Ph. D. Research Project

Complexity in Financial Markets: Modeling Psychological Behavior in Agent-Based Models and Order Book Models

Candidate: Matthieu Cristelli
Supervisor: Prof. Luciano Pietronero

Part I

Introduction

Physics community has shown an increasing interest in the study of Social Sciences and in particular Economics in the last 25 years [1, 2]. The growing popularity of this field arises, roughly speaking, from two main reasons.

The former one consists in the very interesting experimental evidences which are observed throughout financial markets. In fact almost all price time series of financial stocks and indexes approximately exhibit the same statistical properties (at least qualitatively). These regularities, that are very widespread among all financial markets, are usually called Stylized Facts (SF). The interest in these common features derives from the fact that the theoretical framework of classical Economics is not able to explain their origin. Therefore SF set a not trivial conceptual problem that cannot be reconciled within the simple linear cause-effect vision of standard economical approach. But why could Physics provide a new perspective to this problem? The answer can be traced back to Seventies when a new paradigm was introduced to explain the phenomenology of Critical Phenomena. This new approach finally has led to Complex System Physics. Here is the link among Physics and Economics. In fact SF are somehow similar (in the next section we will make clear in which sense this sentence must be intended) to the Physics of Complex System. Economical systems recall adaptive and evolutionary systems which are typically observed in biology. Hence Statistical Physics and Complex System Physics appear as natural candidates to give a quantitative representation and explanation of these phenomena.

The latter reason instead lies in the huge amount of empirical data which are produced by financial markets and which are now easily available. Large data availability is a crucial point because Economics and in general Social Sciences set a further conceptual problem beyond the investigation of the Natural Laws itself: the existence or not of some underlying Natural Laws to investigate. In Physics the existence of these Laws is always assumed, instead in Economics this question is as crucial as the scientific investigation itself. For instance the irrationality of human psychology is an aspects which a priori may prevent the existence of these Natural Laws. However, limiting our attention to financial economics, SF are robust on different timescales and in different stock markets. Moreover the non rational-behavior\(^1\) of investors, that we have said to be a possible element which could prevent the existence of some Laws to investigate, is itself a key element to understand the mechanisms underlying financial crashes. In fact it has recently argued [3] that the non-rational behavior of some hedge funds and small investors in surfing trends has acted as fuel for financial bubbles.

\(^1\)with respect to the economical definition of rationality
Besides in very recent years new kinds of social data have begun to be collected such as all movements of cars, the daily network of interaction of human beings by monitoring their mobile phone and many others. These new observations permit now a quantitative description of real society networks and the interplay between different levels of human activities, in theory, could play a role in global markets.

Finally in order to make a comparison with some traditional fields of Physics, a similar quantity of information is observed only in the output of a big particle accelerator such as LEP, LHC, etc.

**Stylized Facts**

Turning now our attention to the experimental evidences of financial markets, the main SF are:

- absence of simple arbitrage in financial markets, that is, given the price time series up to now, the sign of the next price variation is unpredictable

- the probability density function (pdf) of price variations (called returns) is not a Gaussian and prices do not follow a simple Random Walk. The return pdf is a function characterized by positive kurtosis (a Gaussian would be characterized by zero kurtosis) and the tail behavior of the pdf is approximately a power law with exponent ranging from 2.5 to 5. This SF is usually called Fat Tails.

- the autocorrelation function of the volatility which is defined as the absolute value or the square of the returns, is non zero and is well-described by a power law decay with exponent ranging from −1 to 0. Therefore the process defined by the return series is uncorrelated, i.e. price variations are still unpredictable, but returns are not independent. This means that big fluctuations are more likely followed by big fluctuations and vice-versa. Therefore the volatility results to be clustered over very long time periods, even several months.

Beyond these SF we can state other relevant effects which are widespread in financial markets such as

- the gain/loss asymmetry, i.e. one observes large drawdowns in stock prices and stock index values but not equally large upward movements.

- leverage effect: the volatility of an asset are negatively correlated with the returns of that asset.

- trading volume and volatility are correlated.

See also [4] for a more complete list of SF and their detailed analysis.

**Beyond equilibrium and perfect rationality in Economics**

The traditional approach of Economics to these features (SF, non rational behavior, etc) is rather simplistic and unsatisfactory from a scientific point of view. In fact the traditional models such as the representative agent, the assumption of perfect rationality and the equilibrium hypothesis for markets are not able to reproduce the rich phenomenology and ecology of real markets and their dynamics. In particular these models fail even in reproducing the main SF, that is the deviations from a gaussian regime and the persistence of volatility.

If the architecture of these models is investigated, this failure is not surprising because it is an inborn feature due to the linearity of these ones with respect to the price evolution. They totally neglect the non linear dynamics of markets which is instead the origin of the non trivial properties of price times series.

