

## Pareto tails in socio-economic phenomena: a kinetic description

*Stefano Gualandi and Giuseppe Toscani*

### Abstract

Various phenomena related to socio-economic aspects of our daily life exhibit equilibrium densities characterized by a power law decay. Maybe the most known example of this property is concerned with wealth distribution in a western society. In this case the polynomial decay at infinity is referred to as Pareto tails phenomenon (Pareto, *Cours d'économie politique*, 1964). In this paper, the authors discuss a possible source of this behavior by resorting to the powerful approach of statistical mechanics, which enlightens the analogies with the classical kinetic theory of rarefied gases. Among other examples, the distribution of populations in towns and cities is illustrated and discussed.

(Published in Special Issue [Agent-based modelling and complexity economics](#))

**JEL** C02 C6 C68

**Keywords** Kinetic models; Boltzmann-type equations; multi-agent systems; economic modeling equation

### Authors

*Stefano Gualandi*, Department of Mathematics, University of Pavia, Italy,  
[stefano.gualandi@unipv.it](mailto:stefano.gualandi@unipv.it)

*Giuseppe Toscani*, Department of Mathematics, University of Pavia, Italy

**Citation** Stefano Gualandi and Giuseppe Toscani (2018). Pareto tails in socio-economic phenomena: a kinetic description. *Economics: The Open-Access, Open-Assessment E-Journal*, 12 (2018-31): 1–17. <http://dx.doi.org/10.5018/economics-ejournal.ja.2018-31>

Received November 29, 2017 Published as Economics Discussion Paper December 12, 2017

Revised April 9, 2018 Accepted April 23, 2018 Published May 18, 2018

© Author(s) 2018. Licensed under the [Creative Commons License - Attribution 4.0 International \(CC BY 4.0\)](#)

## **1 Introduction**

Probability distribution functions with power decay appear in numerous biological, physical, social and economic contexts, which look fundamentally different. In biology, power laws have been claimed to describe the distributions of the connections of enzymes and metabolites in metabolic networks, the number of interactions partners of a given protein, and other quantities (Karev et al., 2002; Kuznetsov, 2003). Among others, in the physical context power laws appear when studying the cooling of an inelastic dissipative gas by means of the Boltzmann equation (Ernst and Brito, 2002a,b). Likewise, in network analysis power laws imply evolution of the network with preferential attachment, i.e. a greater likelihood of nodes being added to pre-existing hubs (Barabasi, 1999, 2002; Durrett, 2007; Simkin and Roychowdhury, 2011).

As documented in Gabaix (1999); Newman (2005) through an exhaustive list of references, in addition to the most studied phenomena, power-law distributions occur in an extraordinarily diverse range of situations, which range from firms size (Ishikawa, 2009; Fujimoto et al., 2011), the sizes of cities (Gabaix, 1999), the size of earthquakes (Gutenberg and Richter, 1944) to the frequency of use of words in any human language (Estoup, 1916; Zipf, 1949).

Most of these applications are quite recent, and reflect the increasing interest of mathematicians towards applications to living systems (Marsan et al., 2016; Bassetti and Toscani, 2015; Bellomo et al., 2013, 2008; Bellomo and Soler, 2012; Bellouquid et al., 2013; Ben-Naim et al., 2003a,b; Ben-Naim, 2005; Bertotti and Delitala, 2008; Carrillo et al., 2010a,b; Cucker and Mordecki, 2008; Cucker and Smale, 2007a,b).

As a matter of fact, however, maybe the first example of distributions with power decay is concerned with the original study of Pareto on wealth distribution (Pareto, 1964). In reason of its great importance in the development of modern societies, the unequal distribution of wealth in a population described by Pareto, consequent to the formation of power tails, attracted the interest of a number of economists, physicists and mathematicians (Castellano et al., 2009; Düring et al., 2008, 2009; Gupta, 2006; Naldi et al., 2010; Pareschi and Toscani, 2013) (cf. also Bouchaud and Mézard (2000); Burger et al. (2013, 2014); Cordier et al. (2009); Garibaldi et al. (2007); Hayes (2002); Ispolatov et al. (1998); Scalas et al. (2006)). Among the various models present in the literature, mostly based on the approach furnished by statistical mechanics, kinetic models of socio-economic systems gained a lot of popularity (Bisi et al., 2009; Chakraborti, 2002; Chakraborti and Chakraborti, 2000; Chatterjee et al., 2004, 2005; Comincioli et al., 2009; Cordier et al., 2005; Drăgulescu and Yakovenko, 2000; Düring and Toscani, 2008; Maldarella and Pareschi, 2012; Toscani et al., 2013), in reason of the strong analogies between them and the classical kinetic theory of rarefied gases.

As well-known in the scientific literature (cf. for example Simkin and Roychowdhury (2011)), where a reasoned history of the rediscoveries of original works on preferential attachment published one century ago Willis and Yule (1922); Willis (1922); Yule (1925) is dealt with), the kinetic approach provides an alternative way to look at phenomena already discussed from a different point of view. However, in various cases this approach allows to obtain rigorous analytical results on the large-time behavior of the underlying models, including the formation of polynomial tails

at large times, thus giving strong arguments to retain or reject them (Bassetti and Toscani, 2010, 2014; Matthes and Toscani, 2008).

With respect to the classical kinetic theory of rarefied gases, described by the Boltzmann equation (Boltzmann, 1995), where the equilibrium density is found to be a Gaussian (known as Maxwellian distribution (Bobylev, 1988; Cercignani, 1988; Cercignani et al., 1994)), kinetic models of wealth distribution are restricted to a nonnegative wealth variable  $w$ , and the corresponding equilibrium density, known as Pareto-type distribution (Pareto, 1964), is often represented by a curve that exhibits a polynomial decay at infinity. If in equilibrium the wealth in a multi-agent society is distributed according to a probability density  $f(w)$  (cf. the recent discussion in Düring et al. (2017)), the distribution function of wealth, say  $F(w)$  satisfies, for  $w \gg 1$

$$1 - F(w) = \int_w^{+\infty} f(v) dv \cong w^{-p}, \quad p > 1. \quad (1)$$

The value of the positive constant  $p$  is usually called the Pareto index.

The equilibrium density of type (1) makes evident both the unequal distribution of wealth in the society, and the existence of a (small) class of extremely rich people. Various studies of the real data of western economies allowed to conclude that the Pareto index is varying between 1,5 and 3 (data referred to the year 2000: USA  $\sim 1,6$ , Japan  $\sim 1,8 - 2,2$ , Drăgulescu and Yakovenko (2000)). The main consequence is that typically less than the 10% of the population possesses at least the 40% of the total wealth of the country, and follows that law.

