Noname manuscript No.
(will be inserted by the editor)

Econophysics

Physicists' approaches to a few economic problems

Anirban Chakraborti · Yoshi Fujiwara · Asim Ghosh · Jun-ichi Inoue · Sitabhra Sinha

Received: August 29, 2013/ Accepted: date

Abstract We review some of the recent approaches and advances made by physicists in some selected problems in economics. This rapidly growing interdisciplinary field is now popularly called "Econophysics". These approaches, mainly originating from statistical physics, have not been free of drawbacks and criticisms, but we intend to discuss these advancements and highlight some of the promising aspects of this research. We hope the readers will be able to judge the positive impact that have come out of these efforts, further improve the methods and the results, remove the shortcomings and eventually strengthen the field with their inputs.

Keywords Quantitative finance · Game theory · Statistical physics · Network science · Time-series analysis · Stochastic process

Anirban Chakraborti
Laboratoire de Mathématiques Appliquées aux Systèmes, École Centrale Paris, 92290 Châtenay-Malabry, France.
E-mail: anirban.chakraborti@ecp.fr

Yoshi Fujiwara
Graduate School of Simulation Studies, University of Hyogo, Kobe 650-0047, Japan.
E-mail: yoshi.fujiwara@gmail.com

Asim Ghosh
Theoretical Condensed Matter Physics Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700 064, India.
E-mail: asim.ghosh@saha.ac.in

Jun-ichi Inoue
Graduate School of Information Science and Technology, Hokkaido University, N14-W9, Kita-Ku, Sapporo 060-0814, Japan.
E-mail: j_inoue@complex.ist.hokudai.ac.jp

Sitabhra Sinha
The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113, India.
E-mail: sitabhra@imsc.res.in

1 Background and motivation

Physics is a hardcore observational science; it does not accept theories or explanations unless they are validated by experimental procedures as well. It is not always necessary that the developments in theoretical and experimental physics make simultaneous progress; most often one leads the other. However, the success of either branch depends on the parallel developments in the other; there needs to be an agreement between the theoretical predictions or explanations and the experimental observations or results. The scenario may be quite different in the social sciences like economics. Sometimes the after-effects of some economic policies or theoretical models can be realized only much later; at times, they cannot even be validated by the lack of empirical data. Furthermore, the social scientists do not often agree with each other on the background assumptions, predictions and success of some theoretical model, and they systematically debate upon the counter views and interpretations. On many fronts, the outcomes of their theories are later found to deviate from real-life observations, because the assumptions behind were simplified and non-realistic, though mathematically consistent. Occasionally, in order to reproduce or construct some very realistic features, the social scientists have proposed models with a large number of parameters and variables, which render them mathematically intractable and too complicated.

The econophysicists, among many other things, have advocated that one should rely primarily on the empirical observations in order to construct models and validate them. Thus, a major part of the efforts in econophysics have been the study of empirical data and financial time series analyses. Often, the empirics have guided the theoreticians in designing more realistic and practical models. It has also been found that very simple yet elegant models or mechanisms are able to reproduce much of the features of the observed data. In many cases, these (idealized) models could serve as a test-bed for many complex properties, and the models could be further improved to fit more realistic situations. Recently, due to the advent of very powerful and cheap computation, many multi-agent models could be simulated and tested, without having to wait for a long time to validate the predictions of the models. The branch of statistical physics has successfully combined the principles of classical and quantum dynamics, theory of probability and law of large numbers. Interestingly, very simple models (with a very few parameters and minimal assumptions) inspired from statistical physics, have been easily adapted in recent occasions, to gain deeper understanding and insights of many complex economic problems. There have been huge number of books [1–3], monographs [4–7], edited volumes [8–16], reviews [17–19] and journal articles on this field.

In this article, we have chosen to review certain representative efforts or approaches of the econophysicists; it is certainly not exhaustive by any means. They should only serve the purpose of illustrating or reflecting the statements made in the prededing paragraphs. Neither do we claim that the efforts are the only ones– correct and free from error; nor do we say that they should replace the tools and techniques of mainstream economics. We only modestly suggest that these interdisciplinary approaches may also prove to very effective, and possibly could compliment and strengthen the existing ones. There has been a huge surge of research activities in this field, and the references can only be too many to be possibly included here. Thus, we try to point out in this article, some general books, important reviews and key references, which further contain more details and references of original research in this field. Our goal is to only highlight the positive aspects and important outcomes of these research efforts, in order to arouse the interests of open minded economists and social scientists. We remind the readers that these approaches have their fare share of drawbacks and criticisms. We encourage the readers to go through the referred literature carefully, and then contribute to this new interdisciplinary field so that this field may develop further.

In the following sections, we will present some distinct and disjoint topics, each with sufficiently short introductions or motivations, few important results and summaries. These topics do not comprise the whole field of econophysics, and give only partial (and perhaps biased) glimpses of the research conducted over the years by the authors and their collaborators. The order of discussions of the topics do not reflect in any manner either their chronology or their importance.

2 Income and wealth distributions: Kinetic exchange models

2.1 Introduction

It was the Swiss physicist and mathematician, Daniel Bernoulli, who published *Hydrodynamica* in 1738, that eventually led to the formulation of the "kinetic theory of gases"[1]. Bernoulli had proposed for the first time, that (i) gases consist of a large number of molecules moving in all directions, (ii) their impact on a surface causes the gas pressure, and (iii) heat is simply the "kinetic energy of their motion". Then it was in 1859 when the Scottish physicist, James Clerk Maxwell, formulated the "Maxwell distribution of molecular velocities", after reading a paper on the diffusion of molecules by Rudolf Clausius. This may be considered as the first *statistical law* in physics. Five years later, an Austrian physicist, Ludwig Eduard Boltzmann, was inspired by Maxwell's paper and began developing the subject further. Thus were laid the foundations of "statistical thermodynamics"[2] by greats like Maxwell, Boltzmann, Clausius, and developed further by people like Max Planck and Josiah Willard Gibbs. They started applying probability theory, which contains the mathematical tools for dealing with very large numbers, to the study of the thermodynamic behaviour of physical systems composed of a large number of particles, giving rise to the field of "statistical mechanics"[3]. The subject of statistical mechanics provides a theoretical framework for relating the microscopic properties of individual atoms and molecules to the macroscopic or "bulk" properties of materials that we observe in our everyday life. It can be applied to various systems with an inherently *stochastic* nature in the fields of physics, chemistry, biology, and even economics and sociology. In fact, the application of

[1] Less popular is the fact that Daniel Bernoulli was also the author of *Specimen theoriae novae de mensura sortis* (Exposition of a New Theory on the Measurement of Risk), published in 1738.

[2] The term was coined by the American physical chemist J. Willard Gibbs, and appeared in his book title published in 1902.

[3] The term was first used by the Scottish physicist, James C. Maxwell, in 1871.

statistical physics ideas and tools in modeling economics and sociology, have led to the interdisciplinary fields of "Econophysics"[4] and "Sociophysics"[5].

A natural question that may rise in one's mind is: "*How can such a physical theory like statistical mechanics, which deals with particles, be applied to an economic system, composed of (human) agents?*" Well, that is what we set to describe next, while dealing with the problem of income and wealth distributions in the society[6]. Physicists have come up with some very elegant and simple kinetic exchange models in recent times to the problem of economic inequality, based on the simple philosophy: A single molecule of gas does not have a temperature (T), or a pressure (p). It is simply a point-like particle that moves at a particular speed, depending on how much energy it has, governed by the statistical law of Maxwell-Boltzmann distribution of molecular speeds. However, when there are of the order of 10^{23} or so molecules in an isolated and sealed box of volume V, their collective behaviour can be captured by the equation of state: $pV = RT$, where R is the gas constant; and even though each individual particle is moving at *random*, one can predict with extraordinary accuracy how many of them will, for example, hit the walls of the box at any one time. Similarly, the belief of the physicists is that the economy can be described in terms of simple observables. An individual person is neither an economy, nor has any of the characteristics of the entire economy. However, a million such persons acting individually creates the economy, and may be described by some rules that perhaps allow an economy to be predicted, just as the equation of state mathematically describes pressure and temperature, and predicts the aggregate behaviour of atoms. Also, the standard economic theory would like to consider that the activities of individual agents are driven by the *utility maximization principle*. The alternate picture proposed by physicists is that the agents can be simply viewed as gaseous particles exchanging "money", in the place of energy, and trades as money (energy) conserving two-body scatterings, as in the *entropy maximization* based kinetic theory of gases [21]. This qualitative analogy between the two maximization principles seems to be quite old – both economists and natural scientists had already noted it earlier in many contexts, but this equivalence has gained firmer ground only recently.

Truly, it would be difficult to find any society or country where income or wealth is equally (or fairly) distributed amongst its people. Socio-economic inequality is not just limited to the modern times; it has been a persistent phenomenon and a constant source of irritation to most, since antiquity. It is one of the most fiercely debated subjects in economics, and the economists and philosophers have spent much time on the normative aspects of this issue ([20, 22–24]). The direct and indirect effects of inequality on the society have also been studied extensively, particularly, the effects of inequality on the growth of the economy ([25–28]) and on the econo-political scenario ([29–32]). There are several non-trivial issues and related open questions: *How are income and wealth distributed? What are the forms*

[4] The term was coined by the American physicist, H. Eugene Stanley, in a conference on statistical physics in Kolkata (erstwhile Calcutta) in 1995, and first appeared in its proceedings published in the journal Physica A (1996).

[5] The term was first used by the French physicist, S. Galam, and appeared in an article published in 1982.

[6] This part contains some text overlapping with the recent monograph by Chakrabarti et al. (2013) [7].

of the distributions? Are they universal, or do they depend upon specific conditions of a country?

Such questions have indeed intrigued many great personalities in the past: Vilfredo Pareto, more than a century ago, made extensive studies in Europe and found that wealth distribution follows a power law tail for the richer section of the society ([33]), known now as the *Pareto law*. Separately, Roger Gibrat worked on the same problem and he proposed a "law of proportionate effect" ([34]). Much later, Champernowne also considered this problem systematically and came up with a probabilistic theory to justify Pareto's claim ([35,36]). Subsequently, it was found in numerous studies that the distributions of income and wealth indeed possess some globally stable and robust features (see, e.g., [17] for a detailed review). In general, the bulk of the distribution of both income and wealth seems to fit both the *log-normal* and the *Gamma* distributions, reasonably well. Economists usually prefer the log-normal distribution ([37,38]), whereas statisticians ([39]) and more recently, physicists ([40,41,17]) tend to rely more on alternate forms such as the Gamma distribution (for the probability density) or Gibbs/ exponential distribution (for the cumulative distribution). There is considerably more consensus on the upper end of the distribution, that is the tail of the distribution – described by a power law as was found by Pareto.

These observed regularities in the income distribution may thus indicate a "natural" law of economics. The distribution of income $P(x)$ is defined as follows: $P(x)dx$ is the probability that in the "equilibrium" or "steady state" of the system[7], a randomly chosen person would be found to have income between x and $x + dx$. Detailed empirical analyses of the income distribution so far indicate

$$P(x) \sim x^n \exp(-x/T), \quad \text{for} \quad x < x_c, \tag{1}$$

and

$$P(x) \sim x^{-\alpha-1}, \quad \text{for} \quad x \geq x_c, \tag{2}$$

where n and α are two exponents, and T denotes a scaling factor. The latter exponent α is called the *Pareto exponent* and its value ranges between 1 and 3 (see, e.g., [42,43]). A historical account of Pareto's data and that from recent sources can be found in [44]. The crossover point x_c is extracted from the numerical fittings of the initial Gamma distribution form to the eventual power law tail. One often fits the region below x_c to a log-normal form: $\log P(x) = \text{const} - (\log x)^2$. As mentioned before, although this form is often preferred by economists, the statisticians and physicists think that the Gamma distribution form fits better with the data (see [17], [39] and [45]). Fig. 1 shows schematically the features of the cumulative income or wealth distribution.

Most of the empirical analyses, especially with recent income data, have been extensively reviewed, e.g., in the chapter of book by Chakrabarti et al. [7]. It may be mentioned that compared to the empirical work done on income distribution, relatively fewer studies have looked at the distribution of wealth, which consist of the *net* value of assets (financial holdings and/or tangible items) owned at a given instant. The lack of an easily available data source for measuring wealth, analogous to income tax returns for measuring income, means that one has to resort to

[7] To keep things simple, we will often be using the terms "equilibrium" or "steady state" interchangeably; strictly speaking, for systems that are "non-ergodic", one can only write "steady state".

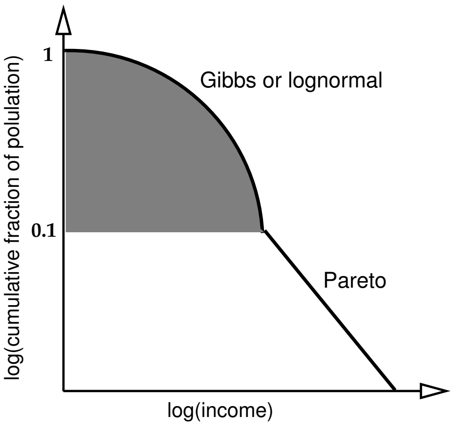


Fig. 1 When one plots the cumulative wealth (income) distribution against the wealth (income), almost $90-95\%$ of the population fits the Gibbs distribution, or often fitted also to log-normal form (Gibrat law) – indicated by the shaded region in the distribution; for the rest (very rich) $5-10\%$ of the population in any country, the number density falls off with their wealth (income) much slowly, following a power law (Pareto law). It is found that about $40-60\%$ of the total wealth of any economy is possessed by $5-10\%$ of the people in the Pareto tail. Taken from Chakrabarti et al. (2013) [7].

indirect methods. Again, one notes that the general feature observed in the limited empirical study of wealth distribution, is that of a power law behavior for the wealthiest 5-10 % of the population, and exponential or log-normal distribution for the rest of the population. The Pareto exponent as measured from the wealth distribution is always found to be lower than that for the income distribution, which is consistent with the general observation that, in market economies, wealth is much more unequally distributed than income ([46]). Interestingly, instead of focussing on the income of individuals when one shifts attention to the income of companies, one still observes the power law tail. A study of the income distribution of Japanese firms ([42]; see also [6]) concluded that it follows a power law (with exponent value near unity, which is also often referred to as the Zipf's law). Similar observation has been reported for the income distribution of companies in the US ([47]). Such strikingly robust features of the distribution $P(x)$, in income or wealth, seem to be well-established from the analyses of the enormous amount of data available today. The important question is – *if inequality is universal (as some of its gross features, studied by Pareto, Gibrat and others, indicate), then what is the reason for such universality? Is it plausible that this only reflects a basic natural law, with simple physical explanation?* Many econophysicists actually believe so. According to them, the regular patterns observed in the income (and wealth) distribution are indeed indicative of a natural law for the statistical properties of a many-body dynamical system representing the entire set of economic interactions in a society, analogous to those previously derived for gases and liquids.

The class of kinetic exchange models ([56, 68, 57–60]) are simple microeconomic models with a large number of "agents" and the "asset" transfer equations among the agents due to "trading" in such an economy, closely resemble the process of "energy" transfer due to "collisions" among "particles" like those in a thermodynamic system of ideal gas. In these models, the system is assumed to be made

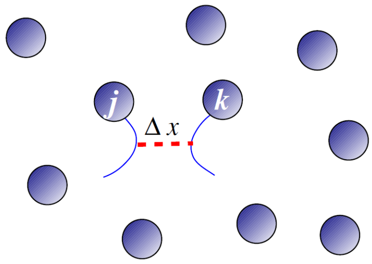


Fig. 2 The kinetic exchange models prescribe a microscopic interaction between two units analogously to a kinetic model of gas in which, during an elastic collision, two generic particles j and k exchange an energy amount Δx, as in (3). Taken from Chakrabarti et al. (2013) [7].

up of N agents with assets $\{x_i \geq 0\}$ $(i = 1, 2, \ldots, N)$. At every trade, an agent j exchanges a part Δx with another agent k chosen randomly. The total asset $X = \sum_i x_i$ is constant, as well as the average asset $\langle x \rangle = X/N$. After the exchange the new values x'_j and x'_k are $(x'_j, x'_k \geq 0)$

$$\begin{aligned} x'_j &= x_j - \Delta x \,, \\ x'_k &= x_k + \Delta x \,. \end{aligned} \tag{3}$$

The form of the function $\Delta x = \Delta x(x_j, x_k)$ defines the underlying dynamics of the model. Fig. 2, shows the schematic picture that captures the essence of these models.

Among the first examples of kinetic exchange models of markets proposed by social scientists, we must mention the works of E. Bennati [52–54], an economist, and those of John Angle [48–51], a sociologist. Independently, physicists also had made several studies. The first kind of models both with multiplicative and additive exchanges, were proposed by Ispolatov et al. [55].

