

Universal patterns of inequality

Anand Banerjee and Victor M Yakovenko¹

Department of Physics, University of Maryland, College Park,
Maryland 20742-4111, USA

<http://www.physics.umd.edu/~yakovenk/>

New Journal of Physics **12** (2010) 075032 (25pp)

Received 2 December 2009

Published 30 July 2010

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/12/7/075032

Abstract. Probability distributions of money, income and energy consumption per capita are studied for ensembles of economic agents. The principle of entropy maximization for partitioning of a limited resource gives exponential distributions for the investigated variables. A non-equilibrium difference of money temperatures between different systems generates net fluxes of money and population. To describe income distribution, a stochastic process with additive and multiplicative components is introduced. The resultant distribution interpolates between exponential at the low end and power law at the high end, in agreement with the empirical data for the USA. We show that the increase in income inequality in the USA originates primarily from the increase in the income fraction going to the upper tail, which now exceeds 20% of the total income. Analyzing the data from the World Resources Institute, we find that the distribution of energy consumption per capita around the world can be approximately described by the exponential function. Comparing the data for 1990, 2000 and 2005, we discuss the effect of globalization on the inequality of energy consumption.

¹ Author to whom any correspondence should be addressed.

Contents

1. Introduction	2
2. Statistical mechanics and thermodynamics of money	3
2.1. Entropy maximization for division of a limited resource	3
2.2. Flow of money and people between two countries with different temperatures .	5
2.3. Thermodynamics of money and wealth	6
2.4. The circuit of money and the circuit of goods	8
3. Two-class structure of income distribution	9
3.1. Introduction	9
3.2. Income dynamics as a combination of additive and multiplicative stochastic processes	10
3.3. Comparison with the personal income data from the Internal Revenue Service (IRS)	11
3.4. The fraction of income in the upper tail and speculative bubbles	12
3.5. The power-law exponent of the upper tail	15
3.6. The Lorenz plot and Gini coefficient for income inequality	16
4. Probability distribution of the global energy consumption	17
4.1. Introduction	17
4.2. Energy consumption distribution as division of a limited resource	18
4.3. Empirical data analysis	18
4.4. The effect of globalization on the inequality of energy consumption	22
5. Conclusions	23
Acknowledgments	24
References	24

1. Introduction

Two types of approaches are utilized in science to describe the natural world around us. One approach is suitable for systems with a small number of degrees of freedom, such as a harmonic oscillator, a pair of gravitating bodies and a hydrogen atom. In this case, the goal is to formulate and solve dynamical equations of motion of the system, be it within Newtonian, relativistic or quantum mechanics. This approach is widely used beyond physics to study dynamical systems in chemistry, biology, economics, etc. In the opposite limiting case, we deal with systems consisting of a very large number of degrees of freedom. In this case, statistical description is employed, and the systems are characterized by probability distributions. In principle, it should be possible to derive a statistical description from microscopic dynamics, but it is rarely feasible in practice. Thus, it is common to use general principles of the theory of probabilities to describe statistical systems, rather than to derive their properties from microscopic equations of motion. Statistical systems are common in physics, chemistry, biology, economics, etc.

Any probability distribution can be thought of as representing some sort of ‘inequality’ among the constituent objects of the system, in the sense that the objects have different values of a given variable. Thus, a study of probability distributions is also a study of inequality developing in a system for statistical reasons. To be specific, let us consider an economic system with a large number of interacting agents. In the unrealistic case where all agents have exactly

the same values of economic variables, the system can be treated as a single agent called the ‘representative agent’. This approach is common in traditional economics, but by construction it precludes a study of inequality among the agents. However, social and economic inequality is ubiquitous in the real world, and its characterization and understanding are a very important issue.

In this paper, we apply the well-developed methods of statistical physics to economics and society in order to gain insights into probability distributions and inequality in these systems. We consider three specific cases: the distributions of money, income and global energy consumption. In all three cases, the common theme is entropy maximization for partitioning of a limited resource among multiple agents. Despite the difference in the nature of the considered variables, we find a common pattern of inequality in these cases. This approach can also be useful for studying other statistical systems beyond the three specific cases considered in this paper.

Applications of these ideas to money and income have been published in the literature before: see review [1]. To introduce these ideas and to make the paper self-contained, we briefly review the applications to money and income in sections 2 and 3. Section 3 also shows the latest available data for income distribution in 2007, not published before. In section 4, we present a quantitative study of the probability distribution of energy consumption per capita around the world. This is a new kind of study that, to the best of our knowledge, has not appeared before in the literature.

2. Statistical mechanics and thermodynamics of money

2.1. Entropy maximization for division of a limited resource

Let us consider a general mathematical problem of partitioning (dividing) a limited resource between a large number of agents. The solution of this problem is similar to the derivation of the Boltzmann–Gibbs distribution of energy in physics [2]. To be specific, let us apply it to the probability distribution of money in a closed economic system.

Following [3], let us consider a system consisting of N economic agents. At any moment of time, each agent i has a money balance m_i . Agents make pairwise economic transactions with each other. As a result of a transaction, the money Δm is transferred from an agent i to an agent j , so their money balances change as follows:

$$\begin{aligned} m_i &\rightarrow m'_i = m_i - \Delta m, \\ m_j &\rightarrow m'_j = m_j + \Delta m. \end{aligned} \quad (1)$$

The total money of the two agents before and after transaction remains the same

$$m_i + m_j = m'_i + m'_j, \quad (2)$$

i.e. there is a local conservation law for money. It is implied that the agent j delivers some goods or services to the agent i in exchange for the money payment Δm . However, we do not keep track of what is delivered and only keep track of money balances. Goods, such as food, can be produced and consumed, so they are not conserved.

Rule (1) for the transfer of money is analogous to the transfer of energy from one molecule to another in molecular collisions in a gas, and rule (2) is analogous to conservation of energy in such collisions. It is important to recognize that ordinary economic agents cannot ‘manufacture’ money (even though they can produce and consume goods). The agents can only receive money

from and give it to other economic agents. In a closed system, the local conservation law (2) implies the global conservation law for the total money $M = \sum_i m_i$ in the system. In the real economy, M may change due to money emission by the central government or central bank, but we will not consider these processes here. Another possible complication is debt, which may be considered as negative money. Here, we consider a model where debt is not permitted, so all money balances are non-negative $m_i \geq 0$.

After many transactions between different agents, we expect that a stationary probability distribution of money would develop in the system. It can be characterized as follows. Let us divide the money axis m into the intervals (bins) of a small width m_* and label them with an integer variable k . Let N_k be the number of agents with the money balances between m_k and $m_k + m_*$.² Then, the probability to have a money balance in this interval is $P(m_k) = N_k/N$. We would like to find the stationary probability distribution of money $P(m)$, which is achieved in statistical equilibrium.

Because the total money M in the system is conserved, the problem reduces to partitioning (division) of the limited resource M among N agents. One possibility is an equal division, where each agent gets the same share M/N of the total money. However, such an equal partition would be extremely improbable. It is more reasonable to obtain the probability distribution of money from the principle of entropy maximization. Let us consider a certain set of occupation numbers N_k of the money bins m_k . The multiplicity Ω is the number of different realizations of this configuration, i.e. the number of different placements of the agents into the bins preserving the same set of occupation numbers N_k . It is given by the combinatorial formula in terms of the factorials³

$$\Omega = \frac{N!}{N_1! N_2! N_3! \dots} \quad (3)$$

The logarithm of multiplicity is called the entropy $S = \ln \Omega$. In the limit of large numbers, we can use the Stirling approximation for the factorials:

$$S = N \ln N - \sum_k N_k \ln N_k = - \sum_k N_k \ln \left(\frac{N_k}{N} \right). \quad (4)$$

In statistical equilibrium, the entropy S is maximized with respect to the numbers N_k under the constraints that the total number of agents $N = \sum_k N_k$ and the total money $M = \sum_k m_k N_k$ are fixed. To solve this problem, we introduce the Lagrange multipliers α and β and construct the modified entropy

$$\tilde{S} = S + \alpha \sum_k N_k - \beta \sum_k m_k N_k. \quad (5)$$

Maximization of entropy is achieved by setting the derivatives $\partial \tilde{S} / \partial N_k$ to zero for each N_k . Substituting equation (4) into (5) and taking the derivatives⁴, we find that the equilibrium probability distribution of money $P(m)$ is an exponential function of m :

$$P(m_k) = \frac{N_k}{N} = e^{\alpha - \beta m_k} = e^{-(m_k - \mu)/T}. \quad (6)$$

² Throughout the paper, we use the indices k and n to label the money bins m_k and the indices i and j to label the individual money balances m_i of the agents.

³ Note that human agents, unlike particles in quantum physics, are distinguishable.

⁴ Note that $N = \sum_k N_k$ in (4) should also be differentiated with respect to N_k .

Here the parameters $T = 1/\beta$ and $\mu = \alpha/T$ are the analogues of temperature and chemical potential for money. Their values are determined by the constraints

$$1 = \frac{\sum_k N_k}{N} = \sum_k P(m_k) = \int_0^\infty \frac{dm}{m_*} e^{-(m-\mu)/T} \Rightarrow \mu = -T \ln \left(\frac{T}{m_*} \right), \quad (7)$$

$$\langle m \rangle = \frac{M}{N} = \frac{\sum_k m_k N_k}{N} = \sum_k m_k P(m_k) = \int_0^\infty \frac{dm}{m_*} m e^{-(m-\mu)/T} = T. \quad (8)$$

We see that the money temperature $T = \langle m \rangle$ (8) is nothing but the average amount of money per agent. The chemical potential μ (7) is a decreasing function of T .

