possibly influence the oil price determination[9–12]. In particular, our agent-based computational model (ABM) is applied on the crude oil futures market context and reproduces the behaviour of financial agents from a speculative standpoint. In fact, the development of computer systems allows a better description of financial markets as interacting groups of rational agents with dynamic heterogeneity represented by their distribution of wealth, different investment strategies or risk aversion parameters[13]. Also, the availability and quality of current financial data gives the possibility to accurately detail the common characteristics of all financial markets such as: absence of autocorrelations in returns, heavy-tailed distribution of returns and volatility clustering[14–18]. These characteristics, known as "stylized facts", are useful to calibrate and validate our model.

This work aims to develop a model, which simulates the different agents as well as the crude oil price evolution consistent with a set of financial stylized facts. In this sense, we intend to answer the question: Is it possible to model the crude oil futures financial market, according to the observable behaviour of its agents?

In the literature on ABMs, it is possible to identify the two main modelling elements of an ABM: heterogeneity and rationality (learning)[19–23]. These elements allow highlighting the different perspectives of the various ABMs according to their realism and analytical tractability which in turn can be classified as either analytically or computationally oriented in general. This classification is justified by the existence of models with a significant number of groups of agents and complex learning mechanisms that are difficult to treat analytically.

The computational model, developed in C++ programming language with object oriented technology, is designed to be as simple as possible but still able to reproduce the stylized facts of crude oil prices. In this sense, our ABM is in line with the original idea of the Santa Fe Institute (SFI) artificial stock market[24], with the realism of Lux and Marchesi[25,26] introduced by the interaction between chartists and fundamentalists, as well as with the evolutionary dynamics of Brock and Hommes[27] to describe endogenous selection of expectations.

In fact, through the application of agent-based modelling, it is possible to design a model which simulates the speculative behaviour of heterogeneous agents (fundamentalists and chartists), who adapt their beliefs and change their investment strategy over time. This heterogeneity is motivated by the empirical evidence of the behaviour of financial agents collected from different kinds of surveys, such as questionnaires and telephone interviews. Fundamentalists base their expectations on economic theory relying on the perceived evolution of market fundamentals. On the other hand, chartists or technicians base their expectations on observed historical price patterns. They extrapolate information from previous prices, expecting trends to continue in the same direction[28,29].

In this model, we introduce the learning ability of financial agents using a form of evolutionary dynamics in which agents adapt their beliefs and change their business strategy over time according to some measure of performance. This simple evolutionary property provides the variation of the market fraction allocated to each type of strategy, allowing to investigate the influence of heterogeneous agents on the crude oil price dynamics. Finally, the application of this computational-oriented modelling in the crude oil futures market is noteworthy and innovative creating a new middle term approach that give a clearer understanding of the importance of speculation on crude oil price formation.

This paper is structured as follows. Section II presents the developed ABM. In section III, we present the whole process of choosing the model parameters as well as the simulation results. The results are evaluated according to the emergence of stylized facts in section IV. In addition, we intend to give some explanations about the origin of the stylized facts. Finally, section V provides some closing outlines and shares some further ideas for future research.

II. MODEL DESCRIPTION

This section presents the simple and heterogeneous ABM that will be used to evaluate the effect of heterogeneous expectations on crude oil prices. In this sense, we focus our attention on the explanation of all the features of our model from the macroeconomic environment to the structure of the market, through the characterization of the individual behaviour of financial agents and their strategies.

Looking at the model from a simple perspective, each agent has a set of strategies resulting from technical or fundamental analysis, each of which converts some economic information into a "buy/sell/inactive" decision. The strategy chosen at time t by a given agent is the one, which would have the best performances in a recent past. Finally, each agent decides the number of investment positions and submits orders to buy or sell futures contracts. After all orders are summed and requests are satisfied, the crude oil price $P(t)$ increases (decreases) whether there is an excess of demand (supply).

A. Economic Information

The ABM is initialized by a sequence of initial economic time series, whose length τ is long enough to allow any forecasting strategy which converts some economic information into a decision $s_i(t)$. Furthermore, the performance measures $U_h(t-1,..,t-m)$ of each strategy are only updated for $t > \tau + m$. Until this period, the association between an agent and a strategy is completely random.

We assume that the simulated crude oil price time series $P(t)$ is initialized by an autoregressive process of order 2 $AR(2)$ given by

$$P(t) = \begin{cases} P_0, & \text{if } t = 0 \\ A_1 P_0 + \epsilon(t), & \text{if } t = 1 \\ A_1 P(t-1) + A_2 P(t-2) + \epsilon(t), & \text{if } t > 1, \end{cases} \tag{1}$$

where P_0 is the initial value, A_1 and A_2 are the autoregressive parameters and $\epsilon(t) \approx N(0,1)$ is a noise term given by the standard normal distribution. The GDP series is generated through a correlation c_{OIL} with the crude oil price

$$\text{GDP}(t) = \begin{cases} \text{GDP}_0, & \text{if } t = 0 \\ \text{GDP}_0 + \epsilon(t), & \text{if } t = 1 \\ \text{GDP}(t-1) + c_{OIL}(P(t-1) - P(t-2)) + \epsilon(t), & \text{if } t > 1. \end{cases} \tag{2}$$

The choice of GDP as dependent variable is based on the fact that nine out of ten of the United States recessions since World War II were preceded by a substantial increase in oil prices[30].

Note that these parameters are chosen in order to generate a fictitious evolution of GDP and crude oil price. The fundamental series of GDP only serve as an external source that supports the decisions of fundamentalists. Since it is not possible to obtain the daily periodicity of GDP, we seek to generate certain time-varying processes to ensure the functioning of the fundamental GDP strategy.

B. Strategies

We assume throughout this model that there are only four types of traders: fundamentalists, chartists, contrarian chartists and noise traders, who in fact are the most widespread types of traders used in ABMs. We distinguish between two types of trading rules: the inserted under the technical analysis or under the

TABLE I: Agents investment strategies.