The steady state distribution for a system with pure random asset exchange is an exponential one, as was found by Gibbs a hundred years ago (see e.g., [41, 17,68]). However, the introduction of "saving propensity" ([56]) brought forth the Gamma-like feature of the distribution $P(x)$ and such a random exchange model with uniform saving propensity for all agents was subsequently shown to be equivalent to a commodity clearing market where each agent maximizes his/her own utility ([61]). A further modification of the model produces ([58]) a power law for the upper or tail end of the distribution of money, as has been found empirically. These are explained in the following sub-section.

2.2 Model with uniform savings

In any trading, savings come naturally [46]. A saving propensity factor λ was introduced in the random exchange model [56], where each trader at time t saves a fraction λ of its money $x_i(t)$ and trades randomly with the rest:

$$x_i(t+1) = \lambda x_i(t) + \epsilon_{ij} \left[(1-\lambda)(x_i(t) + x_j(t))\right], \tag{4}$$

$$x_j(t+1) = \lambda x_j(t) + (1-\epsilon_{ij}) \left[(1-\lambda)(x_i(t) + x_j(t))\right], \tag{5}$$

where

$$\Delta x = (1 - \lambda)[\epsilon_{ij}\{x_i(t) + x_j(t)\} - x_i(t)], \tag{6}$$

where ϵ_{ij} being a random fraction. This randomness reflects the *stochastic* nature of the trading. By definition, λ is a proper fraction, i.e., $0 \leq \lambda \leq 1$.

Interestingly, in this model, the market (non-interacting at $\lambda = 0$ and 1) becomes 'interacting' for any other non-vanishing λ: For fixed λ (uniform for all agents), the steady state distribution $P(x)$ of money is exponentially decaying on both sides with the most-probable money per agent shifting away from $x = 0$ (for $\lambda = 0$) to X/N as $\lambda \to 1$ (Fig. 3). Here, the *self-organizing* feature of the market[8] , induced by sheer *self-interest* of saving by each agent without any global perspective, is quite significant as the fraction of paupers decrease with saving fraction λ and most people end up with some finite fraction of the average money in the market − for $\lambda \to 1$, the economy is ideally 'socialist', and this is achieved just with people's self-interest of saving. Although this fixed saving propensity does not give yet the Pareto-like power-law distribution, the Markovian nature of the scattering or trading processes is effectively lost. Indirectly through λ, the agents get to know (start interacting with) each other and the system co-operatively self-organizes towards a most-probable distribution ($x_p \neq 0$) (see Fig. 3).

Based on numerical results, it has been claimed (through heuristic arguments) that the distribution is a close approximate form of the Gamma distribution [63]:

$$P(x) = \frac{n^n}{\Gamma(n)}\, x^{n-1} \exp(-nx) \tag{7}$$

where $\Gamma(n)$ is the Gamma function whose argument n is related to the savings factor λ as:

$$n = 1 + \frac{3\lambda}{1-\lambda}. \tag{8}$$

This result has also been supported by numerical results in [64]. However, later studies [65,66] analyzed the moments, and found that moments up to the third order agree with those obtained from the form of the Eq. (8), and discrepancies start from fourth order onwards. Hence, the actual form of the distribution for this model is still an open question.

2.3 Model with distributed savings

In a real society or economy, the interest of saving varies from person to person, which implies that λ is a very inhomogeneous parameter. To reproduce this situation, one moved a step closer to the real situation where saving factor λ was widely distributed within the population [58,57,59]. The evolution of money in such a trading can be written as:

$$x_i(t+1) = \lambda_i x_i(t) + \epsilon_{ij}\left[(1-\lambda_i)x_i(t) + (1-\lambda_j)x_j(t)\right], \tag{9}$$

$$x_j(t+1) = \lambda_j x_j(t) + (1-\epsilon_{ij})\left[(1-\lambda_i)x_i(t) + (1-\lambda_j)x_j(t)\right]. \tag{10}$$

[8] Self-organization also occurs in other market models when there is restriction in the commodity market [62].

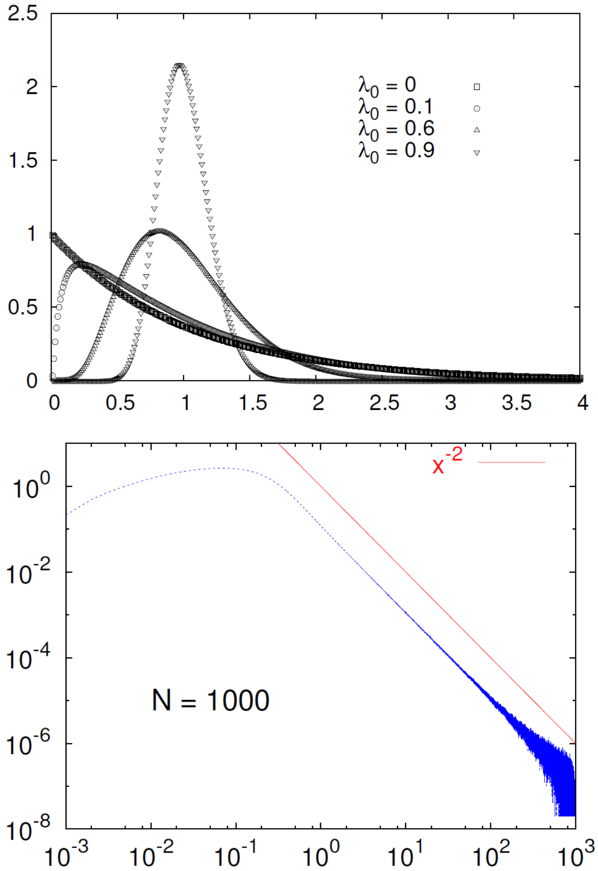


Fig. 3 Steady state money distribution $P(x)$ *vs.* x for: (TOP) The model with uniform savings. The data shown are for different values of λ: $0, 0.1, 0.6, 0.9$ for a system size $N = 100$. All data sets shown are for average money per agent $X/N = 1$. Taken from [41]. (BOTTOM) The distributed λ model with $0 \leq \lambda < 1$ for a system of $N = 1000$ agents. The x^{-2} is a guide to the observed power-law, with $1 + \alpha = 2$. Again, the average money per agent $X/N = 1$. Taken from [41].

The trading rules are similar as before, except that

$$\Delta x = \epsilon_{ij}(1 - \lambda_j)x_j(t) - (1 - \lambda_i)(1 - \epsilon_{ij})x_i(t), \tag{11}$$

where λ_i and λ_j are the saving propensities of agents i and j. In this model, the agents have fixed (over time) saving propensities, distributed independently, randomly and uniformly, within an interval 0 to 1. A particular agent i saves a random fraction λ_i ($0 \leq \lambda_i < 1$) and this λ_i value is *quenched* for each agent, i.e., λ_i does not change with time t.

Starting with an arbitrary initial (uniform or random) distribution of money among the agents, the market evolves with the trading. At each time, two agents are randomly selected and the money exchange among them occurs, following the above mentioned scheme. One checks for the steady state, by looking at the stability of the money distribution in successive Monte Carlo steps t (one Monte Carlo time step is defined as N pairwise exchanges). Eventually, after a typical *relaxation time* the money distribution becomes *stationary.* This relaxation time is dependent on system size N and the distribution of λ (e.g, $\sim 10^6$ for $N = 1000$ and uniformly distributed λ). After this, one averages the money distribution over $\sim 10^3$ time steps. Finally, one takes the configurational average over $\sim 10^5$ realizations of the λ distribution to get the money distribution $P(x)$. Interestingly, this has a strict power-law decay, and the decay fits to Pareto law (Eq. 2) with $\alpha = 1.01 \pm 0.02$ (Fig. 3). One may note, for finite size N of the market, the distribution has a narrow initial growth up to a most-probable value x_p after which it falls off with a power-law tail for several decades. This Pareto law (with $\alpha \simeq 1$) covers almost the entire range in money x of the distribution $P(x)$ in the limit $N \to \infty$. This power law is extremely robust, in the sense that apart from the uniform λ distribution used in these simulations in Fig. 3, this decay can also be reproduced for a distribution

$$\rho(\lambda) \sim |\lambda_0 - \lambda|^\alpha, \quad \lambda_0 \neq 1, \quad 0 < \lambda < 1, \tag{12}$$

of quenched λ values among the agents, for all $\alpha > 0$.

2.4 Summary and discussions

In summary, viewing the economy as a “thermodynamic” system ([67–70]), one can identify the income distribution with the distribution of energy among the particles in a gas. Several attempts by social scientists (see e.g., [52–54,48–51]) also provide impetus to this interdisciplinary approach. In particular, the class of kinetic exchange models ([56,68,57–59]) have provided a simple mechanism for understanding the unequal accumulation of assets. While being simple from the perspective of economics, they have the benefit of gripping a key factor – *savings* – in socio-economic interactions, that results in very different societies converging to similar forms of unequal distribution. Interestingly, the economic inequality is a natural outcome of this framework of stochastic kinetics of trading processes in the market, *independent* of any exogenous factors. Thus, the kinetic exchange models demonstrate how inequality may arise naturally. They also indicate how its effects may be partially reduced by modifying the saving habits.

Several analytical aspects of this class of models have been studied (see e.g, [71–73,5,66]). It is noteworthy that presently this is the only known class of models which, starting from microeconomics of utility maximization and solving for the resultant dynamical equations in the line of rigorously established statistical physics, can quite reliably reproduce the major empirical features of income and wealth distributions in economies.

These developments have, of course, not gone without criticism (see e.g., [74–76]), and subsequent rebuttal ([44]). In view of the embarrassing failure of main stream economic schools to anticipate or correctly analyses the recent economic crisis, there have been some recent interests by the main stream economic schools to revisit such physically motivated models of the market dynamics and their solutions (see, e.g., [77]).

3 Market mechanism: Agent-based models

3.1 Introduction

In this section, we will discuss some games on agent-based model [78,79,3] which may be considered as toy models of the market mechanism. One of the most famous games proposed related to this issue, renowned also for its simplicity, is the El Farol bar problem [80]. Brian Arthur introduced in 1994 the game to illustrate the idea of 'inductive reasoning'. In Santa Fe town, there was a bar named El Farol Bar, where every Thursday a musical programme was organized. People went to the bar for entertainment. But when the bar was too crowded then the bar was not enjoyable. Based on this observation, he proposed the repetitive game model where it was suggested the people in the bar would be enjoying only if less than, say, 60% of capacity was occupied. Assuming that all agents were not interacting with each other, and taking their decisions parallelly, he modeled the problem considering only previous attendance history. He defined a strategy space based on previous history and argued that the attendances of the bar can be around 60% of the total number people. The solution therefore is completely based on inductive reasoning, i.e., people learned strategies from previous history and corrected from past errors.

Later a variant of the El Farol bar problem was introduced by Challet and Zhang, was named as Minority Game [79,81–84,18]. In the Minority Game problem, two restaurants are considered (to allow for symmetric choices). The agents on the less crowded side will be the winners (payoff 1 for each) and agents in the more crowded side will be loser (payoff 0 for each). People again learn from the past history and change their strategies to minimize their error/loss. The steady state fluctuation associated with the population is important and need to be minimized for efficient 'reasoning' or learning by the players. Many methods of collective learning have been proposed and studied for this game to reduce fluctuation in this problem and also the convergence time.

In the next part of this section we will discuss another variation of the (two choices) El Farol Bar problem, called Kolkata Paise Restaurant problem [85–87]. In this problem there are many choices (of restaurants) and many agents. It is again a repetitive game. Every restaurant can serve only one agent each day and the price of a meal is same for all restaurants. The agents here also do not interact with each other and everyone wants to choose an unique restaurant each day. If more than one agent arrive at a restaurant then one of them will be randomly chosen by the restaurant and will be served; the rest will not get any meal for that day. The utilization fraction is defined as the ratio of the average number of agents getting dinner in a day to the total number of the restaurants. The main aim of this problem is finding out a strategy which will give a maximum utilization taking the smaller time to converge to the solution.

3.2 El Farol Bar problem

Before the El Farol Bar problem was introduced, the economists had mainly modeled the problems based on deductive rationality. Although it was useful to generate a theoretical problem, but it was observed that deductive rationality breaks down under some complications [80]. It was pointed out that inductive reasoning was needed to model the problems in a more realistic way. It was argued that in our society, we may make our decisions based on inductive reasoning and not based on deductive rationality. The inductive reasoning gives good sense as an intellectual process and it is very easy to model.

Historically, the El Farol Bar problem was introduced as follows: In Santa Fe, New Mexico, there was a bar and every Thursday night, a finite number of people went to the bar. But as the capacity of the bar was finite, it was no fun to go there if it was too crowded and people therefore decided to go or not depending on their personal past experience. Inspired by this fact, Brian Arthur defined his model and it can be described as follows: Suppose total 100 agents are available to go to the bar and capacity of the bar is only 60. Therefore, if any one finds that the attendance of the bar exceeds the capacity value, in that case, staying at home would be a better choice than going to the bar. Similarly, if the attendance of the bar does not exceed the capacity value, in that case, going to the bar would be a better choice than staying at home. Also the agents do not interact with each other while taking their decisions, but previous attendance history of the bar is available to everyone. Depending upon the previous history, every agent will choose either to go to the bar or to stay at home. By doing computer simulation,Brian Arthur surprisingly found that the mean attendance converges to 60 [80].

3.3 Minority Game

As mentioned already, a symmetric choice variant of the El Ferol Bar problem was introduced by Challet and Zhang [79, 81–84] which is known as the Minority game problem. In their model, they considered two restaurants ($n = 2$) and N ($= 2M+1$; M integer) agents. The agents are going to one of the two restaurants each day. Every day, one restaurant will always be more crowded than the other. The agents in the less crowded restaurant are said to be in minority side (or winner side) and will receive a positive payoff. At the same time, the other restaurant (say majority side or looser side) is more crowded and every agent in that restaurant will get a payoff 0. In this problem, if all the agents choose any one of two restaurants randomly every day then the population distribution of each restaurant will be Gaussian and the peak of the distribution will be around M with a mean-square deviation σ (or fluctuation) which is the order of $\sqrt{N}$. The fluctuation in the problem actually is a measure of the loss of resources in this game. Therefore, the fluctuation in the problem should be optimized and researchers developed algorithm to adjust weights in the strategy space depending upon the past history of his/her success or failure. Now if anyone use only h days as a memory size for taking decision on the next day then total number of possible strategies will be 2^h which is a fast-growing function of h. They observed that the fluctuation in the problem could be reduced to a certain minimum level by increasing the memory size of the agents. The fluctuation is numerically found to have a minimum value (about 1/10 of $\sqrt{N}$, the random case value) between short memory and large memory size of agents. A Monte Carlo simulation of the variability of fluctuation in memory size is shown in fig. 4.

*Minority Game and Market Model:*The main aim of introducing the Minority Game model was to serve as a market model. It was mapped to the market model by

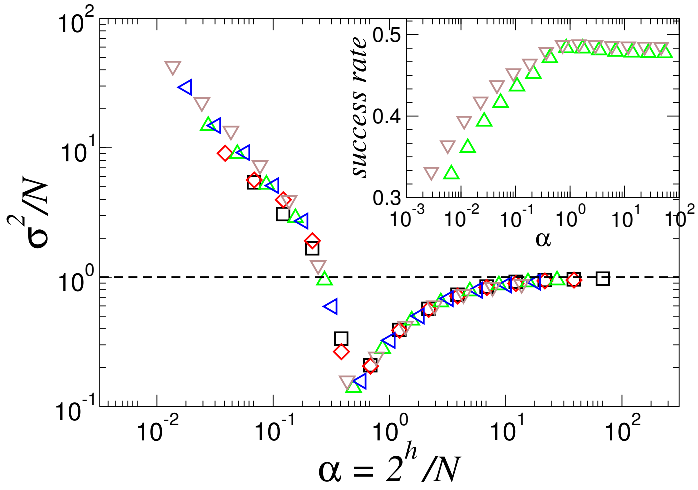


Fig. 4 Square of fluctuation versus $\alpha = 2^h/N$ for different number of agents $N = 101, 201, 301, 501, 701$ (□, ◊, △, ◁, ▽, respectively). Inset: Variation of mean success rate of the agents' with α. From [83].

identifying the two choices as 'buy' or 'sell' options for any trader. More buyers than sellers in any day means gain (positive payoff) for sellers and opposite when they are more in number. Therefore, the total payoff of all agents is equivalent to excess demand of the market. We know that price return of the market is directly related to the excess demand. By doing this, one can relate the problem to the MG (for details see [79]).

3.4 Kolkata Paise Restaurant Problem

The Kolkata Paise Restaurant Problem was introduced by Chakrabarti et al. in 2009 [85] to accommodate many choices by the agents or players. In this problem, the number of choices (restaurants) and number of agents, both are large. Again here, the agents do not interact with each other for making their decisions and they take their decisions parallelly. Specifically one considers n restaurants and N agents where $n \sim \mathcal{O}(N)$. Also the previous history of the game is be available to everyone. Every agent can choose only one restaurant in any day. In this game, price of a meal is same for all the restaurants and each restaurant can serve only one agent each day. So, if more than one customer arrive in a day in a restaurant, one of them will be randomly chosen and will be served and so the rest of the agents will not get food for that day.