Equation (6) shows quite generally that division of a conserved limited resource using the principle of entropy maximization results in the exponential probability distribution of this resource among the agents. In physics, the ‘limited resource’ is the energy E divided between N molecules of a gas, and the result is the Boltzmann–Gibbs distribution of energy [2]. The exponential distribution of money (6) was proposed in [3], albeit without explicit discussion of the chemical potential, as well as in [4]. Various models for kinetic exchange of money are reviewed in [1] and in the popular article [5]. The applicability of the underlying assumptions of money conservation and random exchange of money is discussed in [1, 6]. The analogy between energy and money is mentioned in some physics textbooks [7], but not developed in detail.

2.2. Flow of money and people between two countries with different temperatures

To illustrate some consequences of the statistical mechanics of money, let us consider two systems with different money temperatures $T_1 > T_2$. These can be two countries with different average amounts of money per capita: the ‘rich’ country with T_1 and the ‘poor’ with T_2 .⁵ Suppose that a limited flow of money and agents is permitted between the two systems. Given that the variation $\delta\tilde{S}$ vanishes due to maximization under constraints, we conclude from (5) that

$$\delta S = \beta \delta M - \alpha \delta N \Leftrightarrow \delta M = T \delta S + \mu \delta N. \quad (9)$$

If δM and δN denote the flow of money and agents from system 1 to system 2, then the change in the total entropy of the two systems is

$$\delta S = (\beta_2 - \beta_1) \delta M - (\alpha_2 - \alpha_1) \delta N = \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \delta M + \ln \left(\frac{T_2}{T_1} \right) \delta N. \quad (10)$$

According to the second law of thermodynamics, the total entropy should be increasing, so $\delta S \geq 0$. Then, the first term in (10) shows that money should be flowing from the high-temperature system (rich country) to the low-temperature system (poor country). This is called the trade deficit—a systematic net flow of money from one country to another, which is best exemplified by the trade between USA and China. The second term in (10) shows that the agents would be flowing from high to low chemical potential, which corresponds to immigration from a poor to a rich country. Both trade deficit and immigration are widespread global phenomena. The direction of these processes can be understood also from (8). The two systems are trying to equilibrate their money temperatures $T = M/N$, which can be achieved by changing either the numerators due to money flow or the denominators due to people flow.

⁵ For simplicity, let us assume that the two countries use the same currency or the currency exchange rate is fixed, so that an equivalent currency can be used.

2.3. Thermodynamics of money and wealth

Thermal physics has two counterparts: statistical mechanics and thermodynamics. Statistical mechanics of money was outlined in section 2.1. Is it possible to construct an analogue of thermodynamics for money? Many attempts have been made in the literature, but none is completely successful: see reviews [5] and [8].

One of the important concepts in thermodynamics is the distinction between heat and work. In statistical physics, this distinction can be microscopically interpreted as follows [2]. The internal energy of the system is $U = \sum_k \varepsilon_k N_k$, where ε_k is an energy level and N_k is the occupation number of this level. Suppose the energy levels $\varepsilon_k(\lambda)$ depend on some external parameters λ , such as the volume of a box in quantum mechanics or an external magnetic field acting on spins. Then, the variation of U contains two terms $\delta U = \sum_k (\delta \varepsilon_k) N_k + \sum_k \varepsilon_k (\delta N_k) = \delta W + \delta Q$. The first term has mechanical origin and comes from the variation $\delta \varepsilon_k = (\partial \varepsilon_k / \partial \lambda) \delta \lambda$ of the energy levels due to changes in the external parameters λ . This term is interpreted as the work δW done on the system externally. The second term has statistical origin and comes from the changes δN_k in the occupation numbers of the energy levels. This term is interpreted as the heat δQ .

An analogue of this construction does not seem to exist for money $M = \sum_k m_k N_k$. A variation $\delta M = \sum_k m_k (\delta N_k)$ is possible due to changes in the occupation numbers, but there is no analogue of the variation δm_k of the ‘money levels’ due to changes in some external parameters. Thus, we can only define the heat term, but not the work term in the money variation. Indeed, (9) is the analogue of the first law of thermodynamics for money, but there is no term corresponding to work in this equation.

Nevertheless, statistical mechanics of money can be extended to a form somewhat resembling conventional thermodynamics if we take into account the material property of the agents. Let us define the wealth w_i of an agent i as the sum of two terms. One term represents the money balance m_i , and another term the material property, such as a house, a car or stocks. For simplicity, let us consider only one type of property, so that the agent has v_i physical units of this property. In order to determine the monetary value of this property, we need to know the price P per unit. Then, the wealth of the agent is $w_i = m_i + P v_i$. Correspondingly, the total wealth W in the system is⁶

$$W = M + P V, \quad (11)$$

where $V = \sum_i v_i$ is the total ‘volume’ of the property in the system. If money M is analogous to the internal energy U in statistical physics, then wealth W is analogous to the enthalpy H . The wealth W includes not only the money M , but also the money equivalent necessary to acquire the volume V of property at the price P per unit.

Let us consider the differential of wealth:

$$dW = dM + P dV + V dP = V dP. \quad (12)$$

Here, the first two terms cancel out, and only the last term remains. Indeed, when the volume $dV > 0$ of property is acquired, the money $dM = -P dV < 0$ is paid for the property, i.e. money is exchanged for property. Equation (12) is also valid at the level of individual agents, $dw_i = v_i dP$. These equations show that wealth changes only when the price P changes.

To advance the analogy with thermodynamics, let us consider a closed cycle in the (V, P) plane illustrated in figure 1. This cycle can be interpreted as a model of stock market speculation,

⁶ From now on, we use the letter W to denote wealth, not work.

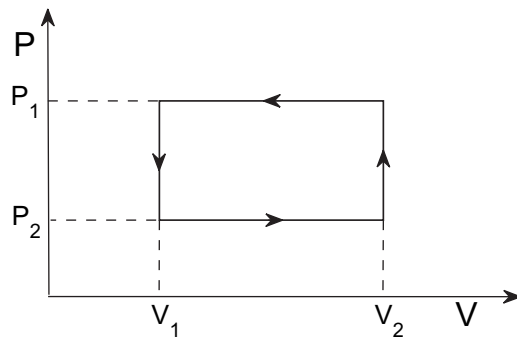


Figure 1. A closed cycle of speculation or trading. V and P represent the volume and price of goods.

in which case V is the volume of stock held by a speculator. Starting from the lower left corner, the speculator purchases the stock at the low price P_2 and increases the owned volume from V_1 to V_2 . Then, the price increases from P_2 to P_1 . At this point, the speculator sells the stock at the high price P_1 , reducing the owned volume from V_2 to V_1 . Then, the price of the stock drops to the level P_2 , and the cycle can be repeated. From (12), we find that the wealth change of the speculator is $\Delta W = \oint V dP$, which is the area $(P_1 - P_2)(V_2 - V_1)$ enclosed by the cycle in figure 1. From (11), we also find that $\Delta W = \Delta M$, because P and V return to the initial values at the end of the cycle. Thus, the monetary profit ΔM is given by the area enclosed by the cycle. This money is extracted by the speculator from the other players in the market, so the conservation law of money is not violated. In the ideal economic equilibrium, there should be no price changes allowing one to make systematic profits, which is known as the ‘no-arbitrage theorem’. However, in the real market, significant rises and falls of stock prices do happen, especially during speculative bubbles.

The cycle in figure 1 also illustrates the trade between China and the USA. Suppose that a trade company pays money M_2 to buy the volume $V_2 - V_1$ of the products manufactured in China at the low price P_2 . After shipping across the Pacific Ocean, the products are sold in the USA at the high price P_1 , and the company receives money M_1 . Empty ships return to China, and the cycle repeats. As shown in [3], the price level P is generally proportional to the money temperature T . Thus, the profit rate in this cycle is

$$\frac{M_1 - M_2}{M_2} = \frac{P_1 - P_2}{P_2} = \frac{T_1 - T_2}{T_2}. \quad (13)$$

By analogy with physics, one can prove that (13) gives the highest possible profit rate for the given temperatures T_1 and T_2 . Indeed, from (9) with $\delta N = 0$, we find that $M_1 = T_1 \Delta S_1$ and $M_2 = T_2 \Delta S_2$. Under the most ideal circumstances, the total entropy of the whole system remains constant, so $\Delta S_1 = \Delta S_2$. Then, $M_1/M_2 = T_1/T_2$, and (13) follows. Here we assumed that the profit money $M_1 - M_2$ has low (ideally zero) entropy, because this money is concentrated in the hands of just one agent or trading company and is not dispersed among many agents of the systems.

Thermal machines have cycles analogous to figure 1, and equation (13) is similar to the Carnot formula for the highest possible efficiency [2, 7]. The China–USA trade cycle resembles an internal combustion engine, where the purchase of goods from China mimics fuel intake, and the sales of goods in the USA mimics expulsion of exhaust. The net result is that goods

are manufactured in China and consumed in the USA. The analogy between trade cycles and thermal machines was highlighted by Mimkes [9, 10]. Although somewhat similar to [10], our presentation emphasizes a conceptual distinction between money and wealth, and explicitly connects statistical mechanics and thermodynamics.