Kinetic models of wealth distribution are based on binary trades between agents (Pareschi and Toscani, 2013), and, as noticed in Düring et al. (2008, 2009); Matthes and Toscani (2008), the structure of the microscopic trade is responsible of the macroscopic behavior. These enlightening studies made possible to identify the types of microscopic interactions which lead to stationary distributions with power tails. A kinetic model for knowledge formation in a western society, still with a power tailed steady distribution has been introduced in Pareschi and Toscani (2014). Next a kinetic model for conviction development was studied in Brugna and Toscani (2015). The common features of all these models is that the classes of very rich people, with very high culture and knowledge or a very high degree of conviction are very thin, even if they retain a large part of their respective traits.

In the rest of this note, we discuss in some details this kinetic modeling, and its possible range of application. In addition to the main known example, we will discuss how the basic principles of kinetic modeling can help to clarify the evolution of the size of a population in towns and cities, thus justifying the formation of Pareto tails in this case.

## **2 Learning from Kac's caricature of a Boltzmann gas**

Let us consider a rarefied gas with molecules that interact pairwise. In the one-dimensional situation, the most general binary interaction between particles which is linear in the entering velocities  $(v, w)$  can be described by assigning the exchange rules

$$v^* = p_1 v + q_1 w, \quad w^* = p_2 v + q_2 w; \quad p_i, q_i > 0, \quad i = 1, 2. \quad (2)$$

In (2)  $(p_i, q_i)$ , for  $i = 1, 2$ , are *mixing parameters*, and can be either fixed constants or random variables, depending on the problem under study. Interactions of type (2) are very general, and include, when the mixing parameters are random, most of the known one-dimensional Boltzmann-type models of Maxwell type, including the famous Kac model (Kac, 1959). The mixing parameters in Kac model are given by

$$p_1 = \cos \theta, \quad q_1 = -\sin \theta, \quad p_2 = -q_1, \quad q_2 = p_1. \quad (3)$$

In (3)  $\theta$  is a random variable uniformly distributed on  $(-\pi, \pi)$ . Note that  $p_1^2 + q_1^2 = p_2^2 + q_2^2 = 1$ , which implies conservation of energy in any collision.

Let  $f(v, t)$  denote the density of particles with velocity  $v$  at time  $t > 0$ . Then, the time evolution of the density consequent to interactions of type (2) can be fruitfully written in weak form as

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}} \varphi(v) f(v, t) dv &= \\ &= \frac{\sigma}{2} \left\langle \int_{\mathbb{R} \times \mathbb{R}} (\varphi(v^*) + \varphi(w^*) - \varphi(v) - \varphi(w)) f(v, t) f(w, t) dv dw \right\rangle. \end{aligned} \quad (4)$$

In (4) we denoted by  $\sigma$  a constant related to the collision frequency, and by  $\langle X \rangle$  the mean value of the random quantity  $X$ . Note that equation (4) allows to compute the evolution in time of all observable quantities  $\varphi$ . Let the initial density  $f_0(v)$  satisfy the normalization conditions

$$\int_{\mathbb{R}} f_0(v) dv = 1, \quad \int_{\mathbb{R}} v f_0(v) dv = 0, \quad \int_{\mathbb{R}} v^2 f_0(v) dv = 1. \quad (5)$$

In the case of Kac model, obtained by mixing parameters (3), evaluating the evolution of the solution density  $f(v, t)$  in correspondence to  $\varphi(v) = 1, v, v^2$  shows that the solution satisfies conditions (5) at any subsequent time  $t > 0$ . Also, the equilibrium density can be explicitly found. One can easily check that the equilibrium density (of mass one) is the Gaussian density (Maxwell-type distribution)

$$M(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}. \quad (6)$$

Indeed, the exchange rule (2) in Kac's case is given by a plane rotation around the origin of the point  $P = (v, w)$  of an angle  $\theta$ , to which it corresponds a Jacobian equal to one.

Then, since  $M(v)M(w) = M(v^*)M(w^*)$ , the identity

$$\begin{aligned} \int_{\mathbb{R} \times \mathbb{R}} (\varphi(v^*) + \varphi(w^*)) M(v)M(w) dv dw &= \\ &= \int_{\mathbb{R} \times \mathbb{R}} (\varphi(v) + \varphi(w)) M(v^*)M(w^*) dv dw \end{aligned}$$

gives the result. Note that this identity depends only of the expression of the mixing parameters (3), and consequently it is independent of the law of  $\theta$ .

Consider now in Kac model, for a given constant  $\varepsilon \ll 1$ , the coupling of an increasing value  $\sigma_\varepsilon = \sigma/\varepsilon$  of the collision frequency with a shrinking of the uniform distribution  $\theta$ , say  $\theta_\varepsilon$ , where

$\theta_\varepsilon$  is uniformly distributed on  $(-\sqrt{3\varepsilon}, \sqrt{3\varepsilon})$ . This gives  $\langle \theta_\varepsilon \rangle = 0$ , and  $\langle \theta_\varepsilon^2 \rangle = \varepsilon$ . Moreover  $\langle |\theta_\varepsilon|^3 \rangle \cong \varepsilon^{3/2}$ . Let  $f_\varepsilon(v, t)$  denote the corresponding density solution of equation (4). Expanding in Taylor's series the functions  $\varphi(v^*)$  and  $\varphi(w^*)$  it is a simple exercise to obtain that the weak form of Kac equation can be rewritten as

$$\frac{d}{dt} \int_{\mathbb{R}} \varphi(v) f_\varepsilon(v, t) dv = \sigma \int_{\mathbb{R}} \left( \frac{1}{2} \varphi''(v) - v \varphi'(v) \right) f_\varepsilon(v, t) dv + R_\varepsilon(t),$$

where the remainder  $R_\varepsilon(t)$  tends to zero as  $t \rightarrow 0$  (cf. Toscani (1998) for details). Hence, in the limit  $f_\varepsilon(v, t)$  converges to  $h(v, t)$  solution of the weak form of the linear Fokker–Planck equation (Risken, 1996)

$$\frac{\partial h}{\partial \tau} = \frac{\sigma}{2} \left( \frac{\partial^2 h}{\partial v^2} + \frac{\partial}{\partial v} (vh) \right). \quad (7)$$

Since the stationary solution of Kac model is left unchanged by the scaling procedure, the Fokker–Planck equation (7) has the Gaussian density (6) as stationary solution. Of course this can be verified directly on the Fokker–Planck equation. Also, the same linear equation is obtained in the limit by substituting one of the two densities in (4) with a fixed background density  $g(v)$  satisfying the normalization conditions (5). In fact, it can be easily checked that the linear Fokker–Planck equation obtained in the limit only depends on  $g$  through its moments at the first two orders.