Type	Strategy
Chartist	Moving Average Crossover
	Exponentially Weighted Moving Average Crossover
	Mean Reverting
	Exponentially Weighted Mean Reverting
Contrarian	Moving Average Crossover
	Exponentially Weighted Moving Average Crossover
	Mean Reverting
	Exponentially Weighted Mean Reverting
Fundamentalist	GDP Forecast
Noise Trader	Random

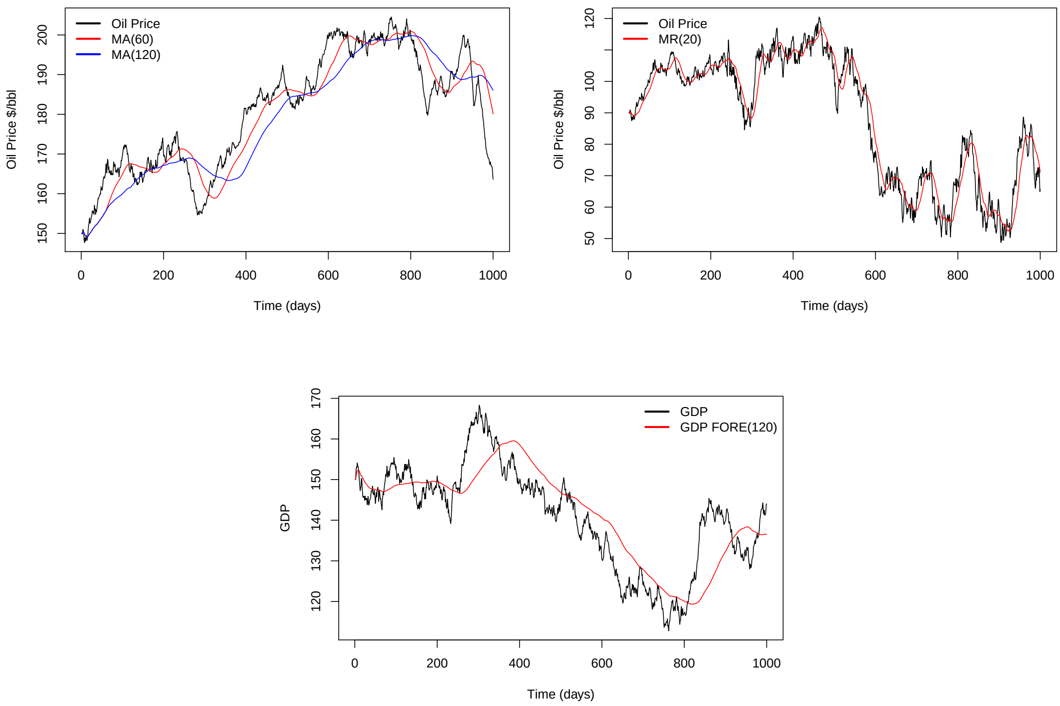

FIG. 1: Example of the evolution of crude oil prices and application of technical trading rules moving average crossover (left), mean reverting moving average (right) and fundamental GDP forecasting rule (bottom).

fundamental analysis. The technical analysis only focuses on statistical analysis of price series while fundamental analysis is essentially based on understanding how economic variables can affect and change asset prices. Therefore, beyond the chartists and the fundamentalists, directly associated with each of these philosophies, there is also the participation of contrarian chartists and noise traders. The contrarian chartists or simply contrarians use technical analysis tools, considering that the price trend starts to reverse. On the other hand, the noise traders are added to the system in order to represent the irrational investor, who randomly selects one strategy.

First, there are several techniques that can be used for technical analysis[31]. In our model there are two indicators based on moving averages to predict the crude oil prices. The first indicator is a moving average crossover and provides a buy (sell) signal when the moving average for a shorter period ST is higher (lower) than the moving average for a longer period LT,

$$s_{MA_t^{ST,LT}} = \begin{cases} 1, & \text{if } \frac{\frac{1}{ST}\sum_{i=0}^{ST-1} P(t-i) - \frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)}{P(t)} > \phi \\ 0, & \text{if } \left|\frac{\frac{1}{ST}\sum_{i=0}^{ST-1} P(t-i) - \frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)}{P(t)}\right| \leq \phi \\ -1, & \text{if } \frac{\frac{1}{ST}\sum_{i=0}^{ST-1} P(t-i) - \frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)}{P(t)} < -\phi. \end{cases} \tag{3}$$

The second indicator is a mean-reverting moving average that checks whether the current oil price moves away from its long term average. In other words, the buy (sell) signal is activated when the oil price is below (above) a moving average for a long period LT,

$$s_{MR_t^{LT}} = \begin{cases} 1, & \text{if } \frac{\frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)-P(t)}{P(t)} > \phi \\ 0, & \text{if } \left|\frac{\frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)-P(t)}{P(t)}\right| \leq \phi \\ -1, & \text{if } \frac{\frac{1}{LT}\sum_{i=0}^{LT-1} P(t-i)-P(t)}{P(t)} < -\phi. \end{cases} \quad (4)$$

Additionally, we also include the exponential weighted "versions" of these two techniques, which basically represent the investors who underweight long term averages, and tend to attach too much weight to recent experiences.