Empirical data on the international trade network between different countries were analyzed in several papers. Serrano *et al* [11] analyzed trade imbalances, defined as the *difference* between exports and imports from one country to another. Their paper classified countries as net consumers and net producers of goods. The typical examples are the USA and China, respectively, as illustrated in figure 2 of [11] for 2000, in qualitative agreement with our discussion above. In contrast, Bhattacharya *et al* [12] studied trade volumes, defined as the *sum* of exports and imports from one country to another. Their paper found that the trade volume s of a country is proportional to the gross domestic product (GDP) of the country: $s \propto (\text{GDP})^\gamma$ with the exponent $\gamma \approx 1$. This means that the trade volume and GDP are *extensive* variables in the language of thermodynamics, so the biggest volumes of trade are between the countries with the biggest GDPs. In thermal equilibrium, money flows between two countries in both directions as payment for traded goods, but the money fluxes in the opposite directions are equal, so there is trade volume, but no trade imbalance. Trade imbalance may develop when the two systems have different values of *intensive* parameters, such as the money temperature. Then, the direction of net money flow is determined by the sign of the temperature difference.

Of course, there may be other reasons and mechanisms for trade imbalance besides the temperature difference. Normally, the flow of money from the high- to the low-temperature system should reduce the temperature difference and eventually bring the systems to equilibrium. Indeed, in global trade, many formerly low-temperature countries have increased their temperatures as a result of such trade. However, the situation with China is special, because the Chinese government redirects the flow of dollars back to the USA by buying treasury bills from the US government. As a result, the temperature difference remains approximately constant and does not show signs of equilibration. The net result is that China supplies vast amounts of products to the USA in exchange for debt obligations from the US government. The long-term global consequences of this process remain to be seen.

2.4. The circuit of money and the circuit of goods

Section 2.3 illustrates that there are two circuits in a well-developed market economy [10]. One is the circuit of money, which consists of money payments between the agents for goods and services. As argued in section 2.1, money is conserved in these transactions and, thus, can be modeled as flow of liquid, e.g. blood in the vascular system. (A hydraulic device, the MONIAC, was actually used by William Philips, the inventor of the famous Philips curve, to illustrate money flow in the economy [5].) The second circuit is the flow of goods and services between the agents. This circuit involves manufacturing, distribution and consumption. The goods and services are inherently not conserved. They represent the material (physical) side of the economy and, arguably, are the ultimate goal for the well-being of a society. In contrast, money represents the informational, virtual side of the economy, because money cannot be physically consumed. Nevertheless, money does play a very important role in the economy by enabling its efficient functioning and by guiding resource allocation in a society⁷. The two

⁷ Here we consider the modern fiat money, declared to be money by the central bank or government. We do not touch the origin of money in the early history as some kind of special goods.

circuits interact with each other when goods and services are traded (exchanged) for money. However, money cannot be physically transformed into goods and vice versa. To illustrate this point, we draw an analogy with fermions and bosons in physics. While the ‘circuits’ of fermions and bosons interact and transfer energy between each other, it is not possible to convert a fermion into a boson or vice versa.

An important consequence of this consideration is that an increase in material production in the circuit of goods and services does not have any direct effect on the amount of money in the monetary circuit. The amount of money in the system depends primarily on the monetary policy of the central bank or government, who have monopoly on issuing money. Technological progress in material production does not produce any automatic increase in money in the system. Thus, the expectation of continuous monetary growth, where the agents would be getting more and more money as a result of technological progress, is false. It is not possible for all businesses to operate with profit on average, i.e. to have a greater total amount of money at the end of a cycle than at the beginning. The agents can get more money on average only if the government decides to print money, i.e. to increase the money temperature $T = M/N$.⁸ Thus, monetary growth of the economy is directly related to the deficit spending by the central government. On the other hand, it is very well possible to have technological progress and an increase in the physical standards of living without monetary growth. The monetary and physical circuits of the economy interact with each other, but they are separate circuits. Unfortunately, this distinction is often blurred in the econophysics and economics literature [10], as well as in the public perception.

3. Two-class structure of income distribution

3.1. Introduction

Although the exponential probability distribution of money (6) was proposed 10 years ago [3, 4], no direct statistical data on money distribution are available to verify this conjecture. Normally, people do not report their money balances to statistical agencies. Given that most people keep their money in banks, the distribution of balances on bank accounts can give a reasonable approximation of the probability distribution of money. However, these data are privately held by banks and not available publicly.

On the other hand, a lot of statistical data are available on income distribution, because people report income to the government tax agencies. To some extent, income distribution can also be viewed as a problem of partitioning of a limited resource, in this case of the total annual budget. Following section 2.1, we expect to find the exponential distribution for income. Drăgulescu and Yakovenko [13] studied the data on income distribution in the USA from the Internal Revenue Service (IRS) and from the US Census Bureau. They found that income distribution is indeed exponential for incomes below 120 k\$ per year. However, in the subsequent papers [14, 15], they also found that the upper tail of income distribution follows a power law, as was first pointed out by Pareto [16]. So, the data analysis of income distribution in the USA reveals coexistence of two social classes. The lower class (about 97% of population) is characterized by the exponential Boltzmann–Gibbs distribution, and the upper class (the top 3% of the population) has the power-law Pareto distribution. Time evolution of the income classes in 1983–2001 was studied by Silva and Yakovenko [17]. They found that the exponential

⁸ For discussion of the issues related to debt, see the review [1].

distribution in the lower class is very stable in time, whereas the power-law distribution of the upper class is highly dynamic and volatile. They concluded that the lower class is in thermal equilibrium, whereas the upper class is out of equilibrium.

Many other papers investigated income distributions in different countries: see the review [1] for references. The coexistence of two classes appears to be a universal feature of income distribution. In this section, we present a unified description of the two classes within a single mathematical model.

3.2. Income dynamics as a combination of additive and multiplicative stochastic processes

The two-class structure of income distribution can be rationalized on the basis of a kinetic approach. Suppose the income r of an agent behaves like a stochastic variable. Let $P(r, t)$ denote the probability distribution of r at time t . Let us consider a diffusion model, where the income r changes by Δr over a time period Δt . Then, the temporal evolution of $P(r, t)$ is described by the Fokker–Planck equation [18]:

$$\frac{\partial P(r, t)}{\partial t} = \frac{\partial}{\partial r}[A(r)P(r, t)] + \frac{\partial^2}{\partial r^2}[B(r)P(r, t)]. \quad (14)$$

The coefficients $A(r)$ and $B(r)$ are the drift and diffusion terms, which are determined by the first and second moments of the income changes Δr per unit time:

$$A(r) = -\frac{\langle \Delta r \rangle}{\Delta t}, \quad B(r) = \frac{\langle (\Delta r)^2 \rangle}{2\Delta t}. \quad (15)$$

The stationary solution $P_s(r)$ of (14) satisfies $\partial_t P_s = 0$; thus we obtain

$$\frac{\partial(BP_s)}{\partial r} = -AP_s. \quad (16)$$

The general solution of (16) is

$$P_s(r) = \frac{c}{B(r)} \exp\left(-\int^r \frac{A(r')}{B(r')} dr'\right), \quad (17)$$

where c is a normalization factor, such that $\int_0^\infty P_s(r) dr = 1$.

For the lower class, the income comes from wages and salaries, so it is reasonable to assume that income changes are independent of income itself, i.e. Δr is independent of r . This process is called the additive diffusion [17]. In this case, the coefficients in (14) are some constants A_0 and B_0 . Then (17) gives the exponential distribution

$$P_s(r) = \frac{1}{T} e^{-r/T}, \quad T = \frac{B_0}{A_0}. \quad (18)$$

On the other hand, the upper-class income comes from bonuses, investments and capital gains, which are calculated in percentages. Therefore, for the upper class, it is reasonable to expect that $\Delta r \propto r$, i.e. income changes are proportional to income itself. This is known as the proportionality principle of Gibrat [19], and the process is called the multiplicative diffusion [17]. In this case, $A = ar$ and $B = br^2$, and (17) gives a power-law distribution:

$$P_s(r) \propto \frac{1}{r^{1+\alpha}}, \quad \alpha = 1 + \frac{a}{b}. \quad (19)$$

The multiplicative hypothesis for the upper-class income was quantitatively verified in [20] for Japan, where tax identification data are officially published for the top taxpayers.

The additive and multiplicative processes may coexist. For example, an employee may receive a cost-of-living raise calculated in percentages (the multiplicative process) and a merit raise calculated in dollars (the additive process). Assuming that these processes are uncorrelated, we find that $A = A_0 + ar$ and $B = B_0 + br^2 = b(r_0^2 + r^2)$, where $r_0^2 = B_0/b$. Substituting these expressions into (17), we find

$$P_s(r) = c \frac{e^{-(r_0/T) \arctan(r/r_0)}}{[1 + (r/r_0)^2]^{1+a/2b}}. \quad (20)$$

The distribution (20) interpolates between the exponential law for low r and the power law for high r , because either the additive or the multiplicative process dominates in the corresponding limit. A crossover between the two regimes takes place at $r \sim r_0$, where the additive and multiplicative contributions to B are equal. The distribution (20) has three parameters: the temperature $T = A_0/B_0$, the Pareto exponent $\alpha = 1 + a/b$ and the crossover income r_0 . It is a minimal model that captures the salient features of the two-class income distribution. A formula similar to (20) was also derived by Fiaschi and Marsili [21] for a microscopic economic model, which is effectively described by (14).