This procedure is well-known under the name of grazing collision limit (Villani, 2002). It corresponds to increase the collision frequency while at the same time the result of a single binary collision is close to leave the incoming velocities unaffected.

One of the main advantages of this procedure is that while maintaining the same equilibrium density, the limit equation is much easier to handle. In other words this asymptotic theory represents a good balance between the microscopic binary collision dynamics, easy to implement from a modeling point of view, and its macroscopic outcome provided by the equilibrium density. Indeed, while this equilibrium density is in general difficult to identify resorting to the bilinear kinetic model, the resulting Fokker–Planck equation always allows to recover explicitly its equilibrium density.

### 3 The case of wealth distribution

The basic model discussed in this section has been introduced in Cordier et al. (2005) within the framework of classical models of wealth distribution in economy, to understand the possible formation of heavy tails, as predicted by the economic analysis of the Italian economist Vilfredo Pareto (Pareto, 1964). This model belongs to a class of models in which the interacting agents are exchangeable. In most of these models an agent's *state* at any instant of time  $t \geq 0$  is completely characterized by his current wealth  $v \geq 0$  (Düring et al., 2008, 2009). When two agents encounter in a trade, their *pre-trade wealths*  $v, w$  change into the *post-trade wealths*  $v^*, w^*$  according to the linear binary rule (2).

Similarly to Kac model, the *mixing parameters*  $p_i$  and  $q_i$  are non-negative random variables. The meaning of these parameters is linked to the economic setting one wants to describe. While

$q_1$  denotes the fraction of the second agent's wealth transferred to the first agent, the difference  $p_1 - q_2$  is the relative gain (or loss) of wealth of the first agent due to market risks. It is usually assumed that  $p_i$  and  $q_i$  have fixed laws, which are independent of  $v$  and  $w$ , and of time. This means that the amount of wealth an agent contributes to a trade is (on the average) proportional to the respective agent's wealth.

In Cordier et al. (2005) the trade has been modeled to include the idea that wealth changes hands for a specific reason: one agent intends to *invest* his wealth in some asset, property etc in possession of his trade partner. Typically, such investments bear some risk, and either provide the buyer with some additional wealth, or lead to the loss of wealth in a non-deterministic way. An easy realization of this idea consists in coupling the saving propensity parameter (Chakraborti, 2002; Chakraborti and Chakrabarti, 2000) with some *risky investment* that yields an immediate gain or loss proportional to the current wealth of the investing agent

$$v^* = (1 - \gamma + \eta_1)v + \gamma w, \quad w^* = (1 - \gamma + \eta_2)w + \gamma v, \quad (8)$$

where  $0 < \gamma < 1$  is the parameter which identifies the saving propensity, namely the intuitive behavior which prevents the agent to put in a single trade the whole amount of his money. In this case

$$p_i = 1 - \gamma + \eta_i, \quad q_i = \gamma \quad (i = 1, 2).$$

The coefficients  $\eta_1, \eta_2$  are random parameters, which are independent of  $v$  and  $w$ , and distributed so that always  $v^*, w^* \geq 0$ , i.e.  $\eta_1, \eta_2 \geq \gamma - 1$ . Unless these random variables are centered, i.e.  $\langle \eta_1 \rangle = \langle \eta_2 \rangle = 0$ , it is immediately seen that the mean wealth is not preserved, but it increases or decreases exponentially (see the computations in Cordier et al. (2005)). For centered  $\eta_i$ ,

$$\langle v^* + w^* \rangle = (1 + \langle \eta_1 \rangle)v + (1 + \langle \eta_2 \rangle)w = v + w, \quad (9)$$

implying conservation of the average wealth. In this case, if the initial density  $f_0(v)$ ,  $v \in \mathbb{R}_+$ , satisfies the normalization conditions

$$\int_{\mathbb{R}_+} f_0(v) dv = 1, \quad \int_{\mathbb{R}_+} v f_0(v) dv = 1, \quad (10)$$

the same conditions are satisfied by the solution  $f(v, t)$  at any subsequent time  $t > 0$ .

Various specific choices for the  $\eta_i$  have been discussed in Matthes and Toscani (2008). The easiest one leading to interesting results is  $\eta_i = \pm r$ , where each sign comes with probability  $1/2$ . The factor  $r \in (0, \gamma)$  should be understood as the *intrinsic risk* of the market: it quantifies the fraction of wealth agents are willing to gamble on. Within this choice, one can display the various regimes for the steady state of wealth in dependence of  $\gamma$  and  $r$ , which follow from numerical evaluation. In the zone corresponding to low market risk, the wealth distribution shows again *socialistic* behavior with slim tails. Increasing the risk, one falls into a *capitalistic*, where the wealth distribution displays the desired Pareto tail. A minimum of saving ( $\gamma > 1/2$ ) is necessary for this passage; this is expected since if wealth is spent too quickly after earning, agents cannot accumulate enough to become rich. Inside the capitalistic zone, the Pareto index decreases from  $+\infty$  at the border with *socialist* zone to unity. Finally, one can obtain a steady wealth distribution

which is a Dirac delta located at zero. Both risk and saving propensity are so high that a marginal number of individuals manages to monopolize all of the society's wealth. In the long-time limit, these few agents become infinitely rich, leaving all other agents truly pauper.

The analysis of Matthes and Toscani (2008) essentially shows that the microscopic interaction (8) considered in Cordier et al. (2005) is such that the kinetic equation (4) is able to describe all interesting behaviors of wealth distribution in a multi-agent society, thus producing various types of equilibria, which are heavily dependent on the details of the trades. This makes the situation completely different from the case described by Kac model, where the equilibrium density is uniquely identified.

This difference is mainly related to the conservation properties of the binary interactions. In Kac model, the mixing parameters  $p_i$  and  $q_i$ ,  $i = 1, 2$  are such that there is pointwise conservation of energy. On the contrary, the mixing parameters of trade (8), as given by (9), only ensure *conservation in the mean* of wealth.