In contrast, the fundamental analysis is based on the study of other available data series that might influence the price of financial assets. The construction of these trading rules is based on a two step procedure. First, each trading rule provides a forecast of the fundamental variable for a given time window LT_f through an autoregressive process $AR(2)$ based on historical data. Secondly, each rule returns a signal to buy or sell futures contracts, comparing the prediction obtained $Y(t)$ in the first step with the current value $X(t)$. In the model, the output signal is associated with the analysis of the factors that influence the evolution of oil prices. For example, the trading rules based on GDP, will provide a signal to buy (sell) when the predicted value for the key variable is below (above) the current value, since a co-movement is observed between GDP and crude oil prices,

$$s_{\text{GDP}_t^{LT_f}} = \begin{cases} 1, & \text{if } \frac{Y(t)^{LT_f}-X(t)}{X(t)} > \phi \\ 0, & \text{if } \left|\frac{Y(t)^{LT_f}-X(t)}{X(t)}\right| \leq \phi \\ -1, & \text{if } \frac{Y(t)^{LT_f}-X(t)}{X(t)} < -\phi. \end{cases} \quad (5)$$

Finally, it should also be noted that it is possible to instantiate additional trading rules changing only two parameters: the temporal dimension of the window and the maximum deviation ϕ from which rules are activated. An example of the trading rules used in the model is presented in figure 1.

According to the strategies previously discussed, agents are randomly distributed to the ten groups ($H = 10$) of traders described in table I.

After evaluating the performance of each strategy, each agent randomly selects one or more time periods of the strategy previously chosen. Consequently, this choice is converted into a trading signal $s_i(t)$, reflecting a buy (1), sell (-1) or refrain (0) decision. In particular, if the selected number of time periods R exceeds one, the trading signal $s_i(t)$ will reflect the most frequent result. For example, if an agent chooses $R = 2$ periods ($s_{\text{GDP}^{20}} = 1$, $s_{\text{GDP}^{40}} = 1$) of a fundamental GDP strategy, one will take a buy (1) decision.

C. Agents

In terms of wealth, each agent i, at each time step, starts with an initial portfolio defined by a distribution that follows a Pareto law with exponent κ,

$$W_i(t) = W_0 U[0,1]^{-1/\kappa} \quad \kappa > 0, \quad (6)$$

where W_0 is a constant measured in monetary units (MU) and κ is an appropriate positive constant[32,33]. U represents the continuous uniform distribution with a finite interval $[0, 1]$.

In the model, we consider the market activity and the trading volume as the consequence of the fact that an agent can reduce his investment due to risk aversion and crude oil price volatility $\sigma(t)$ calculated by the exponentially weighted moving average of the standard deviation from the crude oil price, similar to the approach introduced by JPMorgans RiskMetrics[34]

$$\sigma(t) = \begin{cases} \sqrt{\theta(t)^2}, & \text{if } t = 1 \\ \sqrt{w\sigma^2(t-1) + (1-w)\theta(t)^2}, & \text{if } t > 1, \end{cases} \quad (7)$$

where w is the smoothing constant and $\theta(t)$ corresponds to the absolute crude oil price change between t and $t-1$.

Volatility is a measure that evaluates the price variations of a financial series over time. In this sense, from the investor's point of view, is a variable that reflects the uncertainty of their investments and agents will be discouraged from investing in periods of low fluctuations. In terms of demand specification, each agent attempts at each period to optimize his allocation according to a constant relative risk aversion (CRRA) preference, since increases linearly with agents wealth[35]. In this assumption of investment allocation, agents affect market price proportionally to their relative wealth. In this sense, we assume that investors make investment decisions based on the level of their personal wealth,

$$F_i^{+/-}(t) = \begin{cases} \frac{W_i s_i}{S_F \lambda_i \sigma}, & \text{if } \lambda_i \sigma > 1 \\ \frac{W_i s_i}{S_F}, & \text{if } \lambda_i \sigma \leq 1, \end{cases} \quad (8)$$

where the number of buy (sell) positions in futures contracts $F_i^+(t)$ ($F_i^-(t)$) of agent i depends on the agent's risk aversion λ_i obtained through a continuous uniform distribution with a finite interval $[\lambda_{\min}, \lambda_{\max}]$.

The size of each standardized contract S_F, measured in monetary units, is fixed. In this context, it is important to note that the agent's risk aversion and the price volatility define the proportion of the agent's wealth invested in crude oil futures contracts $1/\lambda_i\sigma$.

Note that a futures contract is only traded if there is always someone on each side. In fact, the effective total long interest $\hat{F_i}^+(t)$ (buy side) in futures contracts must equal the effective total short interest $\hat{F_i}^-(t)$(sell side). This is possible through the mechanism of the SFI artificial stock market[24], where the smaller positions of these two sets of buyers and sellers is satisfied while the other is partitioned according to

$$\hat{F_i}^{+/-}(t) = \begin{cases} F_i^+(t)\frac{S(t)}{D(t)}, & \text{if } D(t) > S(t) \\ F_i^-(t)\frac{D(t)}{S(t)}, & \text{if } D(t) < S(t), \end{cases} \tag{9}$$

in which demand $D(t) = \sum_{i=1}^N F_i^+(t)$ and supply $S(t) = \sum_{i=1}^N F_i^-(t)$.

In our model, the evolutionary selection between different strategies is built upon the adaptive belief system of Brock and Hommes[27], but with memory in the performance measure and asynchronous strategy updating. In this sense, each agent makes a decision taking into account its behavioural features, as well as previous performances of each trading strategy. The chosen metric to evaluate the performance of each strategy is the number of successes in each negotiation session. In fact, studies about the profiles and motivations of habitual speculators in commodity futures markets and survey responses indicate that the investors " ... are not trading solely or even primarily for profit, but may be maximizing excitement or the number of winning trades"[36]. As such, the performance measure of a strategy in a given period is based on the fraction of agents who did not lose money and is given by

$$U_{h,i}(t) = (1-\eta_i)\frac{\text{Successes}_h(t-1)}{N_h(t-1)} + \eta_i\frac{\text{Successes}_h(t-2)}{N_h(t-2)}, \tag{10}$$

where the parameter N_h is the number of agents associated with each strategy. Successes_h is the number of agents who did not lose money in each strategy and $0 \leq \eta_i \leq 1$ represents the memory, measuring the relative weight agents give to past successes of each strategy. In the particular case $\eta_i = 0$, the performance of each strategy is completely determined by the most recent number of successes.