3.4.1 Strategies of the game

In KPR problem, the utilization fraction is defined as the ratio of average number of agents getting food every day to the total number of restaurants. In this part, We will discuss different strategies and their corresponding utilization fraction. The efficient strategy is such that it will give the maximum utilization fraction within a finite time limit.

Random Choice First let us talk about random choice strategy [85]. In this case, an agent will choose any restaurant at random. Therefore, the agents do not use their memory related to previous attendances of the restaurants for making their choices. Also the agents are not discussing with each other to make their choice and all decisions are taken parallelly each day. In this case every agent can choose only one restaurant for a day. we know that every restaurant can serve only one agent for a day so it is not guaranteed that every agent will get food every day. Next part, we will calculate how many agents on an average get served every day.

Suppose there are N agents and n restaurants and in this case all agents choose any restaurant with probability $p = 1/n$. Now the probability of restaurants chosen by m agents for a day is given by

$$\begin{aligned}\Delta(m) &= \binom{N}{n} p^m (1-p)^{N-m}; \qquad p = 1/n \\ &= \frac{(N/m)^m}{m!} exp(-N/n) \quad \text{for} \quad N \to \infty, \quad n \to \infty. \end{aligned} \tag{13}$$

Therefore, the probability of the restaurants not chosen by any agent can be written as $\Delta(0) = exp(-N/n)$. Now we can write an average fraction of restaurants visited by at least one agent as

$$\bar{f} = 1 - \Delta(0) \sim 0.63 \quad \text{for} \quad N = n. \tag{14}$$

Therefore, in the random choice case we get about 63% utilization and obviously the convergence time to reach this utilization fraction value is zero.

Rank dependent Choice Although the price of a meal in all the restaurants is the same, we can assume that all restaurants have different rank depending upon the service quality, food quality etc. . To make a model, we assume that the k-th restaurant has rank of k and probability to go to that restaurant is proportional to k. Again all agents do not use previous history for this strategy. If all agents follow this strategy then it was observed numerically that utilization fraction is about 57% [87]. Therefore, this strategy is less efficient compared to the random choice case. Again no time is needed to reach this steady value means same as the random choice case.

Crowd Avoiding Cases (A) Fully Crowd Avoiding Case: In this case, all the agents use their memory of previous day attendances history for taking their decisions. If all the agents decide to choose only the previous day's vacant restaurants randomly and completely avoiding other restaurants then it was observed numerically that the average utilization fraction becomes about 46% [87]. Therefore, this strategy is much less efficient compared to the previously discussed strategies.

(B) Stochastic Crowd Avoiding Case: In this case, the agents do not avoid the previously crowded restaurants completely. Suppose $n_k(t-1)$ is the number of agents arriving or choosing k-th restaurant on $(t-1)$-th day, then next day (t) these agents will visit the same restaurant with probability $p = 1/[n_k(t-1)]$ and will go to any other restaurants with uniform probability $p\prime = (1-p)/(N-1)$. If all the agents are playing with this strategy then utilization fraction becomes about 80% [87] which is much larger than that for the previously described strategies.

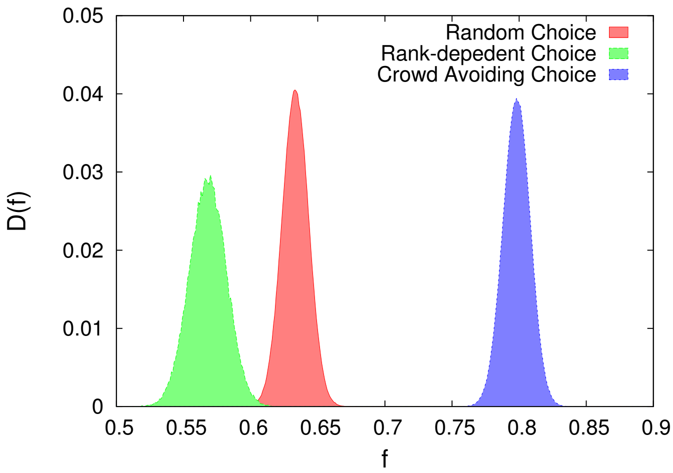


Fig. 5 Probability distribution of utilization for different strategies; using Monte Carlo simulation for $N = n = 10^5$ are shown. From [88].

But the time to reach the steady state value of the efficiency here is of the order of $\log N$ which is much larger than the other strategies (but still much smaller compared to N).

An approximate estimate for above utilization can be made as follows: Suppose in steady state a_i is the fraction of restaurants visited by i agents and we assume $a_i = 0$ for $i \geq 2$; we estimate that number of restaurants where 3 or more customers arrive on any day are negligible. If we take the equal number of agents and restaurants ($N = n$) then we can easily write the equations as

$$a_0 + 2a_2 = 1$$
$$a_0 + a_1 + a_2 = 1 \ .$$

According to this strategy, every agent who visited any restaurant alone on a given day, will surely go the same restaurant next day($p = 1$). But if any restaurant was visited by two agents then the agents will go to the same restaurant with probability $p = 1/2$ the next day. In this process every time $(a_2/4 - a_2a_2/4)$ fraction of restaurants will be vacant from the restaurants previously visited by 2 agents. Similarly a_0a_2 fraction of restaurants will be visited by one agent from the restaurants previously visited by none. Therefore one can write

$$a_0 - a_0a_2 + a_2/4 - a_2a_2/4 = a_0 \ .$$

If the above three equations are solved, we get $a_0 = 0.2$, $a_1 = 0.6$ and $a_2 = 0.2$. So the average fraction of restaurants visited by at least one agent is $a_1 + a_2 = 0.8$, which is very close to the results of the Monte Carlo simulation shown in Fig. 5.

3.4.2 KPR and link with Minority Game

Another important use of the stochastic crowd avoiding strategy is in Minority Game. Dhar et al. [89] showed that fluctuation of Minority Game can reduce to zero within very short time scale by applying a variation of the stochastic crowd avoiding strategy. The strategy used was as follow: Suppose at any instant of time population of two restaurants are $M + \Delta + 1$ and $M - \Delta$ in MG problem. Next day majority side agents will change the choice with probability $p_+ = \Delta/(M + \Delta + 1)$ and minority side people will remain their previous choice ($p_- = 0$). Here agents use previous day information only and Δ will become zero order within $\log\log(2M + 1)$ time.

But in this case there was a problem. At the time when Δ becomes 0 the flipping process will stop which means that majority people ($M + 1$ in number) will remain in the same state for the rest of the game. Therefore though the solution gives zero fluctuation the situation will be very unjust a solution for the majority group. To overcome this problem the probability was later modified as $p_+ = (g\Delta)/(M + g\Delta + 1)$ (g any real positive parameter) and $p_- = 0$ [90]. It was observed that below $g < 2$ the dynamic stops after some time but for $g > 2$ it never stops,thought $\Delta(g) \sim (g - 2)$ can be made arbitrarily small.

3.5 Summary and discussions

Here we have discussed different games and their efficiencies for different strategies. First, we considered the El Farol Bar problem which was a game based on inductive reasoning. In this game, many agents went to a restaurant in a city. The agents would be happy if they found that crowd of the restaurant did not exceed a certain threshold value. In the El Farol Bar problem, the agents would not discuss for making their decisions and one would go to the Bar if he/she was expecting a less crowded situation. The game was actually represented as a toy model of 'inductive learning' strategy. It was also shown that models based on the inductive learning were more effective than the deductive rationality for collective or social decision making contexts.

Later, we have discussed about a variant of El Farol Bar problem which was called Minority Game. In this game there were two restaurants and many agents. All the agents choose one of the this restaurants every day without interacting with each other and the agents in less crowded restaurant would receive payoff and the agents in the crowded restaurant lose. Both restaurants are similar regarding the price of the meal but the agents in crowded side did not get any payoff for that day. In this problem fluctuation associated with attendances of the agents in the restaurants is an important quantity which can be identified as similar to volatility in financial markets. We saw that the strategy based on memory size could reduce the fluctuation up to about 1/10 of the random process fluctuation.

In the next part we have discussed the Kolkata Paise Restaurant (KPR) problem. In this problem, N agents and n ($\sim \mathcal{O}(N)$) restaurants were considered. The agents are again non-interacting and they use the past performance records only for making their decisions. On the other side, price of a meal for each restaurant is assumed to be the same so that no budget restriction or specification can dictate the choice and one restaurant could serve one agent each day and also agents choose only one restaurant in any day. Therefore, if more than one agent had arrived at any restaurant in a day then one of the them would be randomly picked up and served and rest of them would not get food for that day. Here we discussed different strategies of the game and their efficiency (by measuring utilization fraction). We observed that the most efficient strategy was the stochastic crowd avoiding case where the utilization fraction attain a value about 80% within a time, bounded by $\log N$.

In last part, we have seen that a variant of the stochastic crowd avoiding strategy (developed for KPR problem) was applied in the Minority Game problem. Using this strategy, the fluctuation associated with the Minority Game reduced to the order of zero by taking few time steps (magnitude of convergence time in order of $\log \log N$). But it was also observed that the dynamics of the game would be stopped after that. As discussed, this problem can be avoided by taking some noise trader in the game, although the fluctuation remains non zero, thought very small compared to the random choice case.

4 Economic Success and Failures: Analyses and modeling

4.1 Introduction

Economic phenomena rarely give rise to equitable distributions where everyone has the same share of the market. In fact, the same emergence of inequality that marks wealth and income distributions (see Section of this paper), also characterizes other aspects of economic life, such as the respective fates of an ensemble of products (or for that matter, services or ideas) that maybe of more or less similar quality. This applies even to larger economic structures such as organizations and companies which drive economic growth in most societies; Ormerod has pointed out in the book *Why Most Things Fail* [92] that of the successful companies that existed in the past, only a handful have managed to survive to the present. In fact, the relative success of firms appear to have a similar long-tailed distribution described by a power law function that characterizes the distribution of personal income or wealth [93]. It is thus of some importance to identify features that characterize the process of economic success and failure.

What decides whether a newly introduced entity in the market will eventually succeed in face of severe competition from several other competitors is often not so much a result of intrinsic differences between them but rather a result of a social collective consensus that emerges stochastically from the interactions between the choice behavior of individual agents. The choices in many such decision problems often constitute a discrete set, sometimes even having a binary nature, for instance, whether one should cooperate with others or defect by exploiting (free-loading off) the cooperators. These problems resemble the formulation of spin models used by statistical physicists to analyze in detail cooperative phenomena [94, 95]. It is therefore not surprising that the phenomenon of how and why certain entities become successful through competition has been a topic of interest among econophysicists.

To understand how people choose from a set of options, mainstream economics assumes that each individual decides on a specific alternative that maximizes his/her utility function. Moreover, except possibly for the formulation of the utility function, the behavior of other agents are usually not considered as directly affecting the decision to select a particular choice. It has however become apparent over the past few decades, e.g., as exemplified by Schelling's analysis of the reasons for housing segregation along racial lines [96], that in many cases the apparent freedom of choice of an agent may well be illusory. Indeed, the decision taken by an individual is affected by those taken by his/her peers or rather, the actions of neighboring elements in the corresponding social network [97,98]. A physics-based approach which stresses on the role of interactions (and hence the social environment) in taking a decision contrasts with the conventional economic approach of utility maximization by individual rational agents. This can be crucial for explaining why a very popular product can suddenly emerge even though it may difficult to distinguish it from its competitors.

The response of mainstream economic theory to this may well be that it suggests the existence of an unobservable property that should be included as a term in the utility function which differentiates the popular entity from its competitors. However, as this assertion cannot be verified empirically, we cannot comment on its scientific validity. By contrast, an interactions-based mechanism may suggest that although a specific choice did not have any intrinsic advantage over others, stochastic fluctuations may have resulted in a relatively larger number of individuals opting for it initially. This may generate a small advantage in favor of the choice being adopted by others. For example, more people buying a product may make it more economical to produce (economy of scale) or the use of a particular product by some may make it more likely to be adopted by their acquaintances (network externalities). Eventually, through a process of self-reinforcing or positive feedback via interactions, an inexorable momentum is created in favor of the entity that builds into an enormous advantage in comparison to its competitors (see e.g. Ref [99] for an application of this idea into high-technology markets). While the idea of such feedback or externalities has been discussed in economics from quite early on (see the article by Paul Krugman [100] and responses to it by others, including Kenneth Arrow [101]), the quest for analytically tractable linear economic models among mainstream practitioners has meant that a nonlinear interactions-based perspective for analyzing economic phenomena has become popular only recently with the advent of econophysics. The study of economic popularity is one such area that has benefited from this incursion of physics ideas into economics.

4.2 The economic fate of movies: A case study

Movie popularity provides an arena for analyzing the phenomena of economic success or failure arising through interactions among agents - not least because of the availability of large quantities of publicly available digital data. In fact, there have been recent attempts to use temporal patterns in the digital data, e.g., the increase in the number of blog posts on specific movies [102] or in rising activity in the Wikipedia entries for soon-to-be or newly released movies [103], to provide early prediction for the success of a movie. However, we shall here focus on the data about the box-office gross receipts of a movie during its initial run at theaters. Note that, unlike the popularity of several other types of products (e.g., to measure the popularity of a car, we look at how many people are driving it), in the case of a movie it is not completely obvious how to identify a unique observable that will be efficient at capturing all the different dimensions of its popularity. For example, one can consider the average of the ratings given by different film critics in various media, votes received in movie-related forums online or the total number of DVDs bought or rented. For example, we can take the case of popular movies decided by votes of registered users of the Internet Movie Database (IMDb) (http://www.imdb.com), one of the largest movie-related online sites. As voters can give a score between 1 and 10 to a movie, with 1 corresponding to ëawfuli and ë10i to excellent, the rating of a movie can be decided by taking the average over all votes. Unfortunately, there are obvious limitations in using such a score for accurately measuring the popularity of movies. In particular, different scores may be only reflecting the amount of information about the movies available with voters. Thus, the older, so-called "classic" movies may be judged by a completely different yardstick compared to recently released films in view of the differences in the votersi knowledge about them. Possibly more important from a economic agent's point of view is that as it does not cost the user anything to vote for a movie in the online forums, the vital element of competition for viewers that governs which product/idea will eventually become popular is missing from this measure. Therefore, focusing on the box-office gross earnings of movies after they are newly released in theaters is a reasonable measure of their relative popularity, as the potential viewers have a roughly similar kind of information available about the competing items. Moreover, such ëvoting with oneis walleti is arguably a more honest indicator of individual preference for movies.

An important property to note about the distribution of movie income is that it deviates significantly from a Gaussian form with a much more extended tail. In other words, there are many more highly popular movies than one would expect from a normal distribution. This immediately suggests that the process of emergence of popularity may not be explained simply as the outcome of many agents independently making binary (namely 'yes' or 'no') decisions to adopt a particular choice, such as going to see a particular movie. As this process can be mapped to a random walk, we expect it to result in a Gaussian distribution that, however, is not observed empirically. Previous studies of movie income distribution [104–106] have looked at limited datasets and found some evidence for a power-law fit. A more rigorous analysis using data on a much larger number of movies released across theaters in the USA was performed in Ref. [107]. While the tail of the distribution for both the opening gross and the total gross for movies may appear to follow an approximate power law $P(I) \sim I^{-\alpha}$ with an exponent $\alpha \simeq 3$ [107], an even better fit is achieved with a log-normal form [108],

$$P(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\ln x - \mu)^2/2\sigma^2}, \tag{15}$$

where μ and σ are parameters of the distribution, being the mean and standard deviation of the variableis natural logarithm [see Fig. 6 (a)]. The lognormal form has also been seen in the income distribution of movies released in India and Japan [108]. It is of interest to note that a strikingly similar feature has been observed for the popularity of scientific papers, as measured by the number of their citations, where initially a power law was reported for the probability distribution with exponent 3 but was later found to be better described by a log-normal form [109, 110].

Instead of focusing only on the tail (which corresponds to the top grossing movies), if the entire income distribution is considered, we notice another important property: a bimodal nature. There are two clearly delineated peaks, which correspond to a large number of movies having either a very low income or a very high income, with relatively few movies that perform moderately at the box office. The existence of this bimodality can often mask the nature of the distribution, especially when one is working with a small dataset. For example, De Vany and Walls, based on their analysis of the gross for only about 300 movies, stated that log-normality could be rejected for their sample [111]. However, they had clearly assumed that the underlying distribution can be fitted using a single unimodal form. This assumption was evidently incorrect as evident from the histogram of their data. A more detailed and comprehensive analysis with a much larger dataset shows that the distribution of the total (as well as the opening) gross is in fact a superposition of two different log-normal distributions [108].