In analogy with Section 2, for a given constant  $\varepsilon \ll 1$ , let us consider the coupling of an increasing value  $\sigma_\varepsilon = 1/\varepsilon$  of the collision frequency with a shrinking of the quantities characterizing the trade (8), where  $\gamma_\varepsilon = \varepsilon\gamma$ , and, by assuming  $\langle \eta_i^2 \rangle = \lambda$ , for  $i = 1, 2$ ,  $\eta_{i,\varepsilon} = \eta_i\sqrt{\varepsilon}$ ,  $i = 1, 2$ . Let  $f_\varepsilon(v, t)$  denote the corresponding density solution of equation (4). Then (cf. the details in Cordier et al. (2005); Pareschi and Toscani (2014)) one shows that in the limit  $\varepsilon \rightarrow 0$   $f_\varepsilon(v, t)$  converges to  $h(v, t)$ , weak solution of the Fokker–Planck equation

$$\frac{\partial h}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial v^2} (v^2 h) + \gamma \frac{\partial}{\partial v} ((v-1)h). \quad (11)$$

It is immediately recognizable that the solutions of equation (11) satisfy the normalization conditions (10), and that there exists a unique stationary solution of unit mass, given by the Gamma-like distribution (Bouchaud and Mézard, 2000; Cordier et al., 2005)

$$h_\infty(v) = \frac{(\mu-1)^\mu \exp\left(-\frac{\mu-1}{v}\right)}{\Gamma(\mu) v^{1+\mu}}, \quad (12)$$

where

$$\mu = 1 + 2\frac{\gamma}{\lambda} > 1.$$

This stationary distribution exhibits a power-law tail for large values of the wealth variable. Moreover, the size of the tails is dependent of the quotient between the saving rate  $\gamma$  and the variance  $\lambda$  of the risk variable.

**Remark 1** *The asymptotic procedure leading from the bilinear kinetic model of Boltzmann type to the Fokker–Planck description is clear. One considers a system of agents in which the frequency of trades is increasing, while the result of the trades is not varying in a significant way the value of the incoming wealths. In this regime, formation of tails prevails. In other words, formation of Pareto tails is in this picture consequence of the fact that the system exhibits a very large number of binary interactions, most of them of grazing nature (gain or loss in each of them is very low). In this regime, the details of the random variables  $\eta_i$ ,  $i = 1, 2$  are not important, and the equilibrium density only retains the mean value  $\lambda$ .*

#### 4 Distribution of knowledge exhibits tails

A Fokker–Planck equation similar to (11) appears when considering the formation of knowledge in a multi-agent society according to a microscopic kinetic model of Boltzmann type (Pareschi and Toscani, 2014).

This description of knowledge formation was introduced in Pareschi and Toscani (2014) with the aim of a better understanding of the possible effects of knowledge in wealth distribution. Indeed, different degrees of knowledge in a society are usually considered as responsible of wealth inequalities.

Let us briefly explain the main motivations about microscopic interactions which determine the individual knowledge, which can be described as a familiarity with someone or something unknown, which can include information, facts, descriptions, or skills acquired through experience or education. Knowledge is in part inherited from the parents, but the main factor that can enrich it is the environment in which the individual grows and lives (see Teevan and Birney (1965)). Indeed, the experiences that produce knowledge can not be fully inherited from the parents, such as the genome, but rather are acquired over a lifetime of several elements of the environment. The learning process is very complicated and produces different results for each individual in a population. Although all individuals are given the same opportunities, at the end of the cognitive process every individual appears to have a different level of knowledge. Also, the personal knowledge is the result of a selection, which leads to retain mostly the notions that the individuals consider important, and to discard the rest. As noticed in Pareschi and Toscani (2014), this aspect of the process of learning has been recently discussed in a convincing way by Eco (2011), one of the greatest philosophers and contemporary Italian writers. In his fascinating lecture, Eco outlines the importance of a drastic selection of the surrounding quantity of information, to maintain a certain degree of ingenuity.

Resorting once more to the legacy of kinetic theory, one assumes that, as the actual velocity of a gas particle is the result of a huge number of collisions, personal knowledge is the result of a huge number of microscopic variations. Each microscopic variation is interpreted as an interaction where a fraction of the knowledge of the individual is lost by virtue of his selection, while at the same time the external background (the surrounding environment) can move a certain amount of its knowledge to the individual. If we quantify the nonnegative amount of knowledge of the individual with  $v \in \mathbb{R}_+$ , and with  $z \in \mathbb{R}_+$  the knowledge achieved from the environment in a single interaction, the new amount of knowledge can be computed using the interaction

$$v^* = (1 - P(v) + \eta)v + P_E(v)z. \quad (13)$$

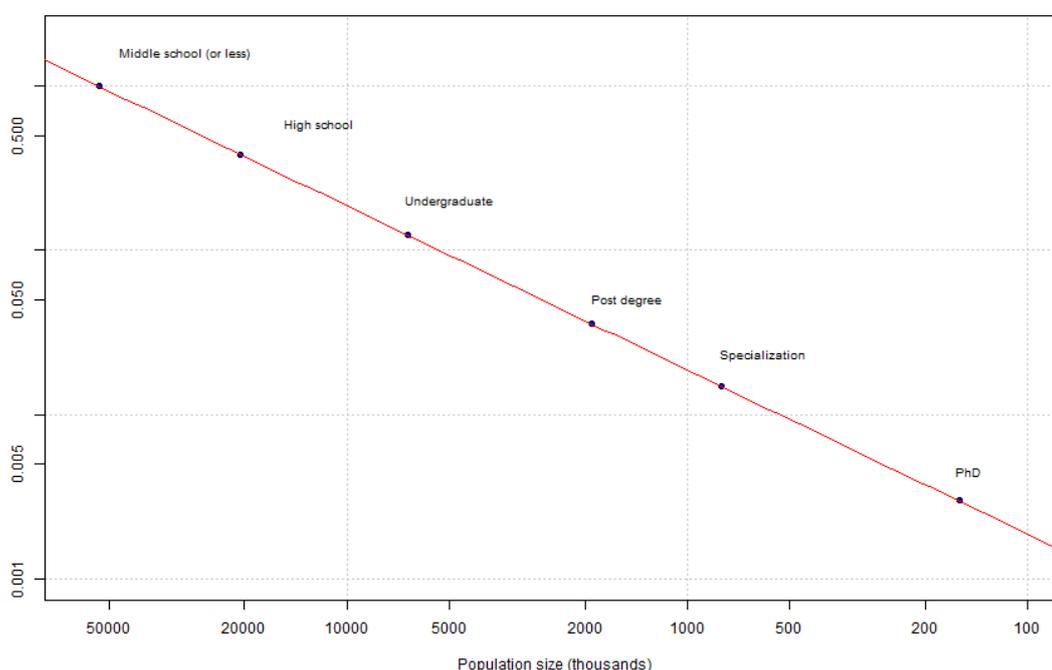
In (13) the functions  $P$  and  $P_E$  quantify, respectively, the amounts of selection and external learning, while  $\eta$  is a random parameter which takes into account the possible unpredictable modifications of the knowledge process. If one assumes that  $\langle \eta^2 \rangle = \lambda$ , and that the average value of the distribution of knowledge in the environment is equal to  $M_E$ , it is immediate to recognize that in the grazing limit procedure described in Sections 2 and 3, the prototype of the Fokker–Planck equation for the density  $k = k(v, \tau)$  of the agents which possess knowledge  $v$  at time  $\tau > 0$  is given by