Given the performance measure, the probability of each agent following a particular trading strategy takes the form

$$\text{Prob}_{h,i}(t) = (1-\gamma_i)\frac{e^{\beta U_{h,i}(t)}}{\sum_{h=1}^H e^{\beta U_{h,i}(t)}} + \gamma_i\frac{e^{\beta U_{h,i}(t-1)}}{\sum_{h=1}^H e^{\beta U_{h,i}(t-1)}}, \tag{11}$$

where $0 \leq \gamma_i \leq 1$ represents the inertia, reflecting the fact that not all the agents update the strategy performance in every negotiation session. Indeed, the higher the parameter intensity of choice β, the faster individuals will switch to more successful strategies. Note that the behavioural parameters memory η_i and inertia γ_i are obtained through a continuous uniform distribution with a finite interval $[\eta_{\min}, \eta_{\max}]$ and $[\gamma_{\min}, \gamma_{\max}]$.

D. Market Structure

Our model environment includes a financial market where agents work on "mark-to-market basis" and use their trading rules to buy or sell futures oil contracts at an announced price. For simplicity, we assume that the maturity of contracts has a single term (one day).

The feedback between prices and agents interactions is performed by a Walrasian mechanism of price update, which is popular in several presented ABMs. In fact, at each time step, the crude oil price changes are a function of excess demand/supply - the difference in the number of buy and sell trades each negotiation day

$$P(t+1) = P(t) + \alpha\frac{D(t)-S(t)}{D(t)+S(t)}, \tag{12}$$

where α is a a positive daily price adjustment parameter. The excess demand/supply moves the price up or down, where the largest return occurs when all traders act in unison, i.e., they all either buy or sell their stocks. $P(t+1)$ is the main output of our ABM. In fact, this value updates the price time series, as well as the computation of market volatility. Finally, we have all the ingredients to compute the profit/losses of each agent

$$\Delta W_i = [P(t+1) - P(t)]\, S_F \hat{F}_i^{+/-}(t). \tag{13}$$

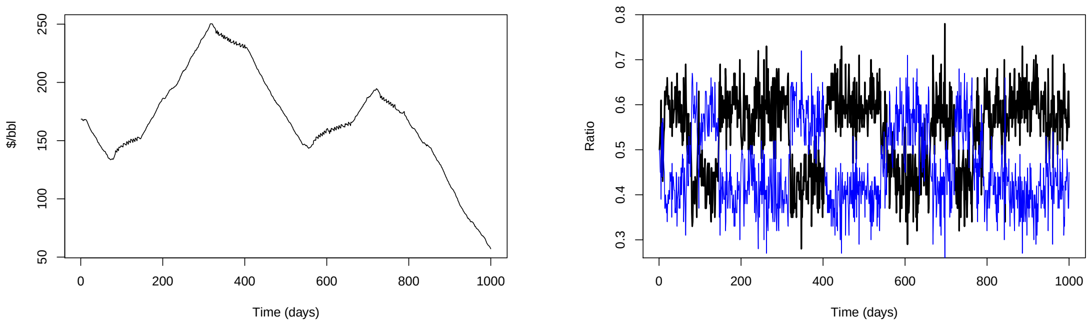

FIG. 2: Simple simulation of 100 market participants with only two available strategies ($H = 2$) during a period of 1000 trading days: moving average crossover chartists (black) vs mean reverting chartists (blue). $\beta = 0.4$.

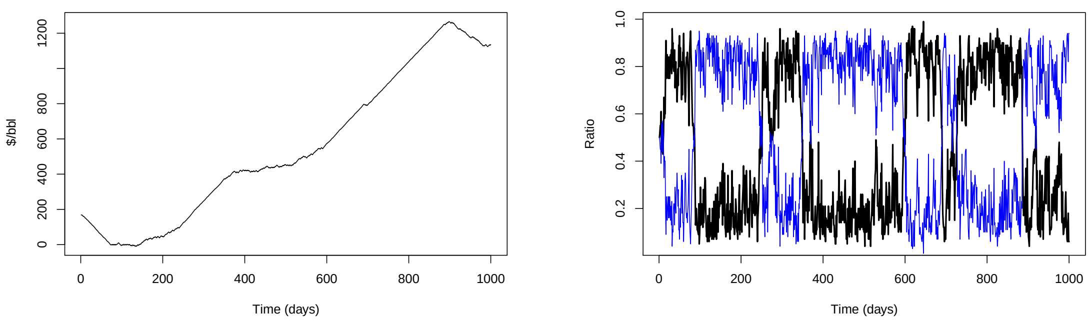

FIG. 3: Simple simulation of 100 market participants with only two available strategies ($H = 2$) during a period of 1000 trading days: moving average crossover chartist (black) vs GDP fundamentalists (blue). $\beta = 6$.

III. SIMULATION TEST

In this section we firstly define the ABM parameters and discuss their influence on the stability of the system. Then, we present all details about the simulation results, including the price dynamics and some market indicators.

A. Simulation Settings

Under initial conditions and parameters listed in table II, the intensity of choice β, the daily price adjustment parameter α and the initial number of agents $N(0)$ have considerable importance to obtain a good representation of financial market dynamics. In this sense, the first step is to fix these parameters in order to control the stability of the system as well as the emergence of the stylized facts.

First of all, the intermittent behaviour that characterizes the oscillations of agents strategies is crucial. In fact, through the intensity of choice parameter, it is necessary to find the "right level" of intermittency between the market fractions of each strategy with the purpose of preventing strategy homogeneity, which therefore creates an uniform tendency to rise or fall in crude oil prices. In addition, the heterogeneity of strategy output signals (buy/sell) is also controlled by the parameter R. Throughout the simulations, we consider that the selected number of time periods in each trading rule is fixed to $R = 7$. This choice ensures the credibility of the signal of each trading rule, since it is assessed taking into account the individual signals of all possible periods.