4.3 Log-normal nature of economic performance

To understand the origin of the bimodal log-normal distribution of the gross income for movies one can of course assume that this is directly related to the intrinsic quality of a movie or some other attribute that is intimately connected to a specific movie (such as how intensely a film is promoted in the media prior to its release). In the absence of any other objective measure of the *quality* of a movie, we can use its production budget as an indirect indicator because movies with higher budget would tend to have more well-known actors, better visual effects and, in general, higher production standards. However, empirically we note that although, in general, movies with higher production budget do tend to earn more, the correlation is not very high (the correlation coefficient r is only 0.63). Thus, production budget by itself is not enough to guarantee economic success. Another possibility is that the immediate success of a movie after its release is dependent on how well the movie-going public have been made aware of the film by pre-release advertising through various public media. Ideally, an objective measure for this could be the advertising budget of the movie. However, as this information is mostly unavailable, one can use instead data about the number of theaters that a movie is initially released at. As opening a movie at each theater requires organizing publicity for it among the neighboring population and wider release also implies more intense mass-media campaigns, we expect the advertising cost to roughly scale with the number of opening theaters. Unfortunately, the correlation between this quantity and per theater movie income is essentially non-existent. In this context, one may note that De Vany and Walls have looked at the distribution of movie earnings and profit as a function of a variety of variables, such as genre, ratings, presence of stars, etc, and have not found any of these to be significant determinants in movie performance [106].

In fact, the bimodal log-normal nature appears as a result of two independent factors, one responsible for the log-normal form of the component distributions and the other for the bimodal nature of the overall distribution. First, turning to the log-normal form, we observe that it may arise from the nature of the distribution of gross income of a movie normalized by the number of theaters in which it is being shown. The income per theater gives us a more detailed view of the popularity of a movie, compared to its gross aggregated over all theaters. It allows us to distinguish between the performance of two movies that draw similar numbers of viewers, even though one may be shown at a much smaller number of theaters than the other. This implies that the former is actually attracting relatively larger audiences compared to the other at each theater and hence is more popular locally. Thus, the less popular movie is generating the same income simply on account of it being shown in many more theaters, even though fewer people in each locality served by the cinemas may be going to see it. The appearance of the log-normal distribution may not be surprising in itself, as it is expected to occur in any linear multiplicative stochastic process. The decision to see a movie (or not) can be considered to be the result of a sequence of independent choices, each of which have certain probabilities. Thus, the final probability that an individual will go to the theater to watch a movie is a product of each of these constituent probabilities, which implies that it will follow a log-normal distribution. It is worth noting here that the log-normal distribution also appears in other areas where the popularity of different entities arises as a result of collective decisions, e.g. in the context of proportional elections [112], citations of scientific papers [110, 113] and visibility of news stories posted by users on an online website [114].

4.4 Bimodality of success and failure

Turning now to the bimodality in the income distribution, this appears to be related to an observed bimodality in the distribution of the the number of theaters in which a motion picture is released [see Fig. 6 (b)]. Thus, most movies are shown either at a handful of theaters, typically a hundred or less (these are usually the independent or foreign movies), or at a very large number of cinema halls, numbering a few thousand (as is the case with the products of major Hollywood studios). Unsurprisingly, this also decides the overall popularity of the movies to an extent, as the potential audience of a film running in less than 100 theaters is always going to be much smaller than what we expect for blockbuster films. In most cases, the former will be much smaller than the critical size required for generating a positive word-of-mouth effect spreading through mutual acquaintances, which will gradually cause more and more people to become interested in seeing the film. There are occasional instances where such a movie does manage to make the transition successfully, when a major distribution house, noticing an opportunity, steps in to market the film nationwide to a much larger audience and a ësleeper hiti is created. An example is the movie *My Big Fat Greek Wedding*, which opened in only 108 theaters in 2002 but went on to become the fifth highest grossing movie for that year, running for 47 weeks and at its peak being shown in more than 2000 theaters simultaneously.

Bimodality has also been observed in other popularity-related contexts, such as in the electoral dynamics of US Congressional elections, where over time the margin between the victorious and defeated candidates has been growing larger [115]. For instance, the proportion of votes won by the Democratic Party candidate in the federal elections has changed from about half of all votes cast to one of two possibilities: either about 35n40% (in which case the candidate lost) or about 60n65% (when the candidate won). This can be explained using a theoretical framework for describing how collective decisions arise from individual binary choice behavior [116, 117]. Here, individual agents take ëyesi or ënoi decisions on issues based on information about the decisions taken by their neighbors and are also influenced by their own previous decisions (adaptation) as well as how accurately their neighborhood had reflected the majority choice of the overall society in the past (learning). Introducing these effects in the evolution of preferences for the agents lead to the emergence of two-phase behavior marked by transition from a unimodal behavior to a bimodal distribution of the fraction of agents favoring a particular choice, as the parameter controlling the learning or global feedback is increased [116]. In the context of the movie income data, we can identify these choice dynamics as a model for the decision process by which theater owners and movie distributors agree to release a particular movie in a specific theater. The procedure is likely to be significantly influenced by the previous experience of the theater and the distributor, as both learn from previous successes and failures of movies released/exhibited by them in the past, in accordance with the assumptions of the model. Once released in a theater, its success will be decided by the linear multiplicative stochastic process outlined earlier and will follow a log-normal distribution. Therefore, the total or opening gross distribution for movies may be considered to be a combination of the lognormal distribution of income per theater and the bimodal distribution of the number of theaters in which a movie is shown.

4.5 Power law decay of income with time

To go beyond the simple blockbusternsleeper distinction and have a detailed view of the time evolution of movie performance, one has to consider the trend followed by the daily or weekly income of a movie over time. This shows an exponential decay with a characteristic rate, a feature seen not only for almost all other blockbusters, but for bombs as well (the rate is different for different movies). The only difference between blockbusters and bombs is in their initial, or opening, gross. However, sleepers may behave differently, showing an initial increase in their weekly gross and reaching the peak in the gross income several weeks after release. For example, in the case of *My Big Fat Greek Wedding* (referred earlier) the peak occurred 20 weeks after its initial opening. It was then followed by exponential decay of the weekly gross until the movie was withdrawn from circulation.

Instead of looking at the income aggregated over all theaters, if we consider the weekly gross income per theater, a surprising universality is observed. As previously mentioned, the income per theater gives us additional information about the movieis popularity because a movie that is being shown in a large number of theaters may have a bigger income simply on account of higher accessibility for the potential audience. Unlike the overall gross that decays exponentially with time, the gross per theater of a movie shows a power-law decay in time measured in terms of the number of weeks from its release, W: $g_W \sim W^{-\beta}$, with exponent $\beta \sim 1$ [108]. Thus, the local popularity of a movie at a certain point in time appears to be inversely proportional to the duration that has elapsed from its initial release. This shares a striking similarity with the time evolution of popularity for scientific papers in terms of citations as the citation probability to a paper published t years ago decays approximately as $1/t$ [110]. In a very different context, namely, the decay over time in the popularity of a website (as measured by the rate of download of papers from the site) and that of individual web pages in an online news and entertainment portal (as measured by the number of visits to the page), power laws have also been reported but with different exponents [118, 119]. More recently, the relaxation dynamics of popularity with a power-law decay have been observed for other products, such as book sales from Amazon.com [120] and the

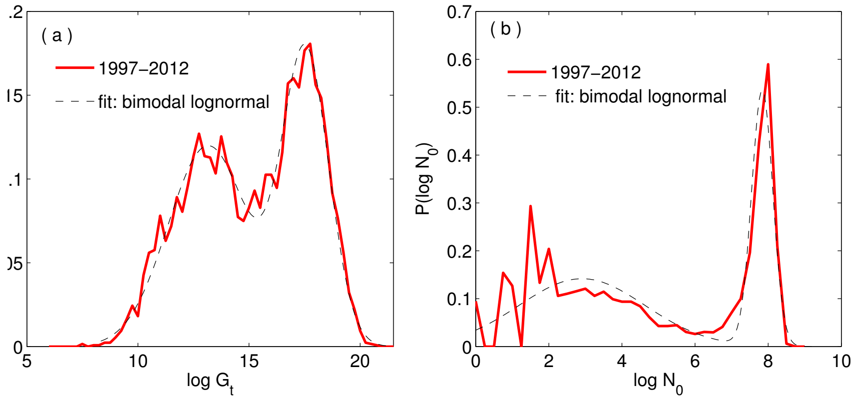


Fig. 6 Distribution of the logarithms of (a) the total gross income of a movie, G_t and (b) the number of theaters in which it opened, N_o, for all movies released in USA during 1997-2012. Fit with bimodal log-normal distributions shows that the empirical data can be well described by this theoretical form.

daily views of videos posted on YouTube [121], where the exponents appear to cluster around multiple distinct classes.

4.6 The stylized facts of "popularity"

Thus, we observe that the complex process of economic success can be understood, at least in the case of movies, in terms of three robust features that (using the terminology of economics) we can term as the *stylized facts of popularity*: (i) log-normal distribution of the success of individual agents (theaters), (ii) the bimodal distribution of the number of agents taking part in a particular round (the theaters in which a movie is shown) and (iii) power-law decay with time of the economic performance of agents (gross income per theater). Some of these features have been seen in areas outside economics in which popularity dynamics play a role, such as citations of scientific papers or political elections. This suggests that it is possible that the above three properties apply more generally to the processes by which a few entities emerge to become a successful product or idea. Possibly a unifying framework may be provided by understanding successful or popular entities as those which have repeatedly survived a sequential failure process [108].

5 Inter-firms and banks-firms network structures: Empirical studies

5.1 Introduction

Credit-debt relationships among economic agents comprise as a whole large-scale networks of the economic system at nation-wide and global scales. There are different layers in such networks even at the core of real-economic and financial systems. One layer is an arena of real economy, namely supplier-customer links among firms as nodes. The firms activities are financed from financial institutions as well as directly from financial markets. The layer of supplier-customer network is thus linked to another layer of financial network between firms and banks. Furthermore, the banks are also creditors and debtors of themselves comprising another layer of inter-banks network.

As a financial system, the inter-bank network resides at the core, which is connected with firms, via banks-firms network, at a periphery of the system; the periphery is a large network of supplier and customer for the engine of real economy. These networks are actually further linked to financial markets, but one may depict the basic picture in a way given in Fig. 7.

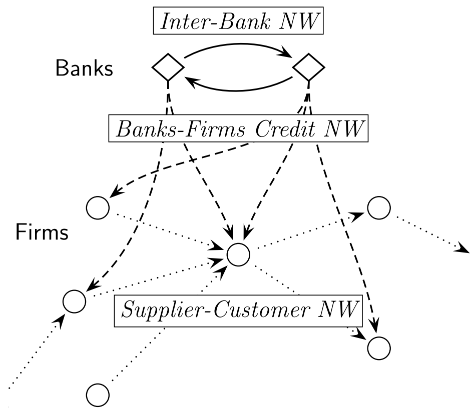


Fig. 7 Inter-bank, banks-firms credit, and supplier-customer networks schematically depicted. Financial institutions or banks (squares) and firms (circles) are creditors and debtors in the links of inter-bank credit (lines), lending-borrowing between banks-firms (dashed lines), and supplier-customer links among firms (dotted lines).

Systemic risk is a network effect caused by failures or financial deterioration of debtors and creditors through the credit-debt links to other nodes even in a remote part of the networks (see [139] and references therein). The systemic risk often has considerable consequences at a nation-wide scale, and sometimes to a world-wide extent, as we experience today in repeated financial crises.

While the understanding of inter-bank network at the core of financial system is crucial (see [125, 129, 135, 140, 134, 124] and references therein for a surge of research focusing on inter-bank networks in different countries), no less important is the propagation of risk from the core of banks to the periphery of firms, vice versa, as well as the propagation of risk among firms. Unfortunately, empirical study based on real-data of banks-firms network or supplier-customer network at a large scale is still lacking. Only recently, there are literature in economics including the studies on US trade network among sectors using input-output (IO) data [123], propagation of sectoral shocks through the IO network [131], US inter-sectoral trade [128], for example; see also reviews [127, 139] and conference reports [122, 136] for collaborative works between physics and economics.

This paper reviews recent empirical studies in Japan on banks-firms lending-borrowing credit network including all financial institutions and listed firms for decades [130, 132], and on supplier-customer network covering a million firms and all bankruptcies in a year [133]. We present new materials here in addition to the description on the unique and exhaustive properties of the Japanese data, but mainly focused is to review recent availability of large-scale networks at nation-wide scale, which can potentially open a new empirical and theoretical studies.

5.2 Banks-firms Credit Network

5.2.1 Data of Lending-Borrowing between Banks and Firms

The dataset is based on a survey of firms quoted in the Japanese stock-exchange markets (Tokyo, Osaka, Nagoya, Fukuoka and Sapporo, in the order of market size). The data were compiled from the firms' financial statements and survey by Nikkei. They include the information about each firm's borrowing obtained from financial institutions such as the amounts of borrowing from 1980 to 2012 including the years of Japanese financial crisis in the late 90s.

For financial institutions, we select commercial banks as a set of leading suppliers of credit. The set comprises long-term, city, regional (primary and secondary), trust banks, insurance companies and other institutions including credit associations. During the examined period, more than 200 commercial banks existed, while the number of listed firms is more than 1600.

5.2.2 Bipartite Network Structure

Annual snapshot of the lending-borrowing network can be regarded as a bipartite graph. Nodes are either banks or firms[9]. Banks and firms are denoted by Greek letters β ($\beta = 1, \ldots, n$) and Latin letters f ($f = 1, \ldots, m$) respectively. n is the number of banks, and m is that of firms. An edge between a bank β and a firm f is defined to be present if there is a credit relationship between them. In addition, a positive number $x_{\beta f}$ is associated with the edge, which is defined to be the amount of the credit. We can depict the network as shown in Fig. 8.

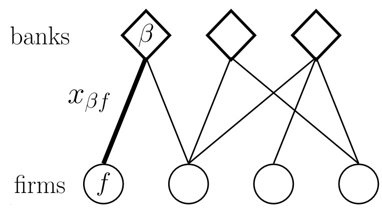


Fig. 8 Credit network as a bipartite graph. An edge connecting between bank β and firm f is associated with an amount of credit $x_{\beta f}$ as a weight.

$x_{\beta f}$ is the amount of lending by bank β to firm i, which precisely equals to the amount of borrowing by firm i from bank β. The total amount of lending by bank

[9] Note that banks are not included in the side of firms, even if they are borrowing from other banks. Because our dataset includes banks' borrowing only partially, the interbank credit is not considered here, though it is no less important than the bank-firm credit studied here.

β is

$$x_\beta = \sum_f x_{\beta f} \ , \tag{16}$$

and the total amount of borrowing by firm f is

$$x_f = \sum_\beta x_{\beta f} \ . \tag{17}$$

The distributions for the amount of credit, x_β, x_f, and the number of borrowers and lenders, denoted by k_β, k_f respectively, have long-tails. They are shown, for the data of credit relationships in the year 2011, in Fig. 9 (a) to (d). There is a significant correlation between w_β and k_β in a natural way, and also for w_f and k_f, as shown in Fig. 9 (e) and (f) respectively. In particular, from the Fig. 9 (e), we can observe an empirical relation of $k_\beta \propto w_\beta^a$, where $a \approx 0.67 \pm 0.04$ (least-square fit; error 95% level). This implies the relation of $w_\beta / k_\beta \propto k_\mu^{0.49 \pm 0.07}$ meaning that the average loan is larger for the larger degree k_μ, or roughly speaking, for the larger banks.

Important properties of the large-scale structure of banks-firms network can be summarized as follows:

- Long-term and city banks are lenders to a number of listed firms. Calculation of Herfindahl index, defined by the sum of squares of shares in the lending amount, shows that they typically lend to 100 firms in the set of 1,600 firms.
- Regional banks have much narrower scope of lending, typically a tenth of long-term and city banks; the lending patterns are closely related to geographical regions.
- From a similarity measure defined by lending patterns of banks, one can construct a minimum-spanning tree (MST), for example (see also [137]). The resulting MST reveals a community structure between banks and firms, the modules of which are related to geographical locations and historical developments of the financial institutions.

See also [130].

5.2.3 Distress Propagation on the Firms-Banks Network

We note that a same value $x_{\beta f}$ has different meanings as a weight to the bank β and to the firm f. For example, even if 90% of the total lending of the bank β goes to the firm f, it may be the case that f depends on β by only 20% for all the loans from banks. It is therefore natural to define $w_{f \to \beta} = x_{\beta f}/x_\beta$, which represents the relative weight of lending by bank β to firm f, or dependency of β on f. Similarly, we define $w_{\beta \to f} = x_{\beta f}/x_f$, which represents the relative amount of borrowing by firm f from bank β, or dependency of f on β. See Fig. 10.

Suppose that bank β has a state of financial deterioration. Then it may shrink the amount of its supplied credit, increase interest-rate, shorten the due time of payment by firms, and so forth. This would eventually influence firm f to an extent that can be quantified by $w_{\beta \to f}$, because it represents the dependency of firm f on bank β for the source of financing. See Fig. 10 (a).