$$\frac{\partial k}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial v^2} (v^2 k) + \frac{\partial}{\partial v} ((P(v)v - P_E(v)M_E)k). \quad (14)$$

It is clear that in the simple case in which the personal amounts of selection and external knowledge are constant, so that  $P(v) = P$  and  $P_E(v) = P_E$ , the stationary solution of equation (14) is explicitly computable and exhibits Pareto tails (cf. the equilibrium (12)). This is a nice way to say that the model is in agreement with the existence in the society of a (very small) class of learned people.

**Distribution of school knowledge.** Figure 1 shows the distribution of school knowledge in Italy, using data collected from the 2011 census. The plot is based on the data given in Table 1, which shows the number of citizens per type of school degree in the second column, and the (inverse) cumulated number of people for school degree in the fourth column. The first basic level of school knowledge includes every citizen who holds the middle school degree as highest degree, which corresponds nowadays to the Italian mandatory school. The second and third levels include people who got a high school degree and an undergraduate degree as highest degree, respectively. The fourth level includes the 1 124 802 Italians who got a “short” (less than 1 year of studies) post graduate degree. Finally, the last two levels give the number of citizens which hold either a “specialization” or PhD as highest school degree, which are respectively 634 503 and 159 455 citizens, a very small percentage of the whole Italian population.

Indeed, Figure 1 shows that this empirical (inverse) cumulated distribution exhibits a tail.



**Figure 1:** Distribution of school knowledge in Italy by highest degree, 2011 census.

Highest degree	Citizens	(%)	Cumulated values	(%)
Middle school	32 906 278	61.6%	53 422 830	100.0%
High school	13 906 688	26.0%	20 516 552	38.4%
Undergraduate	4 691 104	8.8%	6 609 864	12.4%
Post degree	1 124 802	2.1%	1 918 760	3.6%
Specialization	634 503	1.2%	793 958	1.5%
PhD	159 455	0.3%	159 455	0.3%
(totals)	53 422 830	100.0%		

**Table 1:** Tabular data of the distribution of school knowledge in Italy, 2011 census.

## 5 The size of towns and cities

Among social phenomena that lead to tailed distributions, the size of the populations of towns and cities seems to follow a similar behavior. This interesting observation goes back at least to Felix Auerbach (Auerbach, 1913), who noticed it by studying the distribution of the size of German towns (cf. also Newman (2005); Zipf (1949) for more detailed aspects). The time scale of the evolution of population in towns is of many orders greater than the corresponding one used to analyze wealth distribution in a society. What we see in the present time is the end of a process which started many centuries ago, with a population which increased enormously from the beginning. However, the process which is behind the formation and the size of the modern towns can be reasonably modelled using the same microscopic rules adopted for the formation of knowledge.

Let us consider a simplified situation in the mean value of the population of a country does not vary in a sensible way over years, which corresponds to equate the number of deaths with the number of births over a certain period of time, say  $N$  years. In this way we can essentially relate the size of cities to the phenomenon of migration. Then, if we denote by  $s_n > 0$  the size of the population of a town in a certain year  $n$ , the size of the population of the town at the year  $n + 1$  will result from a balance between the negative variation of the population due to migration towards other towns, and the positive variation of the population due to immigration from the rest of the country. In addition, we assume that there is a random variation of the population, proportional to the population itself and of zero mean, which takes into account the modification of the size of the population due to unpredictable events. Following Gualandi and Toscani (2017), we will assume that the rate of emigration  $E(s_n)$  depends of the size of the city, and it is inversely proportional to the size itself

$$E(s_n) = \frac{1 - \lambda + \lambda s_n}{1 + s_n}, \quad (15)$$

where  $0 \leq \lambda \leq 1$  can be both a constant value, or a random variable. Note that, in the case in which  $\lambda$  is a fixed constant,  $E(s_n)$  decreases when  $\lambda < 1/2$ , and increases when  $\lambda > 1/2$ . Thus, the case  $\lambda < 1/2$  represents a society in which the rate of migration from a city is decreasing from the value  $1 - \lambda$  when the size of the city increases, approaching a constant rate of migration  $\lambda$  as  $s_n \rightarrow \infty$ .

If one introduces this simple and natural rate of change, and a constant immigration rate from the environment, the size of the population after one period of time varies according to

$$s_{n+1} = (1 - E(s_n) + \mu)s_n + I_E z. \quad (16)$$

In (16), the variable  $z \in \mathbb{R}_+$  indicates the amount of population that can be achieved from the rest of the country in the period of time, which one can assume to be distributed according to a certain distribution  $\mathcal{E}(z)$ , with finite mean  $M_E$ . Last, the random variable  $\mu$  such that  $\langle \mu \rangle = 0$  and  $\langle \mu^2 \rangle = \alpha$  measures the random unknown variations of the population. This is a very important parameter, which takes into account for example natural events which can force the population to move away, or the foundation of new factories, that can attract a number of people looking for a job. The grazing limit procedure described in Sections 2, 3 and 4, allows to describe the evolution of the size density  $p(s, \tau)$  of towns with size  $s$  at time  $\tau > 0$  in terms of the Fokker–Planck equation

$$\frac{\partial p}{\partial \tau} = \frac{\alpha}{2} \frac{\partial^2}{\partial s^2} (s^2 p) + \frac{\partial}{\partial s} ((E(s)s - I_E M_E) p). \quad (17)$$