As we can see in the simulations of the figures 2 and 3, the fraction of each type of strategy is very important in market dynamics. The choice of β in each simulation is intended to accentuate the dominance of each strategy. According to the dynamics given by the interaction of two types of investors, the moving average crossover strategy has a destabilizing effect on the price of oil, since it extracts price movements from the past and predicts that they continue moving in the same direction. In particular, the moving averages are a source of market instability and can lead to the tendency for the market price to take long excursions away[37]. On the other hand, the mean reverting strategy has a stabilizing effect on the price of oil. In fact, mean reverting chartists predict that the oil price converges on its long term average. Considering this simplest market configuration, we can fix β in such a way that even a single heterogeneity can lead to an interesting dynamics.

On the other hand, the daily price adjustment parameter α amplifies or diminishes the price fluctuations and, more importantly, ensures that the bandwidth for returns over unit time steps should roughly conform to what one usually observes with data from crude oil financial market.

Finally, there is a decrease in volatility with the in-

TABLE II: Summary of the parameter set.

Symbols	Parameter Set
τ	240
m	3
$(P_0, A_1, A_2)_{OIL}$	(90 \$/bbl, 0.03490, 0.00868)
$(\mathrm{GDP}_0, c_{OIL})$	(90 T\$, -0.04 T\$/\$/bbl)
H	10
$N(0)$	100
α	8.5 \$/bbl
W_0	10000 MU
κ	2
w	0.94
β	0.2
$U[\lambda_{\min}, \lambda_{\max}]$	U[1, 10]
$U[\eta_{\min}, \eta_{\max}]$	U[0, 1]
$U[\gamma_{\min}, \gamma_{\max}]$	U[0, 1]
(ST, LT)	$\{(1, 240); (2, 120); (4, 9); (5, 20); (10, 20); (20, 40); (9, 18)\}$
(LT)	$\{5, 10, 20, 40, 60, 120, 240\}$
(LT_f)	$\{5, 10, 20, 40, 60, 120, 240\}$
S_F	1000 MU
R	7
ϕ	0

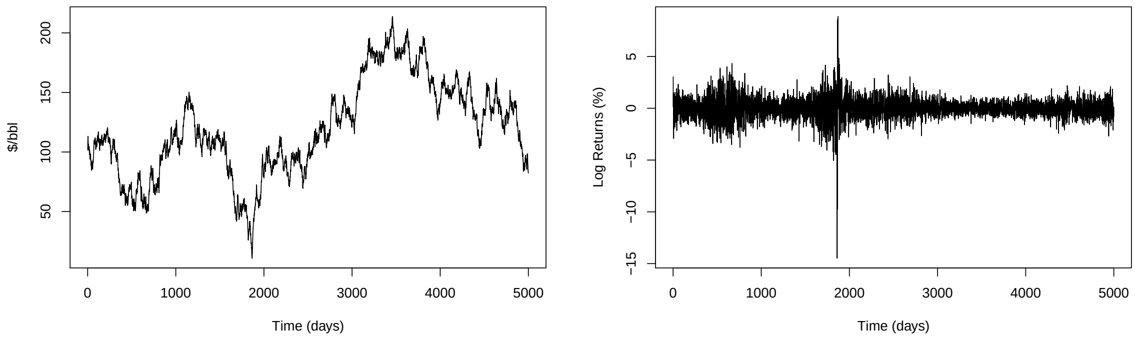

FIG. 4: Simulated crude oil price path (left) and logarithmic returns (right).

creasing initial number of agents N. While a low value of N produces many fluctuations, an higher value of N will prevent the formation of market bubbles and crashes. In this sense, it is important to obtain a range of volatility values that represents the different market behaviours, including calm periods as well as a greater price fluctuation frequency.

In sum, it is necessary to find a combination between these parameters, allowing the emergence of stylized facts of financial markets.

B. Simulation Results

Here we present the detailed analysis of the market activity dynamics according to the parameters listed in table II. We simulate a market with an initial population of 100 agents divided by 10 strategies for 5000 trading days, which corresponds roughly to 20 years of market activity.

Figure 4 shows the price path and the logarithmic returns from a typical simulation

$$r(t) = \frac{P(t) - P(t-1)}{P(t-1)} \approx \ln P(t) - \ln P(t-1). \quad (14)$$

The model simulation presents long calm periods interrupted by sudden bursts of clustered volatility. Moreover, we can evidence the existence of financial bubbles and crashes in simulated crude oil price series.

Figure 5 presents in more detail the internal dynamics involved in a trading period of 5000 trading days. Considering the evolution of market indicators throughout the simulation, we first check the fluctuations of market activity. Note that the fluctuation frequency of market activity increases with an higher volatility range of values. In this case, the combination between risk aversion and high market volatility induces neutral investment positions (see equation 8). This fact influences the amount of traded contracts. The greatest value of volatility induces an higher percentage of agents who do not take any investment position, reflected in refrain fraction. In terms of economic fundamental series, GDP reflects the relation with crude oil price imposed in our model (see equa-

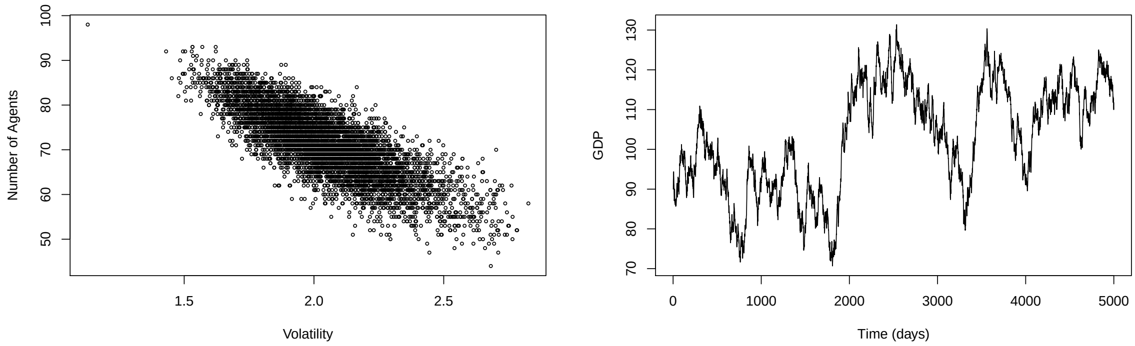

FIG. 5: Simulated market indicators: market activity vs daily volatility (left); GDP time series (right).

tion 2).