3.3. Comparison with the personal income data from the Internal Revenue Service (IRS)

In this section, we compare (20) with the annual income data from the IRS for the years 1996–2007 [22]. Because the IRS releases the data with a delay of a couple of years, 2007 is the latest year for which the data are currently available. The IRS data are given for a set of discrete income levels. Thus, it is more practical to construct the cumulative distribution function (CDF), which is the integral $C(r) = \int_r^\infty P(r') dr'$ of the probability density. For the probability density (20), $C(r)$ is not available in analytical form; therefore it has to be calculated by integrating $P_s(r)$ numerically. We use the theoretical CDF $C_t(r)$ to fit the empirical CDF $C_e(r)$ calculated from the IRS data.

Determining the best values of the three fitting parameters in the theoretical CDF is a computationally challenging task. Thus, we do it step by step. For each year, we first determine the values of T and α by fitting the low-income part of $C_e(r)$ with an exponential function and the high-income part with a power law. Then, keeping these two parameters fixed, we determine the best value of r_0 by minimizing the mean-square deviation $\sum_n \ln^2[C_t(r_n)/C_e(r_n)]$ between the theoretical and empirical functions, where the sum is taken over all income levels r_n for which empirical data are available.

Table 1 shows the values of the fit parameters obtained for different years. The data points for the empirical CDF and their fits with the theoretical CDF are shown in figure 2 in the log–log scale versus the normalized annual income r/T . For clarity, the curves are shifted vertically for successive years. Clearly, the theoretical curves agree well with the empirical data, so the minimal model (20) indeed captures the salient features of income distribution in the USA.

In previous papers [15, 17], fits of the income distribution data were made only to the exponential (18) and power-law (19) functions. The income r_* , where the two fits intersect, can be considered as a boundary between the two classes. The values of r_* are shown in table 1. We observe that the boundary r_* between the upper and lower classes is approximately 3.5 times greater than the temperature T . Given that the CDF of the lower class is exponential, we find

Table 1. T , α and r_0 are the parameters in (20), obtained by fitting the annual income data from the IRS. r_* is the income separating the upper and lower classes. f is the fraction of income going to the upper class, given by (21). G is the Gini coefficient.

Year	T (k\$)	α	r_0 (k\$)	r_* (k\$)	f (%)	G
1996	33	1.63	76	116	11.8	0.55
1997	35	1.57	79	120	13.8	0.56
1998	36	1.55	80	122	16.4	0.57
1999	38	1.54	83	124	16.8	0.58
2000	40	1.34	105	150	18.4	0.59
2001	41	1.46	99	152	14.4	0.56
2002	41	1.51	99	154	12.6	0.55
2003	41	1.48	101	156	13.7	0.56
2004	43	1.41	105	158	16.7	0.58
2005	44	1.36	108	159	19.5	0.59
2006	46	1.36	107	160	20.5	0.60
2007	48	1.34	113	166	21.5	0.60

that the upper class population is approximately $\exp(-r_*/T) = \exp(-3.5) = 3\%$, which indeed agrees with our observations.

3.4. The fraction of income in the upper tail and speculative bubbles

Let us examine the power-law tail in more detail. Although the tail contains a small fraction of population, it accounts for a significant fraction f of the total income in the system. The upper-tail income fraction can be calculated as

$$f = \frac{R - N_e T}{R} \approx \frac{R - NT}{R} = 1 - \frac{T}{\langle r \rangle}. \quad (21)$$

Here, R is the total income, N is the total number of people and $\langle r \rangle = R/N$ is the average income for the whole system. In addition, N_e is the number of people in the exponential part of the distribution and T is the average income of these people. Since the fraction of people in the upper tail is very small, we use the approximation $N_e \approx N$ in deriving the formula (21) for f . The values of f deduced from the IRS data using (21) are given in table 1.

Figure 3(c) shows the historical evolution of $\langle r \rangle$, T and f for the period 1983–2007. We see that the average income T of the lower class increases steadily without any large jumps. In contrast, the fraction f going to the upper class shows large variations and now exceeds 20% of the total income in the system. The maxima of f are achieved at the peaks of speculative bubbles, first at the end of the ‘.com’ bubble in 2000 and then at the end of the subprime mortgage bubble in 2007. After the bubbles collapse, the fraction f drops precipitously. We conclude that the upper tail is highly dynamical and out of equilibrium. The tail swells considerably during the bubbles, whereas the effect of the bubbles on the lower class is only moderate. As a result, income inequality increases during bubbles and decreases when the bubbles collapse.

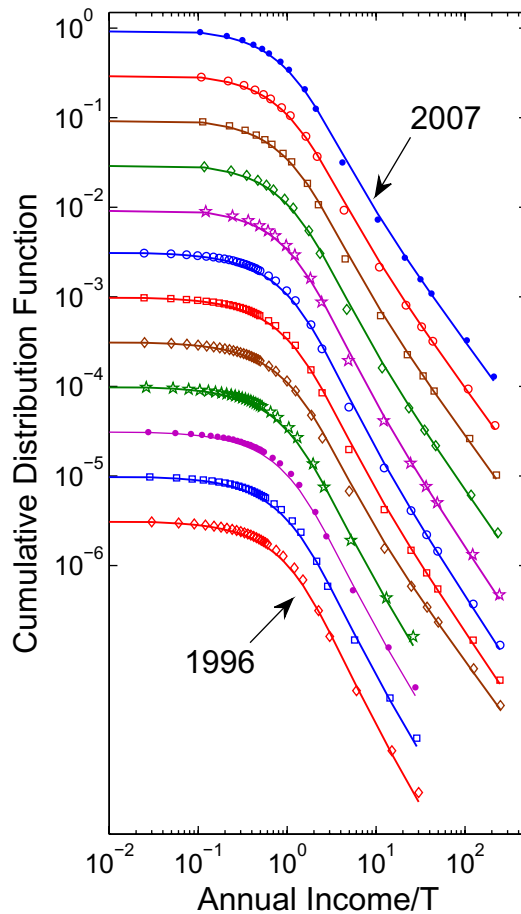


Figure 2. CDFs constructed from the IRS data (symbols) and their fits with the theoretical distribution (20), shown in the log–log scale versus the normalized annual income r/T . Plots for different years are shifted vertically for clarity.

In view of the argument about the conservation of money presented in section 2.1, what is the source of money for the enormous increase in the upper tail income during speculative bubbles? The stock market bubble in the late 1990s was actually predicted in the book [23] published in 1993. The prediction was based on the population data, showing that the demographic wave of aging baby boomers will be massively investing their retirement money in the stock market in the second half of the 1990s, which indeed happened. The stock prices rose when millions of boomers paid for the stocks of ‘.com’ companies. When the demographic wave reached its peak around 2000 and the influx of money to the stock market started to saturate (at its highest level), the market crashed precipitously, and the population was left with worthless stocks. One can see an analogy with the cycle in figure 1. The net result of this bubble is the transfer of money from the lower to the upper class under the cover of ‘retirement investment’.

As it is clear now, the second bubble in 2003–2007 was based on the enormous growth in debt due to proliferation of subprime mortgages. As discussed in [1, 3], debt can be considered as negative money, because debt liabilities are counted with a negative sign towards the net worth of an individual. The conservation law (2) is still valid, but money balances m_i can

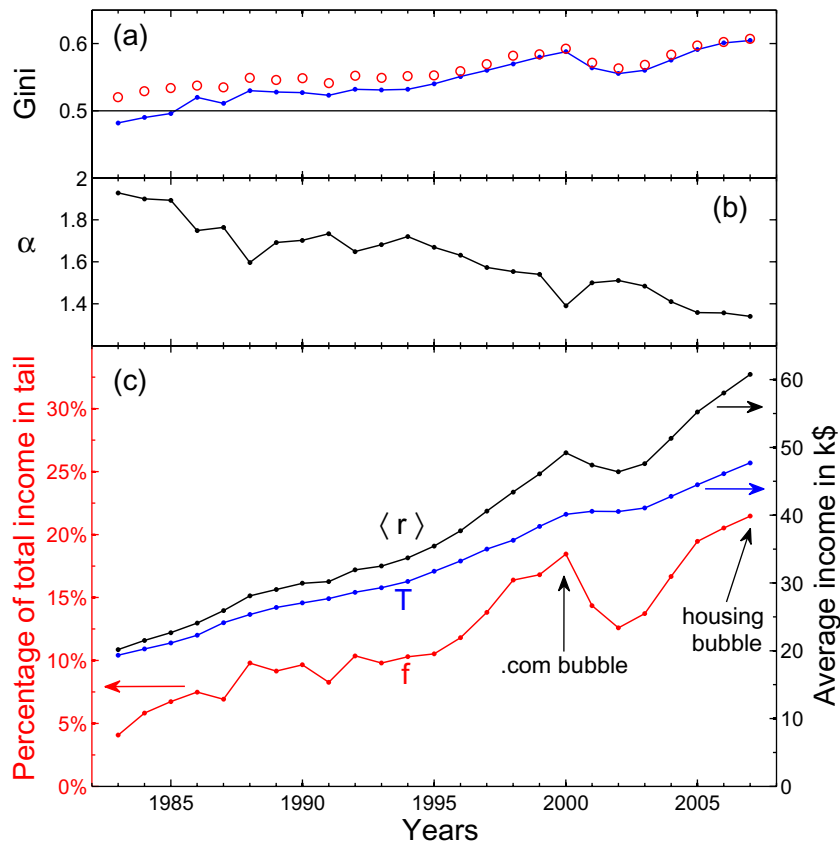


Figure 3. (a) The Gini coefficient G for the income distribution in the USA in 1983–2007 (connected line) and the theoretical formula $G = (1 + f)/2$ (open circles). (b) The exponent α of the power-law tail (19) for income distribution. (c) The average income $\langle r \rangle$ in the whole system, the average income T in the lower class (the temperature of the exponential part), and the percentage of income f going to the upper tail.