If  $E(s)$  is given by (15), the equilibrium solution of the Fokker–Planck equation (17) is given by the solution to the differential equation

$$\frac{\alpha}{2} \frac{\partial}{\partial s} (s^2 p(s)) + \left( \frac{1 - \lambda + \lambda s}{1 + s} s - I_E M_E \right) p(s) = 0. \quad (18)$$

Solving (18) with respect to the unknown function  $s^2 p(s)$  shows that the equilibrium is a probability density of the form

$$p_\infty(s) = \frac{\kappa}{s^2} \left( \frac{(1+s)^{1-2\lambda}}{s^{1-\lambda}} \right)^{2/\sigma} \exp\left(-\frac{2I_E M_E}{\sigma s}\right), \quad (19)$$

The equilibrium distribution (19) has a polynomial rate of decay at infinity given by

$$1 + \gamma = 2 \left( 1 + \frac{\lambda}{\sigma} \right), \quad (20)$$

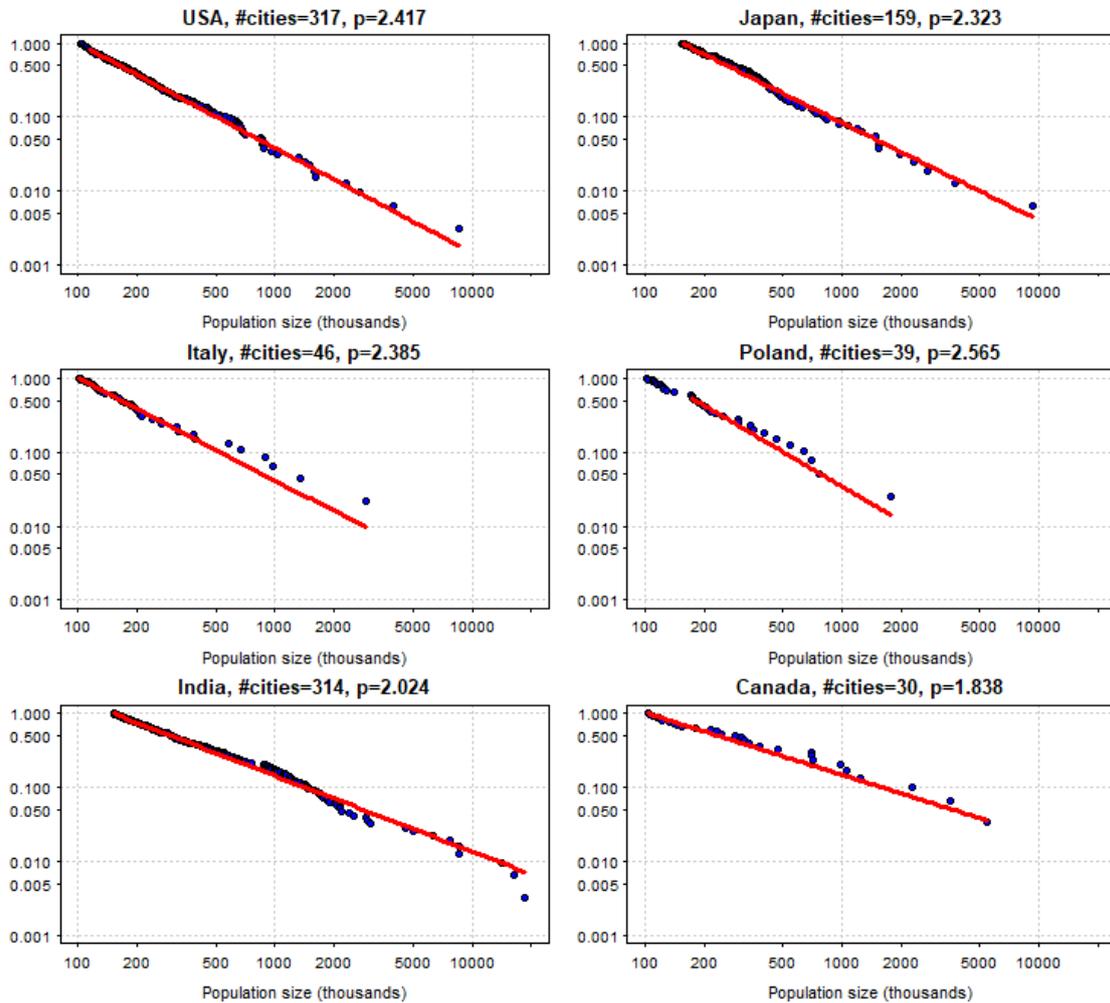
which is related to both the parameters  $\lambda$  and  $\sigma$  denoting respectively the asymptotic value of the rate of migration and the variance of the random migration. Hence, the expected rank  $R(v)$  for large values of  $v$  in this situation is

$$R(s) = \int_s^{+\infty} p(s') ds' \cong s^{-\gamma} = s^{-1-2\lambda/\sigma}. \quad (21)$$

The value  $\gamma = 1$  (Zipf's law) is obtained only for  $\lambda = 0$ , namely in presence of almost no emigration from cities of extremely large size. Among others, as the data that follows clearly show, this behavior seems to be typical of the population of India. The size of population in towns of various countries clearly indicate the presence of tailed distributions. Figure 2 shows the cumulative empirical Pareto distributions in log-log scale of six different countries. For each country the distribution of population size of the largest towns has been collected from the latest official census.<sup>1</sup> We have fitted the data of each country by the method presented in Clauset et al. (2009), using the R implementation therein recommended.<sup>2</sup>

<sup>1</sup> Data collected from <https://citypopulation.de/>. Only for Italy, data have been collected from the Italian National Institute of Statistics (ISTAT) at <http://www.istat.it>

<sup>2</sup> R package for fitting power law distributions: <http://tuvalu.santafe.edu/~aaronc/powerlaws/>



**Figure 2:** Empirical and fitted Pareto distributions of population in towns (in blue the empirical data, and in red the fitted distribution).

## 6 Conclusions

In analogy with some classical argument of kinetic theory, originally developed to establish rigorous relationships between collisional Boltzmann-type and Fokker–Planck type equations in the limit of grazing collisions (Villani, 2002), we introduced and discussed various social and economic phenomena which are characterized by equilibria exhibiting power law tails. In most of these applications, this procedure allows to understand the socio-economic reasons behind the formation of power laws, by relating this phenomenon to few understandable rules characterizing the microscopic interactions. In particular, on the basis of the recent analysis presented in Gualandi and Toscani (2017), a possible microscopic interaction which produces Zipf’s law in the distribution of the size of cities is presented. At difference with the motivations introduced by Gabaix (Gabaix, 1999) to justify this power law behavior, recently criticized in Bee et al. (2013), the equilibrium

distribution found by kinetic arguments seems to cover in a satisfactory way the distribution of all the sizes of the cities (Gualandi and Toscani, 2017).