IV. STYLIZED FACTS

Now we take a closer look at the statistical characteristics of the simulated data set. In particular, we investigate whether or how far the data from our artificial crude oil financial market conform to the stylized facts of crude oil real-life market and to the various financial daily time series[18]. We focus on evaluation of three stylized facts such as heavy-tailed distribution of returns, absence of autocorrelation in returns and volatility clustering. The empirical data used in this section consists of daily WTI crude oil spot prices from 10/06/1994 to 02/01/2014 (5000 trading days) obtained from U.S. Energy Information Administration.

Figure 6 shows the complementary cumulative distribution function of normalized absolute logarithmic returns $|R(t)|$ for empirical and simulated data. For comparison, the solid line represents the complementary cumulative distribution of the standard normal distribution $N(0,1)$. For the two results, it is possible to evidence a clear deviation from Gaussian behaviour with approximate by power law scaling in the tail of the empirical and simulated situations. A log-log regression that satisfy the condition $|R(t)| > 2$ gives the slope $B = 3.25 \pm 0.04$ ($B = 3.53 \pm 0.03$) for empirical (simulated) data. Despite the different values between the empirical and simulated data, the simulated process is able to generate a distribution of returns with a shape, which is markedly different from the Gaussian case, where the probability for large positive or negative fluctuations is larger than for a normal distribution. Moreover, the exponent of the power law fit for the simulated data is in agreement with values found in various financial daily time series, since these fat tails have been fitted in various ways and can be approximated by a power law with an exponent ranging from 2 to 5[18].

In figure 7, we present the autocorrelation function of the absolute returns and raw returns at different time lags. While the autocorrelation function of absolute returns decays as a function of the time lag, the autocorrelation of raw returns decays immediately to 0. In this sense, the absence of significant linear correlations in asset prices means that the sign of the next price fluctuation is unpredictable on average. Therefore, this widely accepted stylized feature of daily returns, shows that each investment always has an associated risk and the market itself has the ability to self-regulate in order to avoid situations of arbitrage opportunity.

The volatility clustering is usually quantified by the autocorrelation function of the absolute returns. Various empirical studies[38,39] indicate that this function shows a slow decay, although still significantly positive. In figure 8, the autocorrelation of the absolute value of returns shows the presence of long-range correlations. However, despite both indicate the presence of volatility clustering, the correlation period of empirical data is higher than simulated data. In fact, the present model can not fully replicate more frequently the situations where large changes tend to be followed by large changes and small changes tend to be followed by small changes. Finally, a power law usually fits the autocorrelation of the absolute value of returns with exponent ranging from about zero to 1. The value found in the power law fit is $B = 0.54 \pm 0.04$ ($B = 0.62 \pm 0.02$) for the empirical (simulated) data.

In sum, the simulated price process exhibits the main features of real markets such as fat tails for the returns distribution, zero autocorrelation for the returns and slow decay of the autocorrelation function of the absolute values of returns. As such, the oscillations of agents strategies, the fluctuations of market activity as well as the intermittent behaviour between high and low volatility periods are crucial to generate the main empirical evidences of crude oil financial markets. In fact, we can begin by noting that while a low number of agents produces too many fluctuations (high volatility), a high number of agents will prevent the formation of financial bubbles and crashes (low volatility). In this sense, we refer the importance of market activity fluctuations to the emergence of volatility clustering. Also, the intermittent behaviour between high and low volatility periods amplifies the fat-tail phenomenon. On the other hand, the results do not show any significant correlation between returns. This fact can be explained by the amount of variables that directly or indirectly influences the

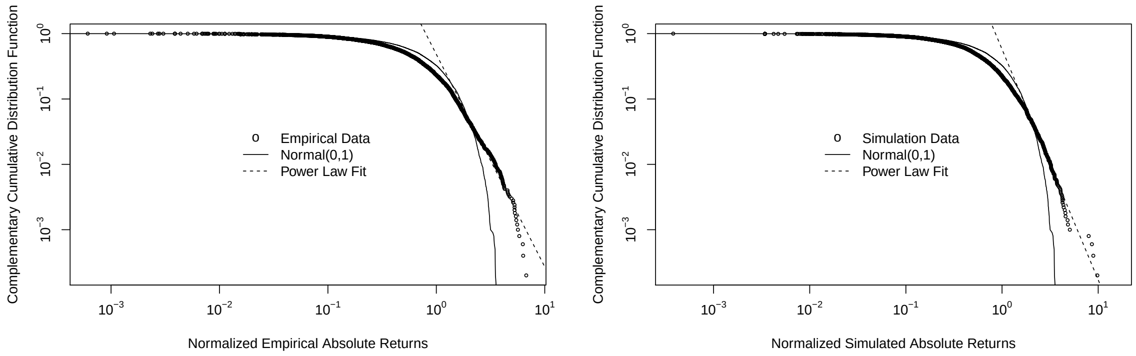

FIG. 6: Complementary cumulative distribution function (ccdf) of normalized absolute logarithmic returns $|R(t)|$ for empirical (left) and simulated (right) data. Normalized returns are computed as $R(t) = (r(t) - M)/SD$, where M and SD are the mean and the standard deviation of $r(t)$, respectively. The dots represent an estimate of the ccdf of $|R(t)|$ related to the empirical and simulated data. The solid line represents the ccdf from the standard normal distribution. The dashed line is the power law fit $P = A|R|^{-B}$ with $B = 3.25 \pm 0.04$ ($B = 3.53 \pm 0.03$) of the tail of the empirical (simulated) ccdf for $|R(t)| > 2$.