take negative values. So, the first moment (the ‘center of mass’) of the money distribution $\langle m \rangle = M/N$ remains constant. However, now some agents can become super-rich with very high positive money balances at the expense of other agents plunging deeply into debt with negative money balances. Thus, relaxing the boundary condition $m \geq 0$ undermines the stability of the Boltzmann–Gibbs distribution (6). This is what happened during the subprime mortgages bubble. The money flowing to the upper tail was coming from the growth of the total debt in the system. Eventually, the bubble collapsed when the debt reached a critical level. Now the bailout effort by the government, effectively, represents the transfer of debt from economic agents to the government. The overall result is that the income growth of the upper class in 2003–2007 was coming from the bailout money that the government is printing now. As emphasized in sections 2.1 and 2.4, the government and central bank are the ultimate sources of new money because of the government monopoly on fiat money.

The discussion and the data presented in this section indicate that, by combining the demographic data with the principle of money conservation, it may be possible to predict, to some degree, the macroeconomic behavior of the economy. In fact, the book [23] predicted in

1993 that ‘the next great depression will be from 2008 to 2023’ (page 16). This is a stunning prediction 15 years in advance of the actual event. For an update, see the follow-up book [24].

3.5. The power-law exponent of the upper tail

Another parameter of the upper tail is the power-law exponent α in (19). Table 1 and figure 3(b) show historical evolution of α from 1983 to 2007. We observe that α has decreased from about 2 to about 1.3. The decrease in α means that the power-law tail is getting ‘fatter’, i.e. the inequality of income distribution increases. It looks like the system is approaching dangerously closely the critical value $\alpha = 1$, where the total income in the tail $\int_{r_*}^{\infty} r P(r) dr$ would formally diverge [25]. On top of the gradual decrease, α dived down and up sharply around 1987 and 2000. The dive-downs of α represent sharp increases in income inequality due to the bubbles, followed by crashes of the bubbles in 1987 and 2000 and subsequent contractions of the upper tail. Thus, the behavior of the tail exponent α is qualitatively consistent with the behavior of the tail fraction f discussed in section 3.4. A similar behavior was found for Japan [20], where α jumped sharply from 1.8 to 2.1 between 1991 and 1992 due to the crash of the Japanese market bubble.

During the times of bubbles, the sharp decrease in α is clearly a dynamical process, which cannot be described adequately by stationary equations. On the other hand, during the time between bubbles, which economists may call ‘recession’ or ‘depression’, the market is quiet, and it may be possible to describe it using a stationary approach. Even during these times, the power-law tail does not disappear, but the exponent α takes a relatively high value. From figure 3(b), it appears that the upper limit for α is about 2. This limiting value is supported by other observations in the literature. Analysis of Japanese data [20] shows that α changes in the range between 1.8 and 2.2. Drăgulescu and Yakovenko [14] found $\alpha = 1.9$ for wealth distribution in the UK for 1996. Thus, we make a conjecture that $\alpha = 2$ is a special value of the power-law exponent corresponding to a quiet, stationary market.

In order to understand what is special about $\alpha = 2$, let us examine the moments of the income change Δr . The first moment $\langle \Delta r \rangle$ is always negative. This condition ensures that $A > 0$ in (15), so that (16) has a stationary solution. The condition $\langle \Delta r \rangle < 0$ indicates that, on average, everybody’s income is decreasing due to the drift term, yet the whole income distribution remains stationary because of the diffusion term. In stochastic calculus, the first $\langle \Delta r \rangle$ and the second $\langle (\Delta r)^2 \rangle$ moments are of the same order in Δt , so they must be treated on an equal footing. Thus, instead of considering the changes in r , let us discuss how r^2 changes in time. Using (15), we find

$$\langle \Delta(r^2) \rangle = \langle (r + \Delta r)^2 - r^2 \rangle = 2r \langle \Delta r \rangle + \langle (\Delta r)^2 \rangle = 2(-rA + B) \Delta t. \quad (22)$$

For the additive stochastic process (18), we find from (22) that $\langle \Delta(r^2) \rangle > 0$ for $r < T$ and $\langle \Delta(r^2) \rangle < 0$ for $r > T$. These conditions indicate a stabilizing tendency of the income-squares to move in the direction of the average income T .

Now, let us apply (22) to the multiplicative process (19). In this case, we find

$$\langle \Delta(r^2) \rangle = 2(-a + b) r^2 \Delta t. \quad (23)$$

For $a = b$, equation (23) gives $\langle \Delta(r^2) \rangle = 0$ for all r . This condition can be taken as a criterion for the inherently stationary state of a power-law tail, because r^2 does not change (on average) for any r in a scale-free manner. From (19), we observe that the condition $a = b$ corresponds to the value $\alpha = 1 + a/b = 2$, which is indeed the upper value of the power-law exponent observed

for stationary, quiet markets:

$$\langle \Delta(r^2) \rangle = 0 \quad \Leftrightarrow \quad a = b \quad \Leftrightarrow \quad \alpha = 2. \quad (24)$$

On the other hand, for $a < b$, we find $\langle \Delta(r^2) \rangle > 0$ and $\alpha < 2$. In this case, the income-square increases on average, which correlates with the upper tail expansion during the boom times. Note that the value $\alpha = 2$ in (24) is different from the value $\alpha = 1$ found for the models of random saving propensity and earthquakes in [5, 6, 26].

3.6. The Lorenz plot and Gini coefficient for income inequality

The standard way of representing income distribution in the economic literature is the Lorenz plot [27]. It is defined parametrically in terms of the two coordinates $x(r)$ and $y(r)$ depending on a parameter r :

$$x(r) = \int_0^r dr' P(r'), \quad y(r) = \frac{\int_0^r dr' r' P(r')}{\int_0^\infty dr' r' P(r')}. \quad (25)$$

Here, $x(r)$ is the fraction of the population with incomes below r , and $y(r)$ is the total income of this population, as a fraction of the total income in the system. When r changes from 0 to ∞ , the variables x and y change from 0 to 1, producing the Lorenz plot in the (x, y) -plane. The advantage of the Lorenz plot is that it emphasizes the data where most of the population is. In contrast, the log-linear and log-log plots, like figure 2, emphasize the upper tail, which corresponds to a small fraction of population, and where the data points are sparse. Another advantage of the Lorenz plot is that all available data are represented within a finite area in the (x, y) plane, whereas in other plots, the upper end of the data at $r \rightarrow \infty$ is inevitably truncated.

For the exponential distribution $P(r) = \exp(-r/T)/T$, it was shown in [13] that the Lorenz curve is given by the formula $y = x + (1-x)\ln(1-x)$. Note that this formula is independent of T . However, when the fat upper tail is present, this formula is modified as follows [15, 17]:

$$y = (1-f)[x + (1-x)\ln(1-x)] + f\Theta(x-1). \quad (26)$$

Here, $\Theta(x-1)$ is the step function equal to 0 for $x < 1$ and 1 for $x = 1$. The jump at $x = 1$ is due to the fact that the fraction of the population in the upper tail is very small, but their fraction f of the total income is substantial.

The data points in figure 4 show the Lorenz plots calculated from the IRS data for 1996 and 2007. The solid lines in figure 4 are the theoretical Lorenz curves (26) with the values of f obtained from (21). The theoretical curves agree well with the data. The distance between the diagonal line and the Lorenz curve characterizes income inequality. We observe in figure 4 that income inequality increased from 1995 to 2007, and this increase came exclusively from the growth of the upper tail, which pushed down the Lorenz curve for the exponential income distribution in the lower class.