The recent financial crisis is a clear example of a system whose break down cannot be explained in terms of a simple linear cause-effect relation and in fact the main causes of this crisis have been traced back to concepts like collective behavior, contagion, network domino effect, coherent portfolios, liquidity crisis, etc.
The crisis also pointed out that these new concepts are usually neglected in the standard risk analysis. In particular while typical risk of stock markets is well measured by standard tools, standard risk analysis does not have any tool to measure the level of systemic risk. In addition the capability of central banks to characterize adequately the robustness of economical systems when large and global drawdowns happen is far from being satisfying. A novel approach to the risk problem is then required and the analysis is shifted from the cause-effect relation to the study of the possible intrinsic instabilities.

In this way we have set two directions (corresponding to two different levels) of development of the research project.

- On one hand we need new basic and fundamental models to understand and explain the origin of SF. The perspective of these models must go over homogeneous representative economical agents. Furthermore these models must be developed with respect to a criterion of increasing (step by step) realism, starting from a minimal and essential framework to reproduce the phenomenon under consideration and then adding more realistic features. In such a way one can fully understand the real mechanisms which give rise to SF and can control the effects and the role of each element of realism added to the model and can avoid to turn the model in an hyper-realistic black-box as it happens for many models developed by Econophysics ([6, 7]).

- However, this kind of models usually can give only a qualitative representation of financial markets and global Economics and cannot tuned with respect to empirical data (except for order book models). Therefore, on the other hand and in a complementary way to the first approach, we must develop new tools to give a quantitative and systematic framework to characterize concepts like systemic instabilities, hidden correlations, etc. These new tools must be complementary to standard risk analysis ones because these last ones are adequate to predict the risk level of financial goods and the economical stability in standard conditions. As the representative agent paradigm has been discarded, we must go over the fundamental analysis of a single and isolated financial good/firm and, for instance, develop an analogue fundamental analysis for the network in which firms are embedded as it happens in a globalized economical system. Up to now nobody knows a satisfying way to realize this new vision of risk analysis and so this is still an open problem and a hot topic after 2007 mortgage crisis.

The highlighted two lines of research have been and will be the core of my Ph. D. thesis and in the next sections I sketch the main results obtained in the first two years and the perspectives of development in this last year.

**Results and perspectives**

This novel vision of economical systems corresponds, within the framework of my research project, on one hand at developing new suitable models to investigate and interpret the origin of the SF (the first line previously discussed and partially the second one), at new models to address the risk evaluation deriving from systemic crashes (second line) and to characterize the dynamics of evolution of bubbles and their breakdown dynamics (second line), and on the other hand at data mining, for instance, at finding new SF (first and second line).

**Agent-Based Model:** The first task is achieved developing two models which are focused on the investigation of two different time scales of market dynamics, the aggregated and the microscopic level. The aggregated level analysis is carried out by the development of a minimal Agent-Based Model (ABM) for heterogeneous agents [8, 9, 10]. The model is minimal in the sense that we try to identify the essential ingredients to reproduce the most important deviations of price time series from a random walk behavior (SF). We focus on four essential ingredients: fundamentalist agents which tend to
stabilize the market; chartist agents which induce destabilization; analysis of price behavior for the two strategies; herding behavior which rules the changes of strategy. Of course real markets are more complex but the study of these elements represents a basis on which one can add more realistic features. For instance in the next section (Order Book Model) I will describe a possible improvement towards realism of this ABM. This permits to obtain a detailed understanding of the origin and nature of SF and also to discuss their self-organization. The main results of our model are:

- detailed understanding of the origin of SF with respect to the microscopic dynamics of the agents;
- demonstration that in this class of model the SF correspond to finite size effects;
- bubbles and crashes correspond to situations dominated by chartists, while fundamentalists provide a long time stability;
- in the minimal version SF are reproduced only qualitatively but, given that the agents strategies can be defined for different time scales, these finite size effects can act at different times and resemble power laws and critical exponents. However, in our perspective, these exponents are essentially a fit to the data without the important property of universality. This situation has important implication both for the microscopic understanding of the SF as well as for the analysis and interpretation of experimental data;
- we have shown the importance of non stationarity in the dynamics of the number of active agents and we propose a new mechanism for the self-organization of this state (intermittent state of SF) which is linked to the existence of a threshold for the agents to be active or not active;
- this threshold and the relative non stationarity are proposed to represent the key element in the self-organization mechanism. This self-organization, however, leads to an intermittency related to finite size effects. For this reasons we define it as self-organized intermittency.

**Order Book Model:** The microscopic level instead consists in the development of a model for the order book dynamics which is the elementary mechanism of price formation. The order book is the double auction mechanism which permits to process and store the orders placed by investors in financial markets. The price formation is the result of the arrival of proposals (orders) of buying or selling. There are two classes of orders: market orders and limit orders. The market ones correspond to orders to buy/sell at the best available price (called best bid/ask), hence they are immediately executed. The limit orders instead are orders to buy or sell at a given price which can be not necessarily the best one. By consequence limit orders may not immediately fulfilled and then they are stored in the order book.