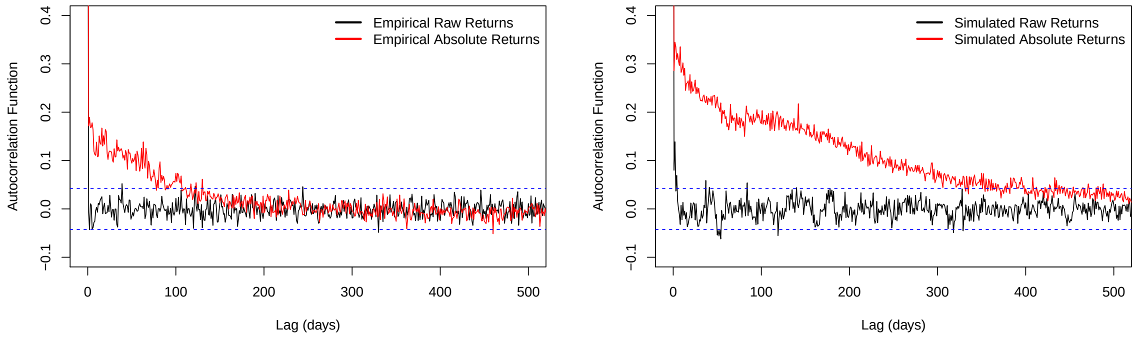

FIG. 7: Behaviour of autocorrelation functions of empirical (left) and simulated (right) raw and absolute logarithmic returns.

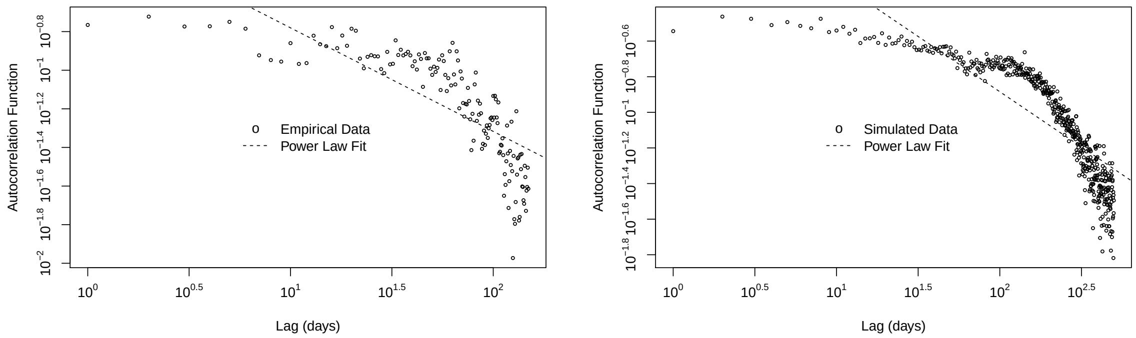

FIG. 8: Estimate of the autocorrelation function of empirical (left) and simulated (right) absolute logarithmic returns. The dots represent the autocorrelation of absolute returns $|r(t)|$. Absolute returns are fitted with a power law decay $P = A|R|^{-B}$ with $B = 0.54 \pm 0.04$ ($B = 0.62 \pm 0.02$) for the empirical (simulated) data.

crude oil price, namely the oil price volatility $\sigma(t)$, the initial number of agents $N(0)$, the number of positions in futures contracts $F_i^{+/-}(t)$ and so on. Indeed, these variables are in general not correlated with the price and lead to a decorrelation of the price increments.

V. CONCLUSIONS AND FUTURE WORK

In fact, through the application of agent-based modelling, it was possible to design a model which simulates the speculative behaviour of heterogeneous agents (fundamentalists, chartists, contrarians and noise traders), who adapt their beliefs and change their investment strategies (technical or fundamental) over time.

The results of implementing such a model, which used C++ as a programming language, were presented and discussed, providing a first look to the potential of the ABM. Thus, the main results of our model are:

- emergence of the three main stylized facts: absence of autocorrelation in returns, heavy-tailed distribution of returns and volatility clustering;
- existence of financial bubbles and crashes, which can be originated by the fluctuations of the fractions of each type of strategy;
- understanding of the origin of stylized facts with respect to the microscopic dynamics of the agents in the market;
- importance of the non stationarity of market activity.

With the present work, the main objective of reproducing the main empirical evidences of financial markets was achieved. However, the limitations of the model may be summarized to a set of vulnerabilities. First, the main vulnerability is related to the fact that results depend heavily on initial conditions of the system. Indeed, due to the computational character of the ABM in contrast to its analytical tractability, it is difficult to identify the system stability limits. In this sense, a fine adjustment of parameters was made in order to avoid the explosion of crude oil price. Moreover, we lack empirical estimates for all behavioural parameters that we introduce in the evolutionary dynamics. On the other hand, since it is not possible to obtain the daily periodicity of GDP, this macroeconomic series does not display a real progress, conditioning the credibility of fundamental strategies.

Considering now the limitations of the model design, we can clearly identify four vulnerabilities. First, due to the multitude of factors that influence the crude oil price, the Walrasian mechanism of market price determination is limited. Indeed, this mechanism does not take into account the supply and demand of real producers and consumers as well as is very dependent on the price adjustment parameter. Second, it is necessary to develop an algorithm that simulates more realistically the Financial Mathematics associated to the futures market, namely by introducing more maturity periods and initial and maintenance margin requirements. Third, agents are limited to the amount of available trading rules and there is no generation of new rules according to market trends. Lastly, there is no prospect of continuity, since a new generation of agents is created in each trading day. The lack of individual wealth update, does not give rise to bankruptcy situations.