Similarly for the reverse direction of influence, from firms to banks. Firm f with some level of financial deterioration may delay its repayment, have defaults,

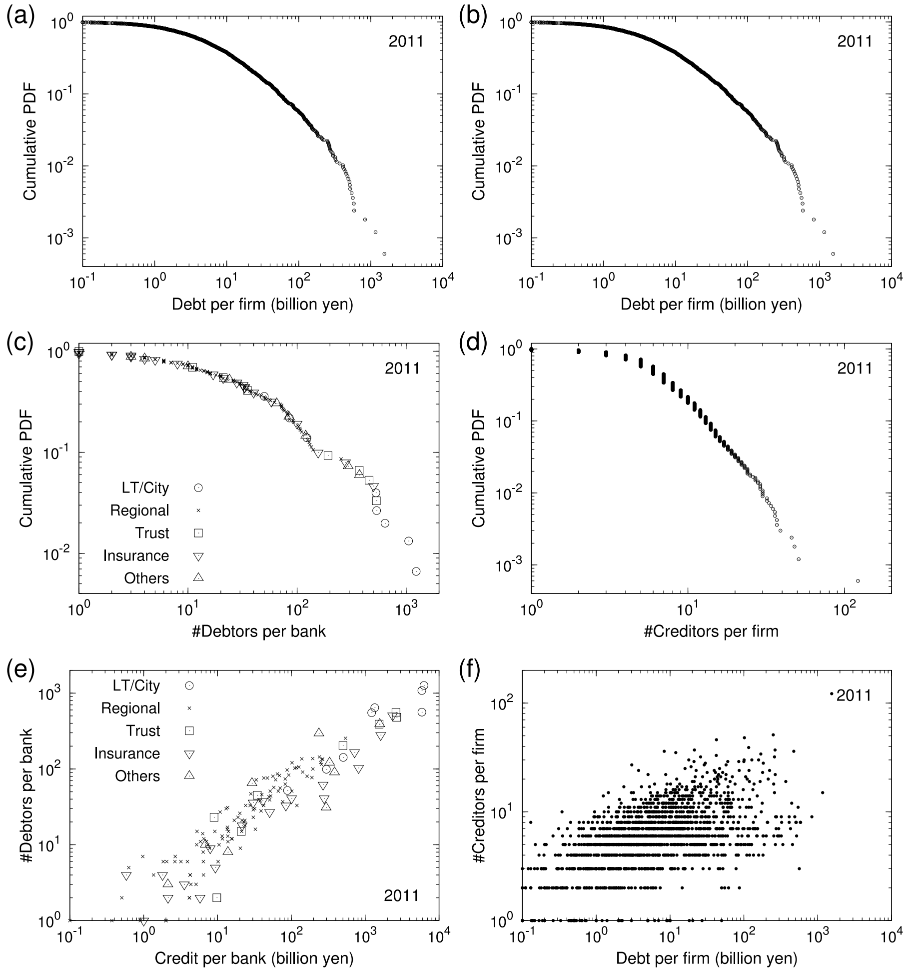
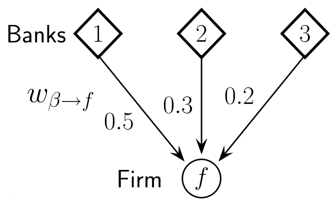
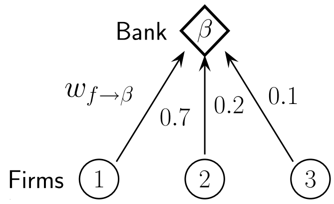


Fig. 9 (a) Cumulative distribution for banks' lending, x_β. (b) For firms' borrowing, x_f. (c) For the number of banks' lending relationships,k_β . (d) For the number of firms' borrowing relationships, k_f. (e) Scatter plot for banks' x_β and k_β. (f) Scatter plot for firms' w_f and k_f. All the plots are for the data in the year 2011. Rank correlations (Kendall's τ) for (e) and (f) are $\tau = 0.814(15.0\sigma)$ and $\tau = 0.384(23.5\sigma)$ respectively (σ calculated under the null hypothesis of statistical independence).



(a) Dependency of firm f on banks β.



(b) Dependency of bank β on firms f.

Fig. 10 Dependency between banks and firms.

even fail into bankruptcy, and so forth. Then the lending banks will not be able to fully enjoy profits in expected amounts due to the delay, may possibly have bad loans partially, if not totally, for the credit given to bankrupted firms. This influence can be quantified by $w_{f \to \beta}$. See Fig. 10 (b).

This consideration can lead one to a methodology of evaluating the level of financial distress that potentially propagate on the network of banks-firms credit. We invented a method based on eigen-vectors and eigen-values structure of the matrices of weights in [132]. By comparing the eigen-structure with that obtained in random bipartite graphs, we found that the largest few (non-trivial) eigenvalues are significant. We performed historical analysis for our datasets, and showed that there are periods when the eigen-structure is stable or unstable, and that a particular set of banks, mostly a few regional banks, have large values of the fragility scores. Drastic change occurs in the late 80s during the bubble and also at the epochs of financially unstable periods including the financial crisis in Japan.

5.3 Production Network among Firms

5.3.1 Data of Large-scale Production Network

Let us say that a directed link is present as $A \to B$ in the production network, where firm A is a supplier to another firm B, or equivalently, B is a customer of A. While it is difficult to record every transaction among suppliers and customers, it is pointless to have a record that a firm buys a pencil from another. Necessary for our study are data of links such that the relation $A \to B$ is crucial for the activity of one or both A and B. If at least one of the firms at either end of a link nominates the other firm as most important suppliers or customers, then the link should be listed. This has a good analogy to a survey of social network, namely "who are your friends important to yourself?"

Our dataset for supplier-customer links has been accumulated on such an idea by one of the leading credit research agencies in Tokyo, which regularly gathers credit information on most of active firms through investigation of financial statements, corporate documents and by hearing-based survey at branch offices located across the nation.

A typical number of active firms in Japan is roughly estimated to be 2 million[10]. We employ a snapshot of production networks compiled in September 2006. In the data, the number of firms is roughly a million, and the number of directional links is more than four million.

5.3.2 Network Structure

The entire network can be represented as a directed graph. To understand the global connectivity, the following graph-theoretical method is useful as was performed in the study of the hyperlink structure of the world-wide web [126].

[10] The National Tax Agency Annual Statistics Report. Other major sources are Establishment and Enterprise Census by the Ministry of Internal Affairs and Communications, and the Ministry of Justice's records on the entry and exit of firms, which are said to have under- or over-estimation problems due to the counting of non-active firms and so forth.

NW= The whole network.
GWCC= Giant weakly connected component: the largest connected component when viewed as an undirected graph. An undirected path exists for an arbitrary pair of firms in the component.
DC= Disconnected components: other connected components than GWCC.
GSCC= Giant strongly connected component: the largest connected component when viewed as a directed graph. A directed path exists for an arbitrary pair of firms in the component.
IN= The firms from which one can reach the GSCC by a directed path.
OUT= The firms that are reachable from the GSCC by a directed path.
TE= "Tendtrils"; the rest of GWCC. Note that TEs may not look like tendrils.

It follows from the definitions that

$$\mathrm{NW} = \mathrm{GWCC} + \mathrm{DC} \tag{18}$$

$$\mathrm{GWCC} = \mathrm{GSCC} + \mathrm{IN} + \mathrm{OUT} + \mathrm{TE} \tag{19}$$

For the benefit of readers, a small example is given in Fig. 11.

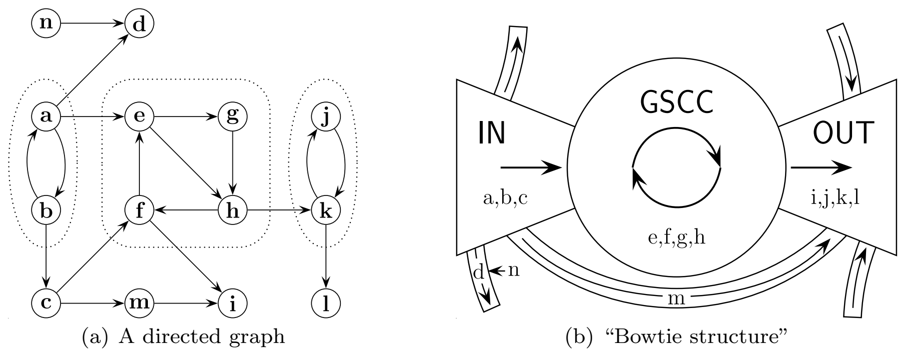


Fig. 11 A simple example of bowtie structure for a graph

The result for the numbers of firms is given as follows:

Component	#firms	Note
NW	1,019,854	
GWCC	1,009,597	99% × NW
DC	10,257	component-size ≤ 4
GSCC	462,563	46% × GWCC
IN	182,018	18% × GWCC
OUT	324,569	32% × GWCC
TE	40,447	4% × GWCC
Total	1,009,597	equal to GWCC

The shortest-path lengths (distances) from the GSCC and firms in the IN and OUT are given by:

Distance from GSCC to IN

Distance	#firms
1	175,855
2	5,949
3	206
4	8
Total	182,018

Distance from GSCC to OUT

Distance	#firms
1	308,572
2	15,441
3	536
4	20
Total	324,569

Any two firms in the network are mutually reachable within 8 directed links as seen from the above table. The inspection of industrial sectors of firms in these components shows that the IN has a relative excess of agriculture, forestry, real estate, and a relative shortage of government-related firms; and that the OUT has a relative excess of medical/health care/welfare, food establishments, and a relative shortage of manufacturing. The majority of manufacturing firms is present in the GSCC.

Important properties of the large-scale structure of production network can be summarized as follows:

- The distributions for the numbers of suppliers/customers (in/out degrees) have long-tails. Denoting the degree by k, the cumulative distributions obey the power-law:

$$P_{>}(k) \sim k^{-\mu} \tag{20}$$

where the exponents $\mu \sim 1.3$.
- There exists a significant positive correlation between the degree and firm-size. Large firms are basically big suppliers/customers, while small and medium enterprises (SME) have relatively smaller numbers of links.
- There exists a weak negative correlation between the degrees at the ends of each link. This means in typical cases of manufacturing sectors that large firms have a number of SMEs as suppliers.
- Transitivity or clustering coefficients, the probability of finding triangles, is small compared with what is expected by random graphs preserving degree-distributions (see [138] for a lucid introduction).
- There exists a hierarchical modular or community structure. For example, manufacturing sectors have communities including electronics, heavy industry, foods, automobiles, construction material, pulp/paper and apparel. The electronics can be recursively divided into sub-communities, which are groups of industrial organization having historical developments and the so-called *keiretsu*, and/or are located in divided geographical sectors. Examples include sub-communities of Hitachi, Fujitsu, NEC; Panasonic, Sharp; Canon, Epson, Nikon, etc.
- Furthermore, these communities can be found to be quasi-cliques in a corresponding bipartite graph as follows. A supplier-customer link $u \to v$ for a set of nodes V $(u, v \in V)$ can be considered as an edge in a bipartite graph that has exactly two copies of V as V_1 and V_2 ($u \in V_1$ and $v \in V_2$). Large and competing firms quite often share a set of suppliers to some extent, depending on the industrial sectors, geographical locations and so on. For example, Honda (v_1), Nissan (v_2) and Toyota (v_3) possibly have a number of suppliers u_i of mechanical parts, electronic devices, chassis and assembling machines, etc., in common. Then the links form a clique or a quasi-clique in the bipartite graph, where most possible links from u_i to v_1, v_2, v_3, ... are present. This forms a portion in the entire graph with a higher density than other portions, which is basically the community structure in the production network.

We refer the readers to the reference [133] and tables and figures therein.

5.3.3 Chain of Bankruptcies

Supplier-customer link is a credit relation [141]. Whenever one delivers goods to others without an immediate exchange of money or goods of full value, credit is extended. Frequently, suppliers provide credit to their customers, who supply credit to their customers and so forth. Also customers can provide credit to their suppliers so as to have them produce an abundance of intermediate goods beforehand. In either case, once a firm goes into financial insolvency state, its creditors will possibly lose the scheduled payment, or goods to be delivered that have been necessary for production. The influence propagates from the bankrupted customer to its upstream in the former cases, and similarly from the bankrupted supplier to its downstream in the latter cases. Thus a creditor has its balance-sheet deteriorated in accumulation, and may eventually go into bankruptcy. This is an example of a *chain of bankruptcy*.

A bankruptcy chain does not occur only along the supplier-customer links. Ownership relation among firms is another typical possibility for such creditor-debtor relationship. It is, however, also frequently observed in our dataset that supplier-customer links are also present between holding and held companies, and sibling and related firms. We assume that most relevant paths along which the chain of bankruptcy occurs are the creditor-debtor links of the production network.

Corresponding to the snapshot of the network taken in September 2006, we employ an exhaustive list of all the bankruptcies for exactly one-year period from October. The number of bankruptcies amounts to roughly 0.13 million, daily mean being 30, and includes a few bankruptcies of listed firms. Nearly half of the bankrupted firms, precisely $N_{\mathrm{b}} \equiv 6264$, were present on the network at the beginning and went into bankruptcy during the period. One can define the probability of bankruptcy by

$$p = N_{\mathrm{b}}/N \approx 0.620\% \tag{21}$$

Note that the probability has inverse of time in its physical dimension. A year was chosen for the time-scale so that it should be longer than the time-scale for financial activities of firms, typically weeks and months, and be shorter than that for the change of network itself.

By using these data, we examined the size distribution for chains of bankruptcies, or avalanche-size distribution. We used a method to evaluate the frequencies of accidental chain in randomized networks, and found that the actual avalanche has a heavy tail distribution in its size. Combining with the large-scale properties and heterogeneity in modular structures, we claim that the effect to a number of creditors, non-trivially large due to the heavy tail in the degree distribution, is considerable in the real economy of the nation [133].

5.4 Summary

We briefly review recent empirical studies on financial networks based on large-scale datasets at nation-wide scale in Japan, banks-firms credit network and production network of suppliers and customers. These datasets provide a quite unique opportunity to investigate the structure and dynamics of networks as well as propagation of financial distress on the them. Because the networks are an arena with different levels of economic agents and relationships among them, on which economic activities take place with possible propagation of financial fragility and distress, it is crucial to understand them based on empirical study. We believe that the current and future collaboration with economists, physicists, computer scientists and practitioners in central banks all over the world would be of great value potentially leading to new ways to monitor and control financial crises that we experience more and more frequently today in the complex connected systems of economy.

6 Financial time-series analyses: Wiener processes and beyond

6.1 Introduction

Nowadays, many people accept the conjecture that there exist probabilistic nature behind almost all of events around us. In economic science, Bachelier [142] dealt with the time-series in financial markets using random walk concept which was five years before Einstein's seminal paper on Brownian motion to determine the Avogadro number. Bachelier attempted to describe the up-down movement of the price changing by means of the Chapman-Kolmogorov equation and found that now-called *Wiener process* is a solution of the equation (see e.g. [143] for the details).

To see the Wiener process, let us define X_t as a price of commodity at time t. Then, we assume that the price updates according to the following rule:

$$X_{t+1} = X_t + Y_t \tag{22}$$

where Y_t is an additive white Gaussian noise satisfying $\langle Y_t \rangle = 0$ and $\langle Y_t Y_s \rangle = \sigma^2 \delta_{t,s}$ under the definition: $\langle \cdots \rangle = \int_{-\infty}^{\infty} (\cdots)(dY_t/\sqrt{2\pi}\sigma)\, \mathrm{e}^{-Y_t^2/2\sigma^2}$. Repeating the above recursion relation (22), we obtain the price at time N as a cumulative return as $X_N \equiv \sum_{t=1}^{N} Y_t$, where we set $X_1 = 0$ for simplicity. It is easy for us to show that X_N also obeys a Gaussian with mean zero and the variance $N\sigma^2$. The model described by (22) to generate a time-series is referred to as Wiener process. It should be noted that even if Y_t does not follow a Gaussian, X_N can posses a Gaussian density with mean zero and the variance $N\sigma^2$ if the variance of each independent component Y_t in X_N is finite, namely, $\langle Y_t Y_s \rangle = \sigma^2 \delta_{t,s} < \infty$ in the limit of $N \to \infty$. In other words, the density of the stochastic variable $Z_N \equiv X_N/\sqrt{N}\sigma$ follows

$$\lim_{N \to \infty} P(Z_N) = \mathcal{N}(0,1). \tag{23}$$

This fact is known as *central limit theorem*.

6.2 Empirical evidence and stable distributions

However, several extensive empirical data analysis (see e.g. [144, 1]) revealed that real financial time series does not follow the Wiener process. This means that the return in financial markets is not generated by a Gaussian distribution and especially the empirical distribution possess a heavy tail. In Fig. 12, we show the the probability distribution of the Standard & Poor's (S&P) 500 index observed at various time intervals Δt (namely, $X_{\Delta t} = \sum_{l=1}^{\Delta t} Y_l$ in terms of (22)) [144].