The standard way of characterizing inequality in the economic literature [27] is the Gini coefficient $0 \leq G \leq 1$, defined as twice the area between the diagonal line and the Lorenz curve. It was shown that $G = 1/2$ for the exponential distribution [13], and

$$G = \frac{1+f}{2} \quad (27)$$

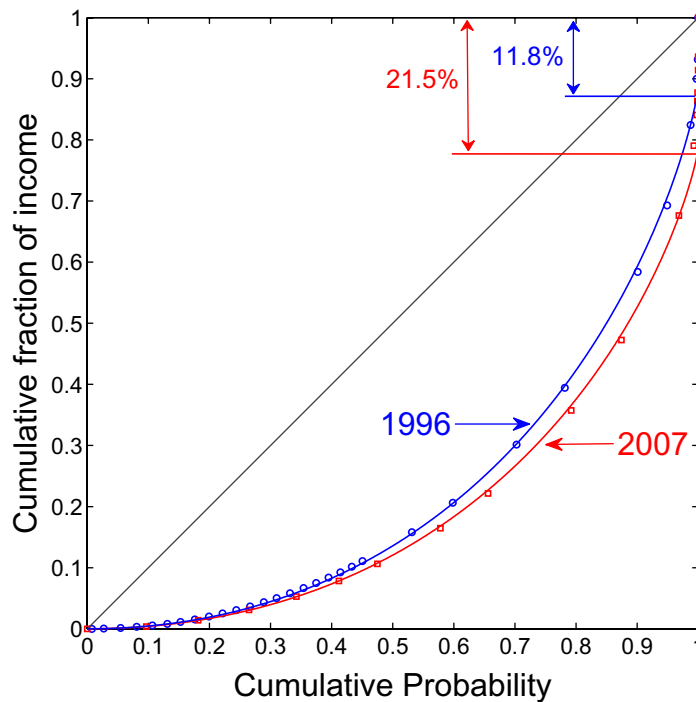


Figure 4. Lorenz plots for income distribution in 1996 and 2007. The data points are from the IRS, and the theoretical curves represent (26) with f from (21).

when taking into account the fraction f going to the upper class on top of the exponential distribution [17]. The values of G deduced from the IRS data are given in table 1 and shown in figure 3(a) by the connected line, along with (27) shown by open circles. The increase in G indicates that income inequality has been rising since 1983. The agreement between the empirical values of G and the formula (27) in figure 3 demonstrates that the increase in income inequality from the late 1990s comes from the upper tail growth relative to the lower class.

4. Probability distribution of the global energy consumption

4.1. Introduction

In the preceding sections, we studied monetary aspects of the economy and discussed probability distributions of money and income. We found that significant inequality of money and income distributions can develop for statistical reasons. Now we would like to discuss physical aspects of the economy. Since the beginning of the industrial revolution several centuries ago, rapid technological development of society has been based on consumption of fossil fuel, such as coal, oil and gas, accumulated in the Earth for billions of years. The whole discipline of thermodynamics was developed in physics to deal with this exploitation. Now it is becoming exceedingly clear that these resources will be exhausted in the not-too-distant future. Moreover, consumption of fossil fuel releases CO_2 to the atmosphere and affects the global climate. These pressing global problems pose great technological and social challenges.

As shown below, energy consumption per capita by human population around the world shows significant variation. This heterogeneity is a challenge and a complication for reaching a

global consensus on how to deal with the energy problems. Thus, it is important to understand and quantitatively characterize the global inequality of energy consumption. In this section, we present such a study using the approach developed in the preceding sections of the paper.

4.2. Energy consumption distribution as division of a limited resource

Let us consider an ensemble of economic agents and characterize each agent i by the energy consumption ε_i per unit time. Note that here ε_i denotes not energy but power, which is measured in kilowatts (kW). Similarly to section 2.1, we can discuss the probability distribution of energy consumption in the system and introduce the probability density $P(\varepsilon)$, such that $P(\varepsilon) d\varepsilon$ gives the probability of having energy consumption in the interval from ε to $\varepsilon + d\varepsilon$. Energy production, based on extraction of fossil fuel from the Earth, is physically limited. So, energy production per unit time is a limited resource, which is divided for consumption among the global population. As argued in section 2.1, it would be very improbable to divide this resource equally. More likely, this resource would be divided according to the entropy maximization principle. Following the same procedure as in section 2.1, with money m replaced by energy consumption ε , we arrive at the conclusion that the probability distribution of ε should follow the exponential law analogous to (6):

$$P(\varepsilon) \propto e^{-\varepsilon/T}, \quad T = \langle \varepsilon \rangle. \quad (28)$$

Here, the ‘temperature’ T is the average energy consumption per capita⁹.

Now we would like to compare the theoretical conjecture (28) with the empirical data for energy consumption around the world. For this purpose, it is convenient to introduce the CDF:

$$C(\varepsilon) = \int_{\varepsilon}^{\infty} P(\varepsilon') d\varepsilon'. \quad (29)$$

Operationally, $C(\varepsilon)$ is the number of agents with energy consumption above ε divided by the total number of agents in the system. If $P(\varepsilon)$ is an exponential function, then $C(\varepsilon)$ is also exponential.

4.3. Empirical data analysis

We downloaded empirical data from the World Resources Institute (WRI) website [28]. The data on energy consumption are listed under the topic ‘Energy and Resources’. We downloaded the variable ‘Total energy consumption’ [29], which contains the annual energy consumption for various countries for the years 1990, 2000 and 2005 (only these years are available). Population data are listed under the topic ‘Population, Health and Human Well-being’. We downloaded the variable ‘Total population, both sexes’ [30], which contains the total population of various countries for the same years. From these two data files, we selected the countries for which both energy and population data are available. Our final data files have 132 countries for 1990 and 135 countries for 2000 and 2005. Then we divided the annual energy consumption in a given country by the population of this country to obtain the average energy consumption per capita ε . The values of ε are listed in table 2 for some countries. A spreadsheet with our complete dataset is available for download as the supplementary online material of this paper.

Then we proceeded to construct the cumulative probability distribution for ε . First, we sorted the countries into ascending order of their energy consumption per capita ε_n , so that $n = 1$

⁹ To make it clear, this effective T is not the temperature as it is known in physics.

Table 2. Energy consumption per capita for the countries labeled in figures 5–7. The units are converted from kilotons of oil equivalent per year to kW (1 ton = 41.85×10^9 J). The last three columns show GDP per capita taken from [34].

Country	Label	Energy use (kW)			GDP per capita (k\$)		
		1990	2000	2005	1990	2000	2005
Australia	AUS	6.9	7.7	7.9	18.9	20.9	36.3
Bahrain	BHR	13.0	12.8	14.9	8.6	12.3	22.1
Brazil	BRA	1.2	1.4	1.5	3.1	3.7	4.7
Canada	CAN	10.0	10.9	11.3	21.0	23.6	35.1
China	CHN	1.0	1.2	1.7	0.3	0.9	1.7
Cuba	CUB	2.1	1.4	1.2			
France	FRA	5.3	5.8	6.0	21.9	22.4	35.0
Germany	DEU	6.0	5.6	5.6	21.6	23.1	33.7
Iceland	ISL	11.3	15.3	16.3	24.5	30.9	54.8
India	IND	0.5	0.6	0.6	0.4	0.4	0.7
Iran	IRN	1.6	2.4	3.1	2.0	1.5	2.8
Israel	ISR	3.6	4.2	3.9	11.6	19.9	19.4
Japan	JPN	4.8	5.5	5.5	24.4	36.7	35.6
Kenya	KEN	0.7	0.6	0.7	0.4	0.4	0.5
Kuwait	KWT	5.3	12.2	13.9	8.6	16.9	29.9
Mexico	MEX	2.0	2.0	2.3	3.1	5.8	7.4
Netherlands Antilles	ANT	10.4	10.2	11.9			
Russia	RUS	7.9	5.6	6.0	3.5	1.8	5.3
Arab Emirates	ARE	16.1	14.7	15.2	18.0	21.7	31.6
The United Kingdom	GBR	4.9	5.3	5.2	17.3	24.5	37.0
The United States	USA	10.0	10.8	10.4	22.5	34.3	41.3
Qatar	QAT	18.1	25.6	26.5	15.8	28.8	53.3
World average		2.2	2.2	2.3	4.2	5.2	7.0

corresponds to the country with the lowest consumption and $n = L$ to the maximal consumption, where L is the total number of countries. We denoted the population of a country n as N_n . Then, the cumulative probability for a given ε_n is

$$C_e(\varepsilon_n) = \frac{\sum_{k=n}^L N_k}{\sum_{k=1}^L N_k}. \quad (30)$$

Effectively, this construction assigns the same energy consumption ε_n to all N_n residents of the country n . Of course, this is a very crude approximation, but it is the best we can do in the absence of more detailed data. The empirically constructed function $C_e(\varepsilon_n)$ is shown in figure 5 by different colors for the years 1990, 2000 and 2005. Table 2 and figure 5 illustrate the great variation and inequality of energy consumption per capita around the world. Let us focus on the data for 2005. In the USA, ε is about 5 times greater than the global average; in China, ε is close to the global average; and in India, ε is about 1/4 of the global average.