Further developments may point to different directions but converge on the goal of having a more powerful model. In this sense, we can choose a line of research that fixes the limitations of the model. For example, we can develop new rules through genetic algorithms in which a set of trading rules would change in time. Moreover, the model may be completed by adding new technical trading strategies like geometric brownian motion and more fundamental ones based on other macroeconomic factors such as interest rates, exchange rates and prices of other commodities. Finally, we can choose a line of research that allows understanding the potential and limits to the undertaken approach, namely the quantitative treatment of other stylized facts as well as the study of the market activity from a self-organization or from statistical mechanics standpoint. In addition, these developments are only possible if there is a parallel code C++ optimization to allow a greater number of simulated series in less time.

In conclusion, the ABM developed as well as the C++ implementation of it allowed to reproduce and capture the main empirical evidences of crude oil financial market by simulating the heterogeneous behaviour of financial agents.

ACKNOWLEDGEMENT

We thank to Prof. Tânia Sousa for helpful guidance and suggestions in the preparation of this work. This work is supported by project "Energy Wars" (QREN 7929), funded by QREN/FEDER.

[1] E. Regnier, Energy Economics **29**, 405 (2007).
[2] J. D. Hamilton, *Understanding Crude Oil Prices*, Tech. Rep. (National Bureau of Economic Research, 2008).
[3] R. K. Kaufmann, Economic Modelling **12**, 165 (1995).
[4] R. K. Kaufmann, Energy Policy **39**, 105 (2011).
[5] T. Helbling, V. Mercer-Blackman, and K. Cheng, Finance and Development **45**, 10 (2008).
[6] G. Calvo, "Exploding Commodity Prices, Lax Monetary Policy, and Sovereign Wealth Funds. VoxEU. 20 June," (2008).
[7] L. Kilian, The American Economic Review **99**, 1053 (2009).
[8] J. D. Hamilton, *Causes and Consequences of the Oil Shock of 2007-08*, Tech. Rep. (National Bureau of Economic Research, 2009).
[9] B. Fattouh, *An Anatomy of the Crude Oil Pricing System* (Oxford Institute for Energy Studies, 2011).
[10] B. Fattouh, L. Kilian, and L. Mahadeva, Energy Journal **34** (2013).
[11] L. Kilian and D. P. Murphy, Journal of Applied Econometrics (2013).
[12] W.-P. Chen, T. Choudhry, and C.-C. Wu, Journal of International Money and Finance (2013).
[13] E. Bonabeau, Proceedings of the National Academy of Sciences of the United States of America **99**, 7280 (2002).
[14] T. Bollerslev, R. Y. Chou, and K. F. Kroner, Journal of Econometrics **52**, 5 (1992).
[15] W. A. Brock and P. J. De Lima, *Nonlinear Time Series, Complexity Theory, and Finance* (Social Systems Research Institute, University of Wisconsin, 1995).

[16] A. Pagan, Journal of Empirical Finance **3**, 15 (1996).
[17] J. Y. Campbell, A. W. Lo, A. C. MacKinlay, and R. F. Whitelaw, Macroeconomic Dynamics **2**, 559 (1998).
[18] R. Cont, Quantitative Finance **1**, 223 (2001).
[19] C. H. Hommes, Handbook of Computational Economics **2**, 1109 (2006).
[20] B. LeBaron, Handbook of Computational economics **2**, 1187 (2006).
[21] C. Chiarella, R. Dieci, and X.-Z. He, Handbook of Financial Markets: Dynamics and Evolution , 277 (2009).
[22] T. Lux, Handbook of Financial Markets: Dynamics and Evolution. North-Holland: Amsterdam , 161 (2009).
[23] S.-H. Chen, C.-L. Chang, and Y.-R. Du, Knowledge Engineering Review **27**, 187 (2012).
[24] W. B. Arthur, *Asset Pricing under Endogenous Expectations in an Artificial Stock Market*, Ph.D. thesis, Brunel University, London (1996).
[25] T. Lux and M. Marchesi, Nature **397**, 498 (1999).
[26] T. Lux and M. Marchesi, International Journal of Theoretical and Applied Finance **3**, 675 (2000).
[27] W. A. Brock and C. H. Hommes, Journal of Economic Dynamics and Control **22**, 1235 (1998).
[28] J. A. Frankel and K. A. Froot, The Economic Record , 24 (1986).
[29] M. P. Taylor and H. Allen, Journal of International Money and Finance **11**, 304 (1992).
[30] J. D. Hamilton, The New Palgrave Dictionary of Economics Palgrave Macmillan, London. Available online at http://www. dictionaryofeconomics. com/dictionary. Jiménez-Rodríguez, Rebeca and Marcelo Sánchez , 201 (2005).
[31] J. J. Murphy, *Technical Analysis of the Futures Markets: A Comprehensive Guide to Trading Methods and Applications, New York Institute of Finance* (Prentice-Hall, 1999).
[32] V. Pareto and C. d. Politique, Rouge, Lausanne (1897).
[33] R. Saucier, *Computer Generation of Statistical Distributions*, Tech. Rep. (DTIC Document, 2000).
[34] J. Longerstaey and M. Spencer, Morgan Guaranty Trust Company of New York: New York (1996).
[35] M. Anufriev and P. Dindo, Journal of Economic Behavior & Organization **73**, 327 (2010).
[36] W. B. Canoles, S. Thompson, S. Irwin, and V. G. France, Journal of Futures Markets **18**, 765 (1998).
[37] C. Chiarella, X.-Z. He, and C. Hommes, Journal of Economic Dynamics and Control **30**, 1729 (2006).
[38] C. Granger and Z. Ding, UCSD Department of Economics Discussion Paper , 94 (1999).
[39] B. Mandelbrot and H. M. Taylor, Operations Research , 1057 (1967).