Obviously, these plots do not look like Gaussian distributions and they have much large kurtosis and exhibit heavy tails. From the central limit theorem, this fact

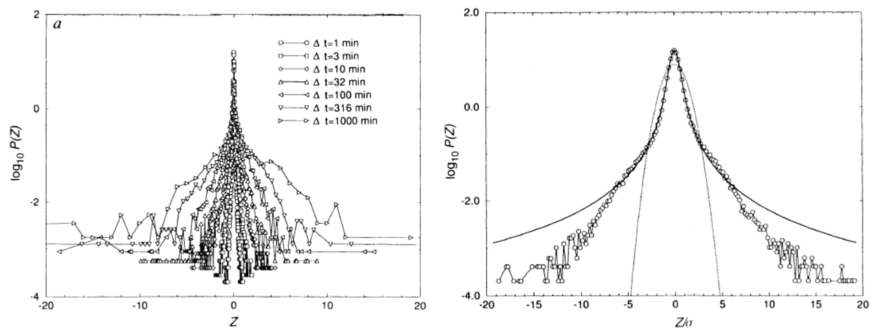


Fig. 12 The left panel shows the probability distribution of the Standard & Poor's (S&P) 500 index observed at various time intervals Δt. From the right panel, we find that the empirical distribution of the left panel is well-described by a Levy stable distribution with $\alpha = 1.4$ and $\gamma = 0.00375$. The broken line is a Gaussian with the same mean 0.0508 as in the empirical distribution of S&P 500 index with $\Delta t = 1$. (The both panels are taken from the reference [144]).

means that the return of the data Y_t does not have a finite variance. Hence, we cannot describe the data in terms of the Wiener process.

In order to describe more generalized stochastic process including the Wiener process, we here introduce the so-called *stable process*. Let us consider that independent stochastic variables $Y_1, Y_2, \cdots, Y_N$ obey the identical distribution $P(y)$. Then, the Fourier transform of the distribution of variable $X(= Y_1 + \cdots + Y_N)$, say, $P_N(X)$ is given by $\phi_N(q) = \{\phi(q)\}^N$, where we defined

$$\phi(q) = \int_{-\infty}^{\infty} P(y) \mathrm{e}^{-iqy} dy. \qquad (24)$$

The inverse transform of $\phi_N(q)$ is immediately written as

$$P_N(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_N(q)\, \mathrm{e}^{iqX} dq. \qquad (25)$$

One can conclude that $P(y)$ is *stable* if $P(y)$ possesses the same function as the $P_N(X)$ does. Especially, for a specific choice $\phi(q) = \mathrm{e}^{-\gamma \Delta t |q|^{\alpha}}$ where Δt is a scaling factor due to the sampling interval, the $P(y)$ leads to

$$P(y : \Delta t) = \frac{1}{\pi} \int_0^{\infty} dq \mathrm{e}^{-\gamma \Delta t |q|^{\alpha}} \cos(qy) \equiv P_L(y : \Delta t) \qquad (26)$$

The $P_L(y : 1)$ is refereed to as *Levy distribution*. We should keep in mind that $P_L(y : 1)$ is identical to a Gaussian when we set $\alpha = 2$, and Lorentzian for $\alpha = 1$. At the tail regime, that is, for $|y| \gg 1$, we have the power-law behavior as $P_L(y : 1) \sim y^{-(\alpha+1)}$.

As shown in Fig. 12 (left), the shape of return distribution is dependent on the sampling intervals Δt. Obviously, if one chooses a large Δt, it is a very rare event to obtain the large $|Y_t|$ and we need huge data points to confirm the shape of the distribution. To avoid this difficulty, we rescale the variables according to Mantegna and Stanley[144], namely, $y_s = y/(\Delta)^{1/\alpha}$, $P_L(y_s : 1) = P_L(y : \Delta t)/(\Delta t)^{-1/\alpha}$. All empirical data having various sampling intervals Δt collapse on the $\Delta t = 1$ distribution by accompanying with the above rescaling with $\alpha = 1.4$ and it is well-described by a Levy distribution $P_L(y) \equiv P_L(y_s : 1)$ as shown in Fig. 12 (right).

6.3 Time-dependent volatility and the prediction models

In the Wiener process, the standard deviation (the volatility in the context of finance) σ is independent of time. However, as empirical analysis for financial time series has revealed, the volatility itself is dependent on time and usually exhibits several distinct behavior, namely, it possess a long memory [1,2]. The long memory is observed through the power-law behavior of the auto-correlation function with respect to the volatility, that is,

$$\overline{\sigma_l \sigma_{l+t}} \equiv \lim_{L\to\infty} \frac{1}{L} \sum_{l=1}^{L} \sigma_l \sigma_{l+t} \sim t^{-\beta} \tag{27}$$

where we defined $\sigma_t^2 \equiv (1/T)\sum_{l=t-1-T}^{t} Y_l^2 - \{(1/T)\sum_{l=t-1-T}^{t} Y_l\}^2$ which is evaluated in the time window with width T assuming the stationarity of Y_t. This is one of the remarkable features of the volatility in comparison with the fact that the auto-correlation function of return Y_t decays exponentially as $\overline{Y_l Y_{l+t}} \sim \mathrm{e}^{-\beta t}$.

For the time-series having the time-dependent volatility, several models to predict the behavior have been proposed. One of the most famous models is refereed to as *ARCH (AutoRegressive Conditional Heteroskedasticity) model* [145] and the simplest version of the model, the so-called ARCH(1) model is given by

$$X_{t+1} = X_t + Y_t, \;\; P(Y_t) = \mathcal{N}(0, \sigma_t), \;\; \sigma_{t+1}^2 = \alpha_0 + \alpha_1 X_t^2 \tag{28}$$

where we should keep in mind that Y_t obeys a Gaussian, however the volatility is not constant but is updated by (28).

The ARCH model is easily extended to the GARCH (Generalized ARCH) model [146]. The update of the volatility in the simplest GARCH(1,1) is described by $\sigma_{t+1}^2 = \alpha_0 + \alpha_1 X_t^2 + \beta_1 \sigma_t^2$ instead of (28).

6.4 Duration between price changes: first-passage process

In the previous subsections, our argument was restricted to the stochastic variables of the price changes (returns) and most of them concern a key-word: *Fat tails* of the distributions or deviation from a Gaussian. However, also the distribution of time intervals can deliver useful information on the markets and it is worth while to investigate these properties extensively [147–153] and if possible, to apply the gained knowledge to financial engineering.

In fact, the *Sony bank rate* is one of the suitable examples. The Sony bank rate is the rate for individual customers of the Sony bank [154] in their on-line foreign exchange trading service via the internet. If the USD/JPY market rate changes

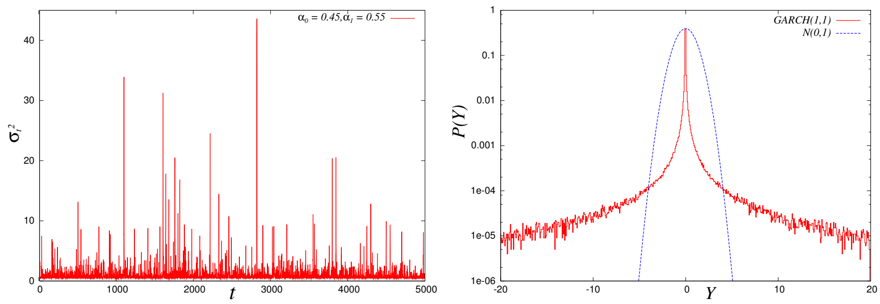


Fig. 13 The time-dependence of the volatility σ_t^2 for setting $\alpha_0 = 0.45, \alpha_1 = 0.55$ (left). The right panel shows the distribution of Y_t for the GARCH(1,1) process with $(\alpha_0, \alpha_1, \beta_1) = (0.4, 0.3, 0.3)$.

by greater or equal to 0.1 yen, the Sony bank USD/JPY exchange rate is updated to the market rate. In this sense, the Sony bank rate can be regarded as a first-passage processes [155–161]. In Fig. 14, we show the mechanism of generating the Sony bank rate from the market rate (this process is sometimes referred to as a *first exit process* [162]). As shown in the figure, the time difference between two consecutive points in the Sony bank rate becomes longer than the time intervals of the market rates. We also should notice that the first passage time fluctuates even if the underlying stochastic process possesses a constant duration.

To qualify the system in terms of the duration, in following, let us suppose that the difference between two consecutive points of the Sony bank rate change, namely, the first-passage time τ follows the distribution with probability density function $P_W(\tau)$[163,164]. Then, the customers observe the rate at time t $(0 \leq t \leq \tau)$ that should be measured from the point at which the rate previously changed. In Fig. 14(right), we show the relation among these variables τ (= first-passage time), t (= observation time) and s (= waiting time) in the time axis. The waiting

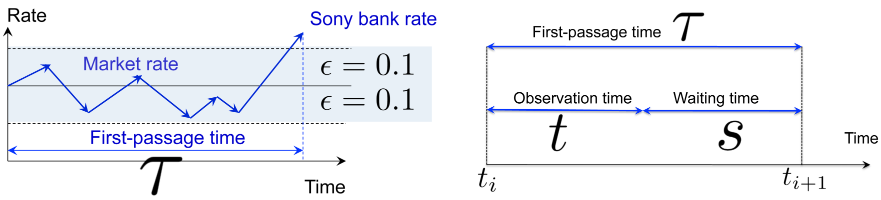


Fig. 14 An illustration of generating the filtered rate by the rate window with width 2ϵ from the market rate. If the market rate changes by a quantity greater or equal to 0.1 yen, the Sony bank USD/JPY exchange rate is updated to the market rate. The right panel shows the relation of the points τ, t and s in time axis. The first-passage time τ is given by $\tau = t_{i+1} - t_i$. The observation time is measured from the point t_i.

time for the customers is naturally defined as $s \equiv \tau - t$. Then, we should notice that the distribution $\Omega(s)$ can be written in terms of the first-passage time distribution (with density $P_W(\tau)$) and the observation time distribution (with density $P_O(t)$)

of the customers as a convolution $\Omega(s) \propto \int_0^\infty d\tau \int_0^\tau dt\, Q(s|\tau,t)P_O(t)P_W(\tau)$. In this equation, the conditional probability density $Q(s|\tau,t)$ that the waiting time takes the value s provided that the observation time and the first-passage time were given as t and τ, respectively, is given by $Q(s|\tau,t) = \delta(s-\tau+t)$, where $\delta(\cdot)$ is Dirac's delta function. Taking into account the normalization constant of $\Omega(s)$, we have

$$\Omega(s) = \frac{\int_0^\infty d\tau P_W(\tau) \int_0^\tau dt\, \delta(s-\tau+t)P_O(t)}{\int_0^\infty ds \int_0^\infty d\tau P_W(\tau) \int_0^\tau dt\, \delta(s-\tau+t)P_O(t)} \tag{29}$$

where t denotes the observation time for the customers. The result of the renewal-reward theorem : $w = \langle s \rangle = E(\tau^2)/2E(\tau)$ (see for example [165,166]) is recovered by inserting a uniformly distributed observation time distribution $P_O(t) = 1$ into the above expression. Indeed, we have

$$w = \langle s \rangle = \int_0^\infty ds s \Omega(s) = \frac{\int_0^\infty ds s \int_s^\infty d\tau P_W(\tau)}{\int_0^\infty ds \int_s^\infty d\tau P_W(\tau)} = \frac{E(\tau^2)}{2E(\tau)} \tag{30}$$

where we defined the n-th moment of the first-passage time $E(\tau^n)$ by $E(\tau^n) = \int_0^\infty ds s^n P_W(s)$. For a Weibull distribution: $P_{m,a}(t) = (mt^{m-1}/a)\exp(-t^m/a)$ which is required from the empirical evidence for the Sony bank rate [167–169], we have

$$\Omega(s) = \frac{m\, \mathrm{e}^{-s^m/a}}{a^{1/m}\Gamma\left(\frac{1}{m}\right)}, \quad w = a^{1/m}\frac{\Gamma\left(\frac{2}{m}\right)}{\Gamma\left(\frac{1}{m}\right)} \tag{31}$$

We show the distribution $\Omega(s)$ with $a = 1$ and $m = 0.59, 1$ and 2 in the left panel of Fig. 15.

Here we encounter the situation which is known as *inspection paradox* [164]. For the Weibull distribution, the paradox occurs for $m < m_c = 1$. Namely, for this regime, we have $\langle s \rangle > \langle \tau \rangle$ (see Fig. 15). In general, it means that the average of durations (first-passage times) is shorter than the average waiting time. This fact

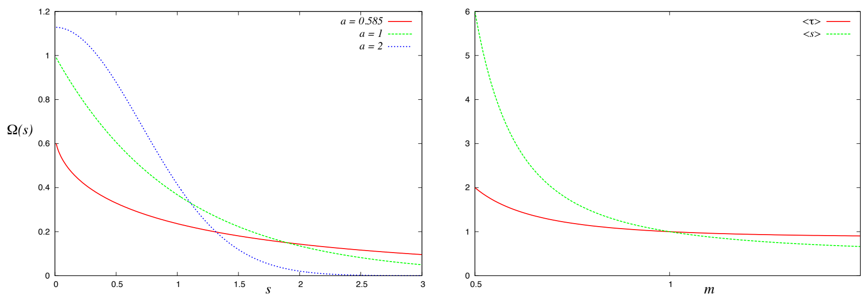


Fig. 15 The distribution of waiting time for a Weibull distributing $\Omega(s)$ with $a = 1$ and $m = 0.59, 1$ and 2. The right panel shows average duration $\langle \tau \rangle$ and average waiting time $\langle s \rangle$ as a function of m for a Weibull duration distribution with $a = 1$. The inspection paradox occurs for $m < m_c = 1$.

is quite counter-intuitive because the customer checks the rate at a time between arbitrary consecutive rate changes. This fact is intuitively understood as follows. When the parameter m is smaller than m_c, the bias of the duration is larger than that of the exponential distribution. As a result, the chance for customers to check the rate within large intervals between consecutive price changes is more frequent than the chance they check the rate within shorter intervals. Then, the average waiting time can become longer than the average duration.

6.5 Microscopic reconstruction of prices

From statistical physics point of view, the price change should be explained from the decision making of huge amount of teaders. Here we show the model proposed by Kaizoji [170] as such an attempt. Recently, we modified the Ising model approach by taking into account the cross-correlations in stocks [171].

As we saw, the return, which is defined as the difference between prices at successive two time steps t and $t+1$ is given by (22). To reconstruct the return Y_t from the microscopic view point, we assume that each trader $(i = 1, \cdots, N)$ buys or sells unit-volume at each time step t and write the total volumes of buying and selling are explicitly given by $\psi_+^{(t)}$ and $\psi_-^{(t)}$, respectively. Then, the return Y_t is naturally defined by means of $\psi_\pm^{(t)}$ as $Y_t = \lambda(\psi_+^{(t)} - \psi_-^{(t)})$, where λ is a positive constant. Namely, when the volume of buyers is greater than that of sellers, $\psi_+^{(t)} > \psi_-^{(t)}$, the return becomes positive $Y_t > 0$ (the price increases from (22)).

We should notice that the making decision of each trader $(i = 1, \cdots, N)$ is obtained simply by an *Ising spin*:

$$S_i^{(t)} = \begin{cases} +1 \text{ (buy)} \\ -1 \text{ (sell)} \end{cases} \tag{32}$$

The return is also simplified as $Y_t = \lambda(\psi_+^{(t)} - \psi_-^{(t)}) = \lambda \sum_{i=1}^{N} S_i^{(t)} \equiv m_t$ where we set $\lambda = N^{-1}$ to make the return:

$$m_t = \frac{1}{N} \sum_{i=1}^{N} S_i^{(t)} \tag{33}$$

satisfying $|m_t^{(t)}| \leq 1$. Thus, m_k corresponds to the so-called *magnetization* in statistical physics, and the update rule of the price is written in terms of the magnetization m_t.