**Acknowledgements** This work has been written within the activities of GNFM and GNAMPA groups of INdAM (National Institute of High Mathematics), and partially supported by MIUR project “Optimal mass transportation, geometrical and functional inequalities with applications”.

## References

- Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, 59: 74–76.
- Barabasi, A. L. (1999). Emergence of scaling in random networks. *Science*, 286: 509–512.
- Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74: 47–97.
- Bassetti, F., and Toscani, G. (2010). Explicit equilibria in a kinetic model of gambling. *Phys. Rev. E*, 81.
- Bassetti, F., and Toscani, G. (2014). Explicit equilibria in bilinear kinetic models for socio-economic interactions. *ESAIM: Proc. and Surveys*, 47: 1–16.
- Bassetti, F., and Toscani, G. (2015). Mean field dynamics of collisional processes with duplication, loss and copy. *Math. Mod. Meth. Appl. Sci.*, 25: 1887–1925.
- Bee, M., Riccaboni, M., and Schiavo, S. (2013). The size distribution of US cities: Not Pareto, even in the tail. *Economics Letters*, 120(2): 232–237.
- Bellomo, N., Knopoff, D., and Soler, J. (2013). On the difficult interplay between life, “complexity”, and mathematical sciences. *Math. Models Methods Appl. Sci.*, 23: 1861–1913.
- Bellomo, N., Li, N. K., and Maini, P. K. (2008). On the foundations of cancer modelling: selected topics, speculations, and perspectives. *Math. Models Methods Appl. Sci.*, 18: 593–646.
- Bellomo, N., and Soler, J. (2012). On the mathematical theory of the dynamics of swarms viewed as complex systems. *Math. Models Methods Appl. Sci.*, Suppl. 22. 1140006 (29 pages).
- Bellouquid, N., De Angelis, E., and Knopoff, D. (2013). From the modelling of the immune hallmarks of cancer to a black swan in biology. *Math. Models Methods Appl. Sci.*, 23: 949–978.
- Ben-Naim, E. (2005). Opinion dynamics: rise and fall of political parties. *Europhys. Lett.*, 69: 671–677.
- Ben-Naim, E., Krapivski, P. L., and Redner, S. (2003a). Bifurcations and patterns in compromise processes. *Physica D*, 183: 190–204.

- Ben-Naim, E., Krapivski, P. L., Vazquez, R., and Redner, S. (2003b). Unity and discord in opinion dynamics. *Physica A.*, 330: 99–106.
- Bertotti, M. L., and Delitala, M. (2008). On a discrete generalized kinetic approach for modelling persuader's influence in opinion formation processes. *Math. Comp. Model.*, 48: 1107–1121.
- Bisi, M., Spiga, G., and Toscani, G. (2009). Kinetic models of conservative economies with wealth redistribution. *Commun. Math. Sci.*, 7: 901–916.
- Bobylev, A. V. (1988). The theory of the nonlinear spatially uniform Boltzmann equation for Maxwellian molecules. *Sov. Sci. Rev. c*, 7: 111–233.
- Boltzmann, L. (1995). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. In *Kinetische Theorie II*, volume 66, pages 275–370. Berkeley: University of California Press (1964) Translated by S.G. Brush. Reprint of the 1896–1898 Edition. Reprinted by Dover Publ.
- Bouchaud, J. F., and Mézard, M. (2000). Wealth condensation in a simple model of economy. *Physica A.*, 282: 536–545.
- Brugna, C., and Toscani, G. (2015). Kinetic models of opinion formation in the presence of personal conviction. *Phys. Rev. E*, 92. 052818.
- Burger, M., Caffarelli, L., Markowich, P. A., and Wolfram, M. T. (2013). On a Boltzmann-type price formation model. In *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* DOI 10.1098/rspa.2013.0126.
- Burger, M., Caffarelli, L., Markowich, P. A., and Wolfram, M. T. (2014). On the asymptotic behavior of a Boltzmann-type price formation model. *Commun. Math. Sci.*, 12: 1353–1361.
- Carrillo, J. A., Fornasier, M., Rosado, J., and Toscani, G. (2010a). Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 42: 218–236.
- Carrillo, J. A., Fornasier, M., Rosado, J., and Toscani, G. (2010b). Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 42: 218–236.
- Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81: 591–646.
- Cercignani, C. (1988). The Boltzmann equation and its applications. *Springer Series in Applied Mathematical Sciences*, 67.
- Cercignani, C., Illner, R., and Pulvirenti, M. (1994). The mathematical theory of dilute gases. *Springer Series in Applied Mathematical Sciences*, 106.
- Chakraborti, A. (2002). Distributions of money in models of market economy. *Int. J. Modern Phys. C*, 13: 1315–1321.
- Chakraborti, A., and Chakrabarti, B. K. (2000). Statistical mechanics of money: effects of saving propensity. *Eur. Phys. J. B*, 17: 167–170.