By construction, $C_e(\varepsilon_n)$ exhibits discontinuities at each ε_n because of the approximation used in our procedure. Given the relatively small number of data points ($L = 135$) and discontinuities of the plot, it is not practical to do a quantitative fit of the data. Nevertheless,

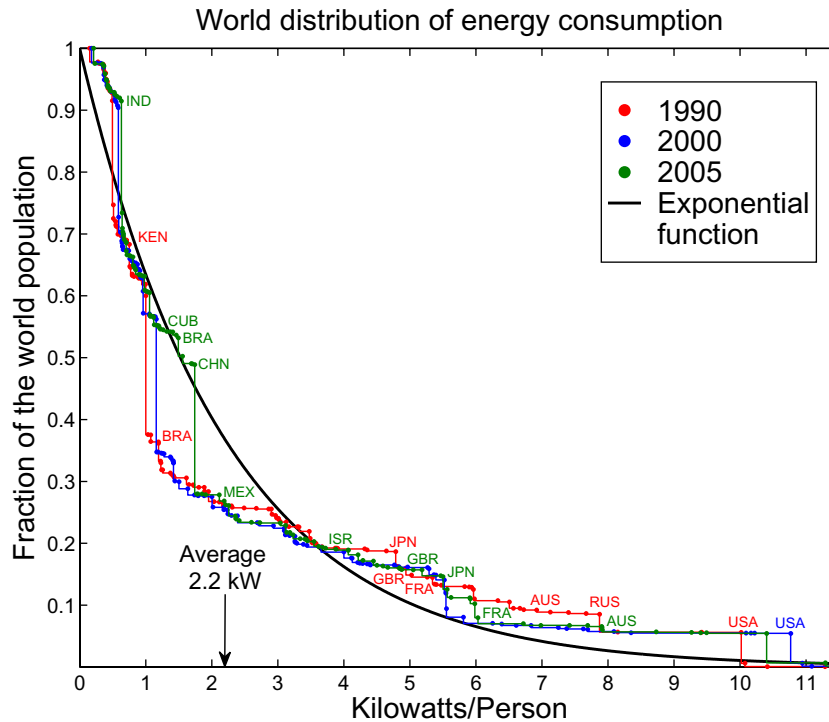


Figure 5. CDFs of energy consumption per capita around the world for 1990, 2000 and 2005. The solid curve is the exponential function.

the empirically constructed function $C_e(\varepsilon)$ can be compared with the theoretical function $C_t(\varepsilon) = \exp(-\varepsilon/T)$, which is shown by the solid line in figure 5. Here, the temperature $T = 2.2$ kW is the average global energy consumption per capita, obtained by dividing the total energy consumption of all countries by their total population. This value is indicated by the arrow in figure 5. (For comparison, the physiological energy consumption at rest by a female of weight 53 kg is 63 W [31].) The exponential function does not fit the data perfectly, but it captures the main features reasonably well, given the crudeness of the data. The agreement is remarkable, given that the solid line is not a fit, but a plot of a function with one parameter T fixed by the global average.

In order to make an additional visual comparison between the theory and the data, the functions $C_e(\varepsilon_n)$ and $C_t(\varepsilon)$ are plotted in figure 6 in the log-linear scale and in figure 7 in log-log scale. In figure 6, we see that the empirical data points oscillate around the theoretical exponential function shown by the straight line. The data jumps for high ε are unnaturally magnified in the logarithmic scale. Figure 7 demonstrates that the empirical data points do not fall on a straight line in the log-log scale, so the energy consumption per capita is not described by a power law. Indeed, energy production and consumption are physically limited and have the characteristic average scale T , so a scale-free power-law distribution would not be expected here.

We have also constructed the plots for CO₂ emission per capita using the data from WRI [28]. They look essentially the same as the plots for energy consumption per capita, in agreement with the findings by other authors [32], because most of the energy in the world is currently generated from fossil fuel.

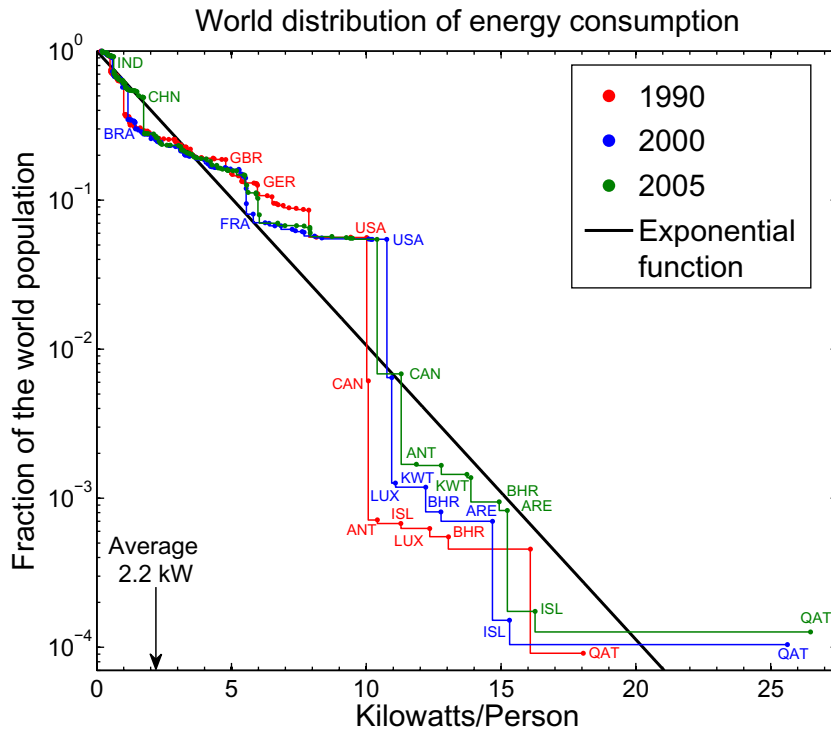


Figure 6. The same data as in figure 5, but plotted in the log-linear scale.

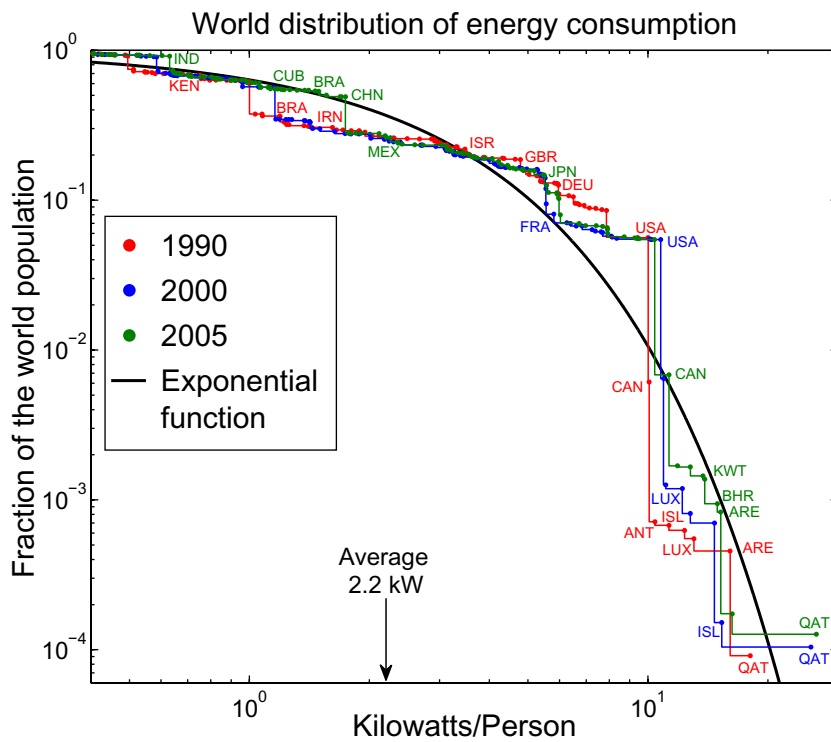


Figure 7. The same data as in figure 5, but plotted in the log-log scale.

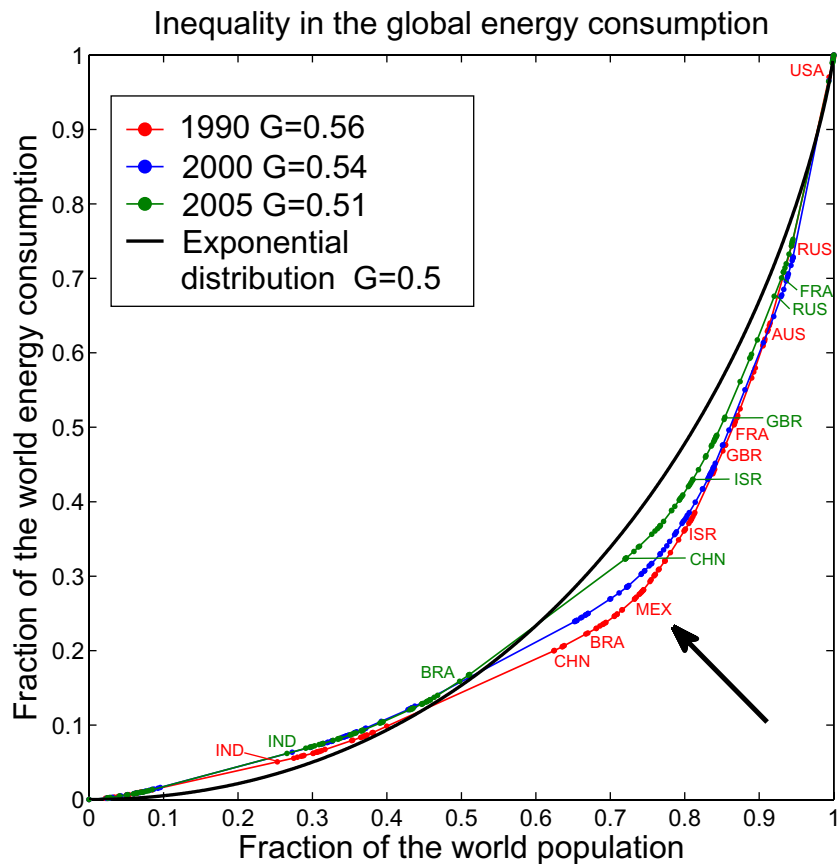


Figure 8. Lorenz plots of energy consumption per capita around the world in 1990, 2000 and 2005, compared with the Lorenz curve for the exponential distribution.

4.4. The effect of globalization on the inequality of energy consumption

Figures 5–7 give different visual representations of $C(\varepsilon)$ but have certain shortcomings. Figure 5 emphasizes the low end of the data, whereas figures 6 and 7 emphasize the high end. All figures suffer from discontinuities.