We next introduce the energy function.

$$E_t(\boldsymbol{S}) = -\frac{J_t}{N} \sum_{ij} S_i S_j - h_t^{(k)} \sum_i \sigma_\tau^{(t)} S_i \tag{34}$$

where the above first term induces human collective behaviour, namely, each agent inclines to take the same decision as the others to decrease the total energy. The effect of this first term on the minimization of total energy might be recognized as the so-called *Keynes's beauty contest*. It means that traders tend to be influenced by the others' decision makings, in particular, at the crisis. The second term appearing in the right hand side of (34) represents the cross-correlation between the decision of trader and market information $\sigma_\tau^{(t)}$. Here we choose the 'trends' :

$$\sigma_\tau^{(t)} = \frac{(p_t - p_{t-\tau})}{\tau} \tag{35}$$

for such information. It should be noticed that the state vectors of the agents: $\boldsymbol{S} = (S_1, \cdots, S_N)$ are determined so as to minimize the energy function (34) from the above argument. For most of the cases, the solution should be unique. However, in realistic financial markets, the decisions by agents should be much more 'diverse'. Thus, here we consider statistical ensemble of traders $\boldsymbol{S}$ and define the distribution of the ensemble by $P(\boldsymbol{S})$. Then, we look for the suitable distribution which maximizes the so-called Shannon's entropy $H = -\sum_{\boldsymbol{S}} P(\boldsymbol{S}) \log P(\boldsymbol{S})$ under two distinct constraints $\sum_{\boldsymbol{S}} P(\boldsymbol{S}) = 1$, $\sum_{\boldsymbol{S}} P(\boldsymbol{S}) E(\boldsymbol{S}) = E$. After some algebra, we immediately obtain the solution as *Gibbs-Boltzmann distribution*:

$$P(\boldsymbol{S}) = \frac{\exp[-\beta E(\boldsymbol{S})]}{\sum_{\boldsymbol{S}} \exp[-\beta E(\boldsymbol{S})]} \tag{36}$$

where β stands for the inverse-temperature given by $\beta = 1/T$. The equation of state at the equilibrium is obtained by $m = \sum_{\boldsymbol{S}} (1/N) \sum_i S_i P(\boldsymbol{S})$, however, in financial markets, it might be assumed that the system is not at the equilibrium. To include the non-equilibrium property, we consider that the system described macroscopically by the following update rule which is based on the equation of state for m as

$$m_t = \tanh(J_t m_{t-1} + h_t \sigma_\tau^{(t)}). \tag{37}$$

It should be noted that the magnetization at the equilibrium is obtained by setting $m_t = m_{t-1} = m$ in the limit of $t \to \infty$.

In order to use the update rule (22) with $Y_t = m_t$ and (37), the information about parameters J_t, h_t appearing in the right hand side of (37) is needed. Hence, we should infer these parameters from the past data set in the financial market. In machine learning framework, the parameters are determined by the gradient descent

$$J_{t+1} = J_t - \eta \frac{\partial \mathcal{E}}{\partial J_t}, \; h_{t+1} = h_t - \eta \frac{\partial \mathcal{E}}{\partial h_t} \tag{38}$$

for the cost function:

$$\mathcal{E}(J_t, h_t) \equiv \frac{1}{2} \sum_{l=1}^{t} \left[\overline{\Delta Z_l} - \tanh \left\{ J_t \overline{\Delta Z_{l-1}} + h_t \sigma_\tau^{(t)} \right\} \right] \tag{39}$$

where η is a learning coefficient and we defined $\overline{\Delta Z_l} \equiv (1/M) \sum_{i=l-M+1}^{l} (Z_{i+1} - Z_i)$ for real (empirical) value of the price Z_t. Namely, the cost function (39) is an error measurement to evaluate how good the update rule of the return (37) is satisfied for the empirically observed return $\overline{\Delta Z_l}$ during the past t-time steps. Learning equations for the cost function (39) coupled with (37) and (22) determine the price at the next step. We show the result of the prediction in Fig. 16. After crush in Fig. 16, the parameters J, h converges to $J \to 1$ and $h \to 0$ which are corresponding to the solution $m = \tanh(Jm)$ as a critical point of the second order phase transition.

6.6 Summary

We briefly showed several examples of analysis for the fluctuation in space and time of financial data sets. We found that the modeling of on-line trading system by means of renewal process and its first-passage process is practically useful

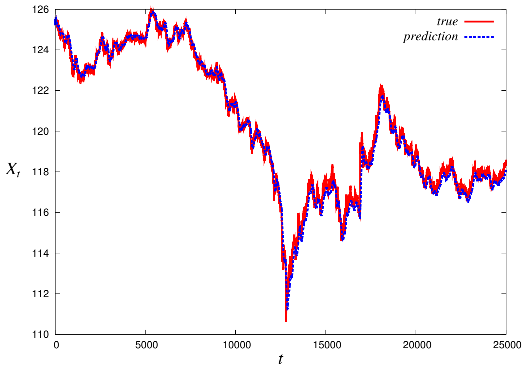


Fig. 16 The result of the prediction. The empirical data (true time-series) is picked-up from EUR/JPY exchange rate from 27th April 2010 to 13th May 2010. We set $\tau = M = 100$ and $\eta = 0.01$.

and one can evaluate several macroscopic quantities such as 'waiting time' for the customers. We also introduced an agent-based modeling of financial markets. We have confirmed that Ising models, which have been used in the literature of statistical physics, are actually effective even in the context of finance. To explain the 'Stylized facts' from the agent-based microscopic viewpoint might be addressed as an important future direction.

7 Outlook

It is well-known that some of the standard assumptions and postulates of traditional economic theory have been:

1. An economy is an "equilibrium" system, where all the markets systematically clear at each point of time; however the equilibrium may be disturbed periodically by exogenous shocks.
2. An "invisible hand" mechanism is at play, where all the selfish and greedy individual agents yield a result that is beneficial to the society as a whole.
3. Agents (individuals, firms, etc.) behave "rationally" – optimise their utilities under specific constraints and their choices satisfy some standard consistency axioms.
4. The behaviour of all the agents together can be treated as corresponding to that of an average or "representative agent".
5. Financial markets are "efficient" such that all the relevant information concerning an asset is reflected by the price of that asset.

However, the last financial crisis and economic slowdown has exposed several limitations of traditional economic theories and models. It has become quite clear that making minor modifications to the standard assumptions or models thereof may not be enough; the whole framework needs to be systematically changed. Firstly, it has been realised that the economy is a complex system and has a network structure, where often minute changes in a certain model's assumptions - or changes in the characteristics of a node or a link - can substantially change the emergent dynamics. This implies that the resulting dynamic interplay can generate unexpectedly large market fluctuations and the inherent risk in any resulting policy-making, is greatly amplified. Secondly, the modelling of "representative agents" in economics often neglects the effects of interaction between those agents. By focusing mainly on "rational behaviour" (individual optimisation of utility or profit), economics has lost the perception that "many is different" – the higher-level aggregate behaviour can have distinct properties that cannot be understood purely on the basis of the constituents at a lower hierarchical level. Built upon this extreme form of reductionism, the established framework of e.g., the banking regulation, has been exclusively micro-oriented and has neglected system-wide effects which can be very important.

In this article, we have presented a few simple models and mechanisms based on statistical mechanics, and shown how they can be adapted to studying socio-economic systems. The knowledge from such models or other studies of self-organised criticality (not discussed here) in the physical sciences, may help in understanding how the structurally similar connections between micro units might lead to similar or emergent collective behaviour in the social systems. We have also discussed in this article that links in economic networks can be interpreted e.g., as relationships between creditors and debtors, or dependency among heterogeneous economic agents as nodes including financial institutions and firms. Once a certain set of agents has deterioration in their financial states, it can potentially propagate as a systemic risk to a considerable scale. It is crucial to understand the large-scale structure of economic networks at nation-wide or globe-wide scales, but such a study has been considered a formidable task so far. We have reviewed the study focusing on banks-firms lending-borrowing credit network and on production network of supplier and customers at a large-scale that covers essentially all domestic banks and firms. There have also been a few other phenomenological studies (which we could not discuss here due to space constraints) of particular segments in the interbank market in the Econophysics literature. It is now necessary to go beyond the first step of analogies, and relatively simple mechanical models, e.g., (i) to examine the behavioural micro-foundations of how the agents involved establish their connections in this financial ecosystem, (ii) to identify and build a proper institutional framework for interactions (with rules and strategies) in order to facilitate beneficial self-organisation, (iii) to understand the globally well-knit, strongly correlated or interdependent socio-economic systems and their resulting complex emergent dynamic behaviour.

It is quite interesting to review how neoclassical economics was influenced by classical physics in earlier times, and now Econophysics has been influenced by statistical physics (kinetic exchange models, Ising model, SOC models, scaling, universality, renormalization, etc.). While statistical physics has been very instrumental in the current development of this field, both from a historical viewpoint and otherwise, it should be noted that Econophysics does not just concern the application of statistical physics. It has gained much from other disciplines as well. The minority game or the Keynes's beauty contest, for example, obviously has an economic origin; the social networks were initiated by sociologists, the random walk formulation of bond prices was formulated first by mathematicians, and so on.

In this article, we have also tried to portray the fact that Econophysics is not just about finance, by choosing many topics outside the realm of finance. Many of the early contributions in Econophysics were very much related to finance, and the phrases "finance", "financial markets", or "speculation" appeared in the titles of the first few books. However, it has constantly and successfully expanded over time from just financial markets to other branches of economics. And probably, in this respect we have seen only the tip of the iceberg!

Acknowledgements

The authors would like to thank all their collaborators and students, whose works have been presented here. AC would like to thank B.K. Chakrabarti, S.R. Chakravarty and A. Chatterjee (the co-authors of the book [7]) for allowing some of the portions of text in Ref. [7] to be used in this manuscript. AG would like to thank B.K. Chakrabarti and S. Biswas for careful reading of the manuscript and useful comments. YF is supported in part by the European Community Seventh Framework Programme (FP7/2007-2013) under Socio-economic Sciences and Humanities, Grant agreement No. 255987 (FOC-II), by the Program for Promoting Methodological Innovation in Humanities and Social Sciences by Cross-Disciplinary Fusing of the Japan Society for the Promotion of Science, by Grant-in-Aid from the Zengin Foundation for Studies on Economics and Finance, and also by Grant-in-Aid for Scientific Research (B) No. 22300080 of the Ministry of Education, Science, Sports and Culture, Japan. JI thanks Enrico Scalas, Naoya Sazuka and Giacomo Livan for critical discussions and useful comments. He also acknowledges the financial support by Grant-in-Aid for Scientific Research (C) of Japan Society for the Promotion of Science, No. 22500195 and No. 25330278.

References

1. R.N. Mantegna and H.E. Stanley, *An Introduction to Econophysics : Correlations and Complexity in Finance*, Cambridge University Press (2000).
2. J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing*, Cambridge University Press (2000).
3. S. Sinha, A. Chatterjee, A. Chakraborti, B. K. Chakrabarti, *Econophysics: An Introduction*, Wiley-VCH, Berlin, 2011.
4. B.M. Roehner, "Patterns of Speculation: A Study in Observational Econophysics", (Cambridge University Press, Cambridge, 2005).
5. U. Garibaldi, E. Scalas, *Finitary probabilistic methods in econophysics* (Cambridge University Press, Cambridge, 2010).
6. H. Aoyama, Y. Fujiwara, Y. Ikeda, H. Iyetomi, and W. Souma, *Econophysics and Companies: Statistical Life and Death in Complex Business Networks*, (Cambridge University Press, New York, 2011).
7. B.K. Chakrabarti, A. Chakraborti, S.R. Chakravarty, and A. Chatterjee, *Econophysics of Income and Wealth Distributions*, (Cambridge University Press, Cambridge, 2013).
8. B.K. Chakrabarti, A. Chakraborti, and A. Chatterjee (Eds) *Econophysics and Sociophysics: Trends and Perspectives*, (Wiley-VCH, Weinheim, 2006).
9. Eds. F. Abergel, H. Aoyama, B.K. Chakrabarti, A. Chakraborti and A. Ghosh, Econophysics of Agent-based models, to be published by Springer-Verlag (Italia), Milan, 2013.
10. Eds. F. Abergel, B.K. Chakrabarti, A. Chakraborti and A. Ghosh, Econophysics of systemic risk and network dynamics (Springer-Verlag (Italia), Milan, 2012).
11. Eds. F. Abergel, B.K. Chakrabarti, A. Chakraborti and M. Mitra, Econophysics of order-driven markets (Springer- Verlag (Italia), Milan, 2011).
12. Eds. L. Pareschi, G. Naldi and G. Toscani, Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences (Birkhauser, Berlin, 2010)

13. Eds. B. Basu, B.K. Chakrabarti, S.R. Chakravarty, K. Gangopadhyay, Econophysics & Economics of Games, Social Choices and Quantitative Techniques (Springer-Verlag (Italia), Milan, 2010)
14. Eds. A. Chatterjee and B.K. Chakrabarti, Econophysics of Markets and Business Networks (Springer-Verlag (Italia), Milan, 2007)
15. Eds. A. Chatterjee and B.K. Chakrabarti, Econophysics of Stock and other Markets (Springer- Verlag (Italia), Milan, 2006)
16. Eds. A. Chatterjee, S. Yarlagadda and B.K. Chakrabarti, Econophysics of Wealth Distributions (Springer-Verlag (Italia), Milan, 2005)
17. V. Yakovenko, J. Barkley Rosser Jr., *Colloquium: Statistical mechanics of money, wealth and income*, Reviews of Modern Physics **81**, 1703 (2009).
18. A. Chakraborti, I. Muni Toke, M. Patriarca and F. Abergel, Quantitative Finance **11:7**, 991 (2011).
19. A. Chakraborti, I. Muni Toke, M. Patriarca and F. Abergel, Quantitative Finance **11:7**, 1013 (2011).
20. A. Sen, *Poverty and famines* (Oxford University Press, New Delhi, 1999).
21. A. Chakraborti and M. Patriarca, *Variational principle for the Pareto power law*, Phys. Rev. Lett. **103**, 228701 (2009).
22. M. Foucault, *Le Pouvoir psychiatrique:Cours au College de France 1973 - 74* (Gallimard,Paris, 2003)
23. R. Scruton, *Thinkers of the new left* (Longman, 1985).
24. J. Rawls, *A Theory of Justice* (Harvard University Press, Cambridge, 1971).
25. P. Aghion, E. Caroli, C. Garca-Pealosa, *Inequality and economic growth: The perspective of the new growth theories*, Journal of Economic Literature **37**, 1615 (1999).
26. R. Barro, *Inequality, growth, and investment*, Working Paper 7038, National Bureau of Economic Research (1999).
27. R. Benabou, *Human capital, inequality, and growth: A local perspective*, European Economic Review **38**, 817 (1994).
28. K. Forbes, *A Reassessment of the Relationship between Inequality and Growth*, American Economic Review **90**, 869 (2000).
29. A. Alesina, D. Rodrik, *Distribution, political conflict, and economic growth: A simple theory and some empirical evidence* in *Political Economy, Growth and Business Cycles*, A. Cukierman, Z. Hercowitz, L. Leiderman (Eds) (MIT Press, Cambridge, 1992).
30. R. Benabou, *Unequal societies: Income distribution and the social contract*, American Economic Review **90**, 96 (2000).
31. A. Alesina, R. Perotti, *Income distribution, political instability, and investment*, European Economic Review **40**, 1203 (1996).
32. J. Blau, P. Blau, *The cost of inequality: Metropolitan structure and violent crime*, American Sociological Review **47**, 114 (1982).
33. V. Pareto, *Cours dfeconomie politique* (Rouge, Lausanne, 1897).
34. R. Gibrat, *Les Ingalites conomiques* (Libraire du Recueil Sirey, Paris, 1931).
35. D. Champernowne, *A Model of Income Distribution*, The Economic Journal **63**, 318 (1953).
36. D. Champernowne, F. Cowell, *Economic Inequality and Income Distribution* (Cambridge University Press, Cambridge, 1998).
37. E. Montroll, M. Shlesinger, *On 1/f noise and other distributions with long tails*, Proceedings of the National Academy of Sciences **79**, 3380 (1982).
38. C. Gini, *Measurement of inequality of incomes*, The Economic Journal **31**, 124 (1921).
39. R. Hogg, J. Mckean, A. Craig, *Introduction to mathematical statistics* (Pearson Education,Delhi, 2007).
40. A. Chatterjee, S. Yarlagadda, and B.K. Chakrabarti, *Econophysics of Wealth Distributions*, (New Economic Windows Series, Springer-Verlag, Milan, 2005).
41. A. Chatterjee, B. Chakrabarti, *Kinetic exchange models for income and wealth distributions*, European Physical Journal B **60**, 135 (2007).
42. H. Aoyama, Y. Nagahara, M. Okazaki, W. Souma, H. Takayasu, M. Takayasu, *Pareto's law for income of individuals and debt of bankrupt companies*, Fractals 8, 293 (2000).
43. S. Sinha, *Evidence for power-law tail of the wealth distribution in India*, Physica A **359**, 555 (2006).
44. P. Richmond, S. Hutzler, R. Coelho, P. Repetowicz, *A Review of Empirical Studies and Models of Income Distributions in Society*, in *Econophysics and Sociophysics: Trends and Perspectives*, Chakrabarti, B.K., Chakraborti, A. and A. Chatterjee (Eds) (Wiley-VCH, Weinheim, 2006), pp. 131-159.