- Chatterjee, A., Chakrabarti, B. K., and Manna, S. S. (2004). Pareto law in a kinetic model of market with random saving propensity. *Physica A.*, 335: 155–163.
- Chatterjee, A., Chakrabarti, B. K., and Stinchcombe, R. B. (2005). Master equation for a kinetic model of trading market and its analytic solution. *Phys. Rev. E*, 72. 026126.
- Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51: 661–703.
- Comincioli, V., Della Croce, L., and Toscani, G. (2009). A Boltzmann-like equation for choice formation. *Kinetic Rel. Mod.*, 2: 135–149.
- Cordier, S., Pareschi, L., and Piatecki, C. (2009). Mesoscopic modelling of financial markets. *J. Stat. Phys.*, 134: 161–184.
- Cordier, S., Pareschi, L., and Toscani, G. (2005). On a kinetic model for a simple market economy. *J. Stat. Phys.*, 120: 253–277.
- Cucker, F., and Mordecki, E. (2008). Flocking in noisy environments. *J. Math. Pures Appl.*, 89: 278–296.
- Cucker, F., and Smale, S. (2007a). Emergent behavior in flocks. *IEEE Trans. Automat. Control*, 52: 852–862.
- Cucker, F., and Smale, S. (2007b). On the mathematics of emergence. *Jpn. J. Math.*, 2: 197–227.
- Drăgulescu, A., and Yakovenko, V. M. (2000). Statistical mechanics of money. *Eur. Phys. Jour. B.*, 17: 723–729.
- Düring, B., Georgiou, N., and Scalas, E. (2017). *A stylised model for wealth distribution*, pages 135–157. Singapore: Springer Singapore.
- Düring, B., Matthes, D., and Toscani, G. (2008). Kinetic equations modelling wealth redistribution: A comparison of approaches. *Phys. Rev. E.*, 78. 056103.
- Düring, B., Matthes, D., and Toscani, G. (2009). A Boltzmann-type approach to the formation of wealth distribution curves. *Riv. Mat. Univ. Parma.*, 8: 199–261. Notes of the Porto Ercole School, June 2008.
- Düring, B., and Toscani, G. (2008). International and domestic trading and wealth distribution. *Commun. Math. Sci.*, 6: 1043–1058.
- Durrett, R. (2007). *Random graph dynamics*. Cambridge: Cambridge University Press.
- Eco, U. (2011). Memoria e dimenticanza. In <http://www.3dnews.it/node/1824>.
- Ernst, M. H., and Brito, R. (2002a). High energy tails for inelastic Maxwell models. *Europhys. Lett.*, 58: 182–187.

- Ernst, M. H., and Brito, R. (2002b). Scaling solutions of inelastic Boltzmann equation with over-populated high energy tails. *J. Stat. Phys.*, 109: 407–432.
- Estoup, J.-B. (1916). *Gammes sténographiques*. Institut Sténographique de France.
- Fujimoto, S., Ishikawa, A., Mizuno, T., and Watanabe, T. (2011). A new method for measuring tail exponents of firm size distributions. *Economics: The Open-Access, Open-Assessment E-Journal*, 5: 1–20.
- Gabaix, X. (1999). Zipf’s law for cities: an explanation. *The Quarterly journal of economics*, 114(3): 739–767.
- Garibaldi, U., Scalas, E., and Viarengo, P. (2007). Statistical equilibrium in simple exchange games II. The redistribution game. *Eur. Phys. Jour. B.*, 60: 241–246.
- Gualandi, S., and Toscani, G. (2017). Size distribution in cities: A kinetic explanation. (Preprint).
- Gupta, A. K. (2006). Models of wealth distributions: a perspective. *Econophysics and Sociophysics: Trends and Perspectives*, pages 161–190.
- Gutenberg, B., and Richter, R. (1944). Frequency of earthquakes in California. *Math. Models Methods Appl. Sci.*, 34: 185–188.
- Hayes, B. (2002). Follow the money. *American Scientist*, 90: 400–405.
- Ishikawa, A. (2009). Power-law and log-normal distributions in temporal changes of firm-size variables. *Economics: The Open-Access, Open-Assessment E-Journal*, 3: 1–25.
- Ispolatov, S., Krapivsky, P. L., and Redner, S. (1998). Wealth distributions in asset exchange models. *Eur. Phys. Jour. B.*, 2: 267–276.
- Kac, M. (1959). *Probability and related topics in the physical sciences*. New York Interscience, London.
- Karev, G. P., Wolf, Y. I., Rzhetsky, A. Y., Berezovskaya, F. S., and Koonin, E. V. (2002). Birth and death of protein domains: A simple model of evolution explains power law behaviour. *BMC Evol. Biol.*, 2.
- Kuznetsov, V. A. (2003). Statistics of the numbers of transcripts and protein sequences encoded in the genome. In *Computational and Statistical Approaches to Genomics*, pages 125–171. Springer.
- Maldarella, D., and Pareschi, L. (2012). Kinetic models for socio-economic dynamics of speculative markets. *Physica A.*, 391: 715–730.
- Marsan, G. A., Bellomo, N., and Gibelli, L. (2016). Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics. *Mathematical Models and Methods in Applied Sciences*, 26(06): 1051–1093.

- Matthes, D., and Toscani, G. (2008). On steady distributions of kinetic models of conservative economies. *J. Stat. Phys.*, 130: 1087–1117.
- Naldi, G., Pareschi, L., and Toscani, G. (2010). *Mathematical modeling of collective behavior in socio-economic and life sciences*. Springer Verlag, Heidelberg.
- Newman, M. E. (2005). Power laws, Pareto distributions and Zipf’s law. *Contemporary physics*, 46(5): 323–351.
- Pareschi, L., and Toscani, G. (2013). *Interacting multiagent systems. Kinetic equations and Monte Carlo methods*. Oxford University Press, Oxford.
- Pareschi, L., and Toscani, G. (2014). Wealth distribution and collective knowledge. A Boltzmann approach. *Phil. Trans. R. Soc. A*, 372. 20130396.
- Pareto, V. (1964). *Cours d’économie politique*, volume 1. Librairie Droz.
- Risken, H. (1996). *The Fokker–Planck equation. Methods of solution and applications*. Springer-Verlag. Berlin.
- Scalas, E., Garibaldi, U., and Donadio, S. (2006). Statistical equilibrium in the simple exchange games I. Methods of solution and application to the Bennati–Dragulescu–Yakovenko (BDY) game. *Eur. Phys. J. B.*, 53: 267–272.
- Simkin, M., and Roychowdhury, V. (2011). Re-inventing Willis. *Physics Reports*, 502(1): 1 – 35.
- Teevan, R. C., and Birney, R. C. (1965). *Readings for introductory psychology*. Harcourt, Brace and World, New York.
- Toscani, G. (1998). The grazing collision asymptotics of the non cut-off Kac equation. *Math. Mod. Num. Anal.*, 32: 763–772.
- Toscani, G., Brugna, C., and Demichelis, S. (2013). Kinetic models for the trading of goods. *J. Stat. Phys.*, 151: 549–566.
- Villani, C. (2002). A review of mathematical topics in collisional kinetic theory. *Handbook of Mathematical Fluid Dynamics*, 1(71-305): 3–8.
- Willis, J. (1922). *Age and area: A study in geographical distribution and origin of species*. University Press, Cambridge(UK).
- Willis, J., and Yule, G. (1922). Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature*, 109(2728): 177–179.
- Yule, G. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London B*, 213: 21–87.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison–Wesley, Reading (MA).

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2018-31>

The Editor