This model is focused on the investigation of the role of liquidity with respect to the price fluctuations in financial markets [11, 12]. We found that granularity (i.e. the proxy of liquidity which is used in the model) operates as a strong amplifier of price variations when a liquidity crisis takes place. In particular one of the main result is that the system response to an incoming order diverges in the limit of vanishing granularity. It would be interesting to compare the Price Impact Surface, i.e. the price response to an order, found in the model with the empirical one so that we recently acquired all the records of all operations performed in 2002 at London Stock Exchange (LSE). We intend to study the behavior of the experimental Price Impact Surface for different proxies of liquidity. Furthermore agent-based models for financial markets usually do not take into account the problem of finite liquidity. Consequently we aim at introducing this dependence on granularity in the agent based model introduced in [8, 9, 10]. In this framework I expect that even small unbalances of the market can produce large price adjustments if a liquidity crisis is present. Therefore the role of the amplification introduced by granularity could be one of the explanation of the breaking of cause-effect relation.

**Data Mining, New Stylized Facts and Technical Analysis:** A further goal is the identification of new SF. This fact would allow to establish a wider set of experimental evidences to test and falsify the innumerable models which have been developed in the last twenty years. In this research of new SF we have found that financial prices time series exhibit a very interesting scaling relation between skewness and kurtosis (the normalized third and fourth central moment of a
pdf respectively). In fact differently from a parabolic scaling relation which is observed for a random walk or in general for not Fat Tailed pdf we find a different scaling law. Preliminary results suggest that this effect derives from large fluctuations/returns which are the leading contribution to skewness and kurtosis [14].

On the other hand the data mining approach of my research points out at characterizing the statistical properties of the so-called technical trading. Even if technical trading does not have any scientific background, it is a well established fact that some investors rely on these heuristic methods to define their investment strategies. This simple observation implies that they could deeply influence the price dynamics and then they could produce some measurable statistical properties that could be quantitatively studied. The main goal of this research would be to show that some special prices such as round prices, opening prices and closing prices, etc act as repulsive or attractive barriers for price movements.

To sum up, I aim at a systematic analysis of these ideas based on agent models and order book models together with the statistical analysis of experimental data. This research point at defining the fundamental properties of new concepts, such as contagion dynamics, psychological trading, domino effect, system instability, system coherence, etc, starting from the models, and then to identify their role in the real financial markets.

Part II

There is more than a power law in Zipf Law [15]

Zipfs Law which defines a relation among some variables and their ranks in a ranked list of these variables is widely assumed to be ubiquitous with respect to those cases where it is possible to rank the relevant quantity of a phenomenon (for instance sizes, income, frequencies, etc). In other words Zipf Law somehow appears as an universal output for ranking procedures. In this contest we intend the word universal in the sense that the Zipf Law usually appears in scientific literature as a general result independently on the details of the phenomenon under consideration. A simple observation instead reveals that the question of the validity of Zipf Law is subtler than expected. In fact let us suppose to consider a set of \( N \) numbers for which Zipf Law holds. If we divide this set into two ordered set, Zipf Law cannot hold for the set corresponding to the second half of the initial set once the rank is rescaled. In general if Zipf Law holds for a sample, then it usually cannot hold for a some subsamples or if Zipf Law holds for two samples or more, it cannot be the same for the union of these sets. Moreover we observe the origin of Zipf Law can be explained in term of an underlying power law distribution only asymptotically and not for all values of \( k \). Therefore there is more than a power law in Zipf Law because we cannot trivially interpret a sample that satisfies a Zipfian rank size rule in terms of independent draws from a power law distribution (in practice from \( 1/x^2 \)).

Results

I have introduced a simple phenomenological model to obtain a \( 1/k \) behavior from a random sampling procedure from the \( 1/x^2 \) pdf. The mechanism consists in conditioned draws, that is each time a random variable is drawn, the draws from a set of values around this variable is now prevented. Let us focus on the case of city sizes of US. The reason of such a mechanism is that once New York City is drawn, we must be no more able to extract a city with a size similar to the one of NYC. This conditioning mechanism acts as a screening effect and in details the strength of this screening is very strong for large cities while it is negligible for small cities. A similar behavior is expected since we have said the \( 1/k \) behavior can be asymptotically recovered by a random sampling.

This simple model implies any system which obeys to Zipf Law must have internal consistency in its sample. We call this consistency coherence and what is in real this consistency is a new open problem which I will try to address in the remaining of my Ph. D thesis.
Perspectives

The conditioned draw mechanism is clearly an ad hoc assumption to show that a random independent sampling and therefore an independent dynamics for the evolution of ranked quantities cannot reproduce a pure Zipf law for all values of $k$. But the mechanism which produces Zipf-like rank-size law is still unclear and this field lacks of convincing models. One of the goal of my research consists in finding a simple model for the evolution of those phenomena that exhibit Zipf law behavior.

References