45. A. Salem, T. Mount, *A Convenient Descriptive Model of Income Distribution: The Gamma Density*, Econometrica **42**, 1115 (1974).
46. P.A. Samuelson, *Economics*, (Mc Graw Hill, Auckland, 1998).
47. R. Axtell, *Zipf Distribution of U.S. Firm Sizes*, Science **293**, 1818 (2001).
48. J. Angle, "The surplus theory of social stratification and the size distribution of personal wealth," in *Proceedings of the American Social Statistical Association, Social Statistics Section*, pp. 395–400, Alexandria, VA (1983).
49. J. Angle, "The surplus theory of social stratification and the size distribution of personal wealth," Social Forces **65**, 293–326 (1986).
50. J. Angle, "The statistical signature of pervasive competition on wage and salary incomes," J. Math. Sociol. **26**, 217–270 (2002).
51. J. Angle, "The inequality process as a wealth maximizing process," Physica A **367**, 388–414 (2006).
52. E. Bennati, *La simulazione statistica nell'analisi della distribuzione del reddito: modelli realistici e metodo di Monte Carlo*, ETS Editrice, Pisa (1988).
53. E. Bennati, "Un metodo di simulazione statistica nell'analisi della distribuzione del reddito," Rivista Internazionale di Scienze Economiche e Commerciali **35**, 735–756 (1988).
54. E. Bennati, "Il metodo Monte Carlo nell'analisi economica," Rassegna di lavori dell'ISCO X, 31–79 (1993).
55. S. Ispolatov, P. L. Krapivsky, and S. Redner, "Wealth distributions in asset exchange models," Eur. Phys. J. B **2**, 267–276 (1998).
56. A. Chakraborti, B.K. Chakrabarti, *Statistical mechanics of money: how saving propensity affects its distribution*, European Physical Journal B **17**, 167 (2000).
57. A. Chatterjee, B. Chakrabarti, S. Manna, *Money in gas-like markets: Gibbs and Pareto laws*, Physica Scripta **T106**, 36 (2003)
58. A. Chatterjee, B. Chakrabarti, S. Manna, *Pareto law in a kinetic model of market with random saving propensity*, Physica A **335**, 155 (2004).
59. B. Chakrabarti, A. Chatterjee, *Ideal gas-like distributions in economics: Effects of savings propensity*, in *Applications of Econophysics*, H. Takayasu (Eds), (Springer, Tokyo, 2004), pp. 280-285.
60. M. Patriarca and A. Chakraborti, Am. J. Phys. **81**, 618 (2013).
61. A. Chakrabarti, B.K. Chakrabarti, *Microeconomics of the ideal gas like market models*, Physica A **388**, 4151 (2009).
62. A. Chakraborti, S. Pradhan, and B.K. Chakrabarti, Physica A **297**, pp. 253-259 (2001).
63. M. Patriarca, A. Chakraborti, and K. Kaski, *Statistical model with a standard gamma distribution*, Phys. Rev. E **70**, 016104 (2004).
64. K. Bhattacharya, G. Mukherjee, and S.S. Manna, *Detailed simulation results for some wealth distribution models in econophysics*, in *Econophysics of Wealth Distributions*, A. Chatterjee, S. Yarlagadda, and B.K. Chakrabarti (Eds), pp. 111-119, (Springer, 2005).
65. P. Repetowicz, S. Hutzler, P. Richmond, *Dynamics of money and income distributions*, Physica A **356**, 641 (2005).
66. M. Lallouache, A. Jedidi, A. Chakraborti, *Wealth distribution: To be or not to be a Gamma?*, Science and Culture (Kolkata, India) **76**, 478 (2010).
67. B. Chakrabarti, S. Marjit, *Self-organisation and complexity in simple model systems: Game of life and economics*, Indian Journal of Physics B **69**, 681 (1995).
68. A. Drăgulescu, V. Yakovenko, *Statistical mechanics of money*, European Physical Journal B **17**, 723 (2000).
69. B. Hayes, *Follow the money*, American Scientist **90**, 400 (2002).
70. M. Patriarca, E. Heinsalu, A. Chakraborti, *Basic kinetic wealth-exchange models: common features and open problems*, The European Physical Journal B **73**, 145 (2010).
71. S. Ispolatov, P. Krapivsky, S. Redner, *Wealth distributions in asset exchange model*, European Physical Journal B **2**, 267 (1998).
72. B. Düring, D. Matthes, G. Toscani, *Kinetic Equations modelling Wealth Redistribution: A comparison of Approaches*, Physical Review E **78**, 056103 (2008).
73. G. Toscani, C. Brugna, *Wealth redistribution in Boltzmann-like models of conservative economies* in *Econophysics and Economics of Games, Social Choices and Quantitative Techniques*, B. Basu, S. Chakravarty, B. Chakrabarti, K. Gangopadhyay (Eds), (New Economic Windows, Springer, Milan, 2010), pp. 71-82.
74. T. Lux, *Emergent statistical wealth distributions in simple monetary exchange models: A critical review*, in *Econophysics of Wealth Distributions*, A. Chatterjee, S. Yarlagadda, B. Chakrabarti (Eds) (New Economic Windows, Springer Verlag, Milan, 2005), pp. 51-60.

75. J. Hogan, *There's only one rule for the rich*, New Scientist, p. 21 March (2005).
76. M. Gallegati, S. Keen, T. Lux, P. Ormerod, *Worrying Trends in Econophysics*, Physica A **370**, 1 (2006).
77. T. Lux, F. Westerhoff, *Economics crisis*, Nature Physics **5**, 2 (2009).
78. R. N. Mantegna, H. E. Stanley, *Introduction to econophysics: correlations and complexity in finance*, Cambridge University Press, Cambridge, 1999.
79. D. Challet, M. Marsili and Y. C. Zhang, *Minority Games*, Oxford Univ. Press, Oxford, 2005.
80. W. Brian Arthur, Amer. Econ. Rev. (Papers and Proc.) **84** 406, (1994).
81. D. Challet and Y. C. Zhang, Physica A, **246** 407 (1997).
82. D. Challet and Y. C. Zhang, Physica A, **256** 514 (1998).
83. E. Moro, in *Advances in Condensed Matter and Statistical Mechanics*, Eds. E. Korutcheva and R. Cuerno (Nova Science Publishers, New York, 2004). arXiv:cond-mat/0402651v1 [cond-mat.dis-nn].
84. W. Kets, preprint: arXiv:0706.4432v1 [q-fin.GN].
85. A. S. Chakrabarti, B. K. Chakrabarti, A. Chatterjee and M. Mitra, Physica A **388** 2420 (2009).
86. A. Ghosh, A. S. Chakrabarti, B. K. Chakrabarti, *Econophysics & Economis of Games, Social Choices & Quantitative Techniques*, Eds. B. Basu, B. K. Chakrabarti, S. R. Chakravarty, K. Gangopadhyay, New Economic Windows, Springer, Milan, 3 (2010).
87. A. Ghosh, A. Chatterjee, M. Mitra and B. K. Chakrabarti, New J. Phys. **12** 075033 (2010).
88. A. Ghosh, S. Biswas, A. Chatterjee, A. S. Chakrabarti, T. Naskar, M. Mitra, B. K. Chakrabarti *Econophysics of Systemic Risk & Network Dynamics*, Eds. F. Abergel, B. K. Chakrabarti, A. Chakraborti, A. Ghosh, New Economic Windows, Springer-Verlag Italia, Milan 173 (2013).
89. D. Dhar, V. Sasidevan, B. K. Chakrabarti, Physica A **390**, 3477 (2011).
90. S. Biswas, A. Ghosh, A. Chatterjee, T. Naskar, B. K. Chakrabarti, Phys. Rev. E **85**, 031104 (2012).
91. A. Ghosh, D. De Martino, A. Chatterjee, M. Marsili, B. K. Chakrabarti, (2012), Phys. Rev. E **85**, 021116.
92. P. Ormerod, *Why Most Things Fail* (Faber & Faber, London 2005)
93. R. L. Axtell, Science **293,** (2001) 1818-1820.
94. S. N. Durlauf, Proceedings National Academy Sciences USA **96,** (1999) 10582.
95. S. V. Vikram and S. Sinha, Physical Review E **83,** (2011) 016101.
96. T. C. Schelling, *Micromotives and Macrobehavior* (Norton, New York 1978)
97. S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge 1994)
98. F. Vega-Redondo, *Complex Social Networks* (Cambridge University Press, Cambridge 2007)
99. W. B. Arthur, The Economic Journal **99,** (1989) 116-131.
100. `http://www.slate.com/articles/business/the_dismal_science/1998/01/the_legend_of_arthur.html`
101. `http://www.slate.com/articles/briefing/articles/1998/01/krugmans_life_of_brian.html`
102. A. Ishii, H. Arakaki, N. Matsuda, S. Umemura, T. Urushidani, N. Yamagata and N. Yoshida, New Journal of Physics **14** (2011) 063018.
103. M. Mestyan, T. Yasseri and J. Kertesz, arXiv:1211.0970.
104. A. De Vany and W. D. Walls, Journal of Cultural Economics **23,** (1999) 285.
105. D. Sornette and D. Zajdenweber, European Physical Journal B **8,** (1999) 653.
106. A. De Vany, *Hollywood Economics* (Routledge, London 2003)
107. S. Sinha and S. Raghavendra, European Physical Journal B **42,** (2004) 293-296.
108. R. K. Pan and S. Sinha, New Journal of Physics **12,** (2010) 115004.
109. S. Redner, European Physical Journal B **4,** (1998) 131.
110. S. Redner, Physics Today **58,** (2004) 49.
111. A. De Vany and W. D. Walls, Economic Journal **106,** (1996) 1493.
112. S. Fortunato and C. Castellano, Physical Review Letters **99,** (2007) 138701.
113. F. Radicchi, S. Fortunato and C. Castellano, Proceedings National Academy Sciences USA **105,** (2008) 17268.
114. F. Wu and B. A. Huberman, Proceedings National Academy Sciences USA **104,** (2007) 17599.
115. D. Mayhew, Polity **6,** (1974) 295
116. S. Sinha and R. K. Pan, *Econophysics and Sociophysics: Trends and Perspectives* (Wiley, Weinheim 2006) 417

117. S. Sinha and S. Raghavendra, *Advances in Artificial Economics: The Economy as a Complex Dynamic System* (Springer, Berlin 2006) 177
118. A. Johansen and D. Sornette, Physica A **276,** (2000) 338.
119. Z. Dezso, E. Almaas, A. Lukacs, B. Racz, I. Szakadat A.-L. Barabasi, Physical Review E **73,** (2006) 066132.
120. D. Sornette, F. Deschatres, T. Gilbert and Y. Ageon, Physical Review Letters **93,** (2004) 28701.
121. R. Crane and D. Sornette, Proceedings National Academy Sciences USA **105,** (2008) 15649.
122. F. Abergel, B.K. Chakrabarti, A. Chakraborti, A. Ghosh (eds.): *Econophysics of Systemic Risk and Network Dynamics* (Springer-Verlag, Italia, 2012).
123. D. Acemoglu, A. Ozdaglar and A. Tahbaz-Salehi,v *The network origins of aggregate uctuations*, Econometrica **80** (5), pp. 1977-2016 (2012).
124. S. Battiston, M. Puliga, R. Kaushik, P. Tasca and G. Caldarelli, *DebtRank: Too central to fail? Financial Networks, the FED and Systemic Risk*, Scienttific Reports **2**, Nature Publishing Group (2012).
125. M. Boss, H. Elsinger, M. Summer and S. Thurner, *Network topology of the interbank market*, Quantitative Finance **4** (6), 677-684 (2004).
126. A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, *Graph structure in the Web*, Computer Networks **33** (1-6), 309-320 (2000).
127. G. Caldarelli, *Scale-Free Networks*, (Oxford University Press, 2007).
128. V.M. Carvalho, *Aggregate uctuations and the network structure of intersectoral trade*, Working Paper (2008).
129. R. Cont, A. Moussa and E.B. Santos, *Network structure and systemic risk in banking systems* (2010).
130. G. De Masi, Y. Fujiwara, M. Gallegati, B. Greenwald and J.E. Stiglitz, *An Analysis of the Japanese Credit Network*, Evolutionary and Institutional Economic Review **7** (2), 209-232 (2011).
131. A. Foerster, G.S. Pierre-Daniel and M.W. Watson, *Sectoral versus aggregate shocks: a structural analysis of industrial production*, Journal of Political Economy **119** (1), 1-38 (2011).
132. Y. Fujiwara, H. Aoyama, Y. Ikeda, H. Iyetomi, W. Souma, *Structure and temporal change of the credit network between banks and large firms in Japan*, Economics E-Journal **3** (2009-7) (2009). `http://www.economics-ejournal.org/economics/journalarticles/2009-7`
133. Y. Fujiwara and H. Aoyama, *Large-scale structure of a nation-wide production network*, European Physical Journal B **77** (4), pp. 565-580 (2010).
134. K. Imakubo, Y. Soejima, *Network of fund transaction in call money market*, Monetary and Economic Studies, Bank of Japan **26** (2008).
135. G. Iori, G.De Masi, O. Precup, G. Gabbi and G. Caldarelli, *A Network Analysis of the Italian Overnight Money Market*, To appear in Journal of Economic Dynamics and Control (2007).
136. Latsis Symposium: *Economics on the Move, trends and challenges from the natural sciences* (2012). ETH Zurich; `http://web.sg.ethz.ch/Latsis_2012/main/main.html`; accessible in January 2013.
137. R. Mantegna, *Hierarchical structure in financial markets*, European Physical Journal B **11** (1), pp. 193-197 (1999).
138. M.E.J. Newman, *The structure and function of complex networks*, SIAM Review **45** (2), pp. 167-256 (2003).
139. F. Schweitzer, G. Fagiolo, F. Sornette, F.Vega-Redondo, A. Vespignani, D.R. White, *Economic Networks: the new challenges*, Science **325**, pp. 422-425 (2009).
140. K. Soramäki, M.L. Bech, J. Arnold, R.J. Glass and W.E. Beyeler, *The topology of interbank payment flows*, Physica A **379** (1), pp. 317-333 (2007).
141. J.E. Stiglitz and B. Greenwald, *Towards a New Paradigm in Monetary Economics*, (Cambridge University Press, 2003).
142. L. Bachelier, *Theorie de la speculation*, *Annales Scientifique de l'Ecole Normale Superieure*, **III-17**, pp.21-86 (1900).
143. J. Voit, *The Statistical Mechanics of Financial Markets*, Springer (2001).
144. R.N. Mantegna and H.E. Stanley, *Nature* **376**, pp. 46-49 (1995).
145. R.F. Engle, *Econometrica* **50**, 987 (1982).
146. T. Ballerslev, *Econometrics* **31**, 307 (1986).

147. R. F. Engle and J. R. Russel, *Econometrica* **66**, 1127 (1998).
148. F. Mainardi, M. Raberto, R. Gorenflo and E. Scalas, *Physica A* **287**, 468 (2000).
149. M. Raberto, E. Scalas and F. Mainardi, *Physica A* **314**, 749 (2002).
150. E. Scalas, R. Gorenflo, H. Luckock, F. Mainardi, M. Mantelli and M. Raberto, *Quantitative Finance* **4**, 695 (2004).
151. T. Kaizoji and M. Kaizoji, *Physica A* **336**, 563 (2004).
152. E. Scalas, *Physica A* **362**, 225 (2006).
153. R. Gorenflo and F. Mainardi, *The asymptotic universality of the Mittag-Leffler waiting time law in continuous random walks*, Lecture note at WE-Heraeus-Seminar on Physikzentrum Bad-Honnef (Germany), 12-16 July (2006).
154. http://moneykit.net
155. S. Redner, *A Guide to First-Passage Processes*, Cambridge University Press (2001).
156. N.G. van Kappen, *Stochastic Processes in Physics and Chemistry*, North Holland, Amsterdam (1992).
157. C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences*, Springer Berlin, (1983).
158. H. Risken, *The Fokker-Plank Equation : Methods of Solution and Applications*, Springer Berlin, (1984).
159. I. Simonsen, M.H. Jensen and A. Johansen, *Eur. Phys. J. B* **27**, 583 (2002).
160. S. Kurihara, T. Mizuno, H. Takayasu and M. Takayasu, *The Application of Econophysics*, H. Takayasu (Ed.), pp. 169-173, Springer (2003).
161. J. Inoue and N. Sazuka, *Physical Review E* **76**, 021111 (2007).
162. M. Montero and J. Masoliver, *European Journal of Physics B* **57**, 181 (2007).
163. N. Sazuka, J. Inoue and E. Scalas, *Physica A* **388**, 2839 (2009).
164. J. Inoue and N. Sazuka, *Quantitative Finance* **10**, 121 (2010).
165. H.C. Tijms, *A first Course in Stochastic Models*, John Wiley & Sons (2003).
166. S. Oishi, *Queueing Theory*, CORONA PUBLISHING CO., LTD (in Japanese) (2003).
167. N. Sazuka, *Eur. Phys. J. B* **50**, 129 (2006).
168. N. Sazuka, *Physica A* **376**, 500 (2007).
169. N. Sazuka and J. Inoue, *Physica A* **383**, 49 (2007).
170. T. Kaizoji, *Physica A* **287**, 493 (2000).
171. T. Ibuki, S. Higano, S. Suzuki, J. Inoue and A. Chakraborti, to appear in *J. Phys. Conf. Ser.* (2013).