A smoother visualization can be achieved in the Lorenz plot for energy consumption per capita. As in (25), the empirical Lorenz curve is constructed parametrically:

$$x(\varepsilon_n) = \frac{\sum_{k=1}^n N_k}{\sum_{k=1}^L N_k}, \quad y(\varepsilon_n) = \frac{\sum_{k=1}^n \varepsilon_k N_k}{\sum_{k=1}^L \varepsilon_k N_k}. \quad (31)$$

The horizontal coordinate $x(\varepsilon_n)$ gives the fraction of global population with energy consumption per capita below ε , and $y(\varepsilon_n)$ gives the total energy consumption of this population as a fraction of the global consumption. When n runs from 1 to L , we obtain a set of points in the (x, y) -plane representing the Lorenz plot.

The empirically constructed Lorenz plots for 1990, 2000 and 2005 are shown in figure 8 using different colors. By construction, the Lorenz plots are continuous without jumps, although the slope (the derivative) of the $y(x)$ curve is discontinuous. Another advantage of the Lorenz

plot is that it emphasizes the data where most of the population is, i.e. the range from the bottom 5% to the top 95% of the population sorted according to their energy consumption per capita.

The black solid line shows the theoretical Lorenz curve $y = x + (1 - x)\ln(1 - x)$ for the exponential distribution. We observe that, in the first approximation, the theoretical curve captures the data reasonably well, especially given that the curve has no fitting parameters at all. Upon closer examination, we notice a systematic historical evolution of the empirical curves. From 1990 to 2005, the data points moved closer to the diagonal, which indicates that global inequality of energy consumption decreased. This is confirmed by the decrease in the calculated Gini coefficient G , which is listed in figure 8.

On the Lorenz plot for 1990, we notice a kink or a knee indicated by the arrow, where the slope of the curve changes appreciably. This point represents the boundary between developed and developing countries. Indeed, below this point we find Mexico, Brazil, China and India, whereas above this point we find Britain, France, Australia, Russia and the USA. The conclusion is that the difference between developed and developing countries lies in the degree of energy consumption and utilization. This criterion provides a physical measure for such a distinction, as opposed to more ephemeral monetary measures, such as dollar income per capita.

Comparing the Lorenz plots for 2000 and 2005 with the plot for 1990, we observe that the kink in the plots is progressively smoothed out. It means that the gap in energy consumption per capita between developed and developing countries is shrinking. We attribute this result to rapid globalization of the world economy in the last 20 years. Nevertheless, the distribution of energy consumption per capita around the world still remains highly unequal. We observe in figure 8 that the Lorenz plot has moved closer to the solid curve representing the exponential distribution. Based on the general arguments about partitioning of a limited resource, we expect that the result of a well-mixed globalized world economy would not be an equal energy consumption, but the exponential distribution. Thus, it is not likely that the energy consumption inequality will be eliminated in the foreseeable future.

It is generally known that energy consumption per capita and GDP per capita are positively correlated, and energy consumption is the physical basis for economic prosperity [32]. Brown *et al* [33] found a power-law relation $\varepsilon \propto (\text{GDP per capita})^{0.76}$ between these two variables by analyzing the data for different countries around the world (see figure 3A in [33]). The last three columns in table 2 show the data for GDP per capita [34]. Although this variable is generally correlated with the energy consumption per capita, the monetary and physical measures are not always well aligned. The movement of sustainable economics [35] criticized GDP as a useful measure of economic prosperity.

5. Conclusions

In this paper, we study probability distributions of money, income and energy consumption per capita for ensembles of economic agents. Following the principle of entropy maximization for the partitioning of a limited resource among many agents, we find exponential distributions for the investigated variables. Using an analogy with thermodynamics, we discuss trade deficit and immigration between two countries with different money temperatures. Considering a cycle similar to a thermal engine, we discuss how a monetary profit can be extracted in the presence of non-equilibrium due to a temperature difference.

Then we study a Fokker–Planck equation for income diffusion with additive and multiplicative components. The resulting probability distribution of income interpolates

between the exponential function (Boltzmann–Gibbs) at the low end and the power law (Pareto) at the high end. This function agrees well with the empirical income distribution data in the USA obtained from the IRS. While the exponential distribution in the lower class remains stable in time, the income fraction f going to the upper tail expands dramatically during speculative bubbles and shrinks when the bubbles burst. Overall, income inequality in the USA has increased significantly from 1983 to 2007, so that now f exceeds 20% of the total income in the system. We also discuss reasons why the Pareto exponent tends to have the value about $\alpha = 2$ in the steady state in the absence of bubbles.

Finally, we analyze the probability distribution of energy consumption per capita around the world using the data from the WRI. We find that the distribution is reasonably described by the exponential function with the average global consumption as the effective temperature. A closer examination finds a gap in energy consumption between developed and developing countries, which tends to shrink as time progresses. We attribute this effect to globalization of the world economy. The inequality of energy consumption decreased from 1990 to 2005, while the corresponding Lorenz plot moved closer to the exponential distribution.

In conclusion, we observe that statistical problems of different nature have a common mathematical description and exhibit similar and universal patterns of inequality. Thus, the statistical approach gives an insight into the persistent and ubiquitous nature of inequality in the world around us. The approach presented here can also be applied to other statistical problems.

Acknowledgments

We are grateful to Doyne Farmer for hospitality at the Santa Fe Institute, where a part of this work was performed, and to the participants of the working group ‘Universal Diversity Patterns Across the Sciences’ at the Santa Fe Institute in February 2009. VMY thanks Ted Jacobson for a stimulating discussion.

References

- [1] Yakovenko V M and Rosser J B 2009 Colloquium: statistical mechanics of money, wealth, and income *Rev. Mod. Phys.* **81** 1703
- [2] Wannier G H 1987 *Statistical Physics* (New York: Dover)
- [3] Drăgulescu A A and Yakovenko V M 2000 Statistical mechanics of money *Eur. Phys. J. B* **17** 723
- [4] Mimkes J 2000 Society as a many-particle system *J. Therm. Anal. Calorimetry* **60** 1055
Mimkes J and Willis G 2005 Lagrange principle of wealth distribution *Econophysics of Wealth Distributions* ed A Chatterjee, S Yarlagadda and B K Chakrabarti (Milan: Springer) pp 61–9
- [5] Sinha S and Chakrabarti B K 2009 Towards a physics of economics *Physics News (Bulletin of Indian Physics Association)* vol 39 no. 2 (April) pp 33–46
- [6] Chatterjee A, Sinha S and Chakrabarti B K 2007 Economic inequality: is it natural? *Curr. Sci.* **92** 1383
- [7] Schroeder D V 2000 *An Introduction to Thermal Physics* (New York: Addison-Wesley Longman) p 89
Yoshioka D 2007 *Statistical Physics: An Introduction* (Berlin: Springer) p 25
Saha M and Srivastava B N 1931 *A Treatise on Heat* (Allahabad: Indian Press) p 105
- [8] Mirowski P 1989 *More Heat than Light: Economics as Social Physics, Physics as Nature’s Economics* (Cambridge: Cambridge University Press)
- [9] Mimkes J and Aruka Y 2005 Carnot process of wealth distribution *Econophysics of Wealth Distributions* ed A Chatterjee, S Yarlagadda and B K Chakrabarti (Milan: Springer) pp 70–8

- Mimkes J 2006 A thermodynamic formulation of economics *Econophysics and Sociophysics: Trends and Perspectives* ed B K Chakrabarti, A Chakraborti and A Chatterjee (Berlin: Wiley-VCH) pp 1–33
- [10] Mimkes J 2010 Stokes integral of economic growth: calculus and the Solow model *Physica A* **389** 1665
- [11] Serrano M A, Boguñá M and Vespignani A 2007 Patterns of dominant flows in the world trade web *J. Econ. Interact. Coord.* **2** 111
- [12] Bhattacharya K, Mukherjee G, Saramäki J, Kaski K and Manna S S 2008 The International Trade Network: weighted network analysis and modelling *J. Stat. Mech.* **P02002**
- [13] Drăgulescu A A and Yakovenko V M 2001 Evidence for the exponential distribution of income in the USA *Eur. Phys. J. B* **20** 585
- [14] Drăgulescu A A and Yakovenko V M 2001 Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States *Physica A* **299** 213
- [15] Drăgulescu A A and Yakovenko V M 2003 Statistical mechanics of money, income, and wealth: a short survey *Modeling of Complex Systems: Seventh Granada Lectures (AIP Conf. Proc. vol 661)* ed P L Garrido and J Marro (New York: AIP) p 180
- [16] Pareto V 1897 *Cours d'Économie Politique* (L'Université de Lausanne)
- [17] Silva A C and Yakovenko V M 2005 Temporal evolution of the 'thermal' and 'superthermal' income classes in the USA during 1983–2001 *Europhys. Lett.* **69** 304
- [18] Lifshitz E M and Pitaevskii L P 1981 *Physical Kinetics* (Oxford: Pergamon)
- [19] Gibrat R 1931 *Les Inégalités Economiques* (Paris: Sirely)
- [20] Fujiwara Y, Souma W, Aoyama H, Kaizoji T and Aoki M 2003 Growth and fluctuations of personal income *Physica A* **321** 598
- Aoyama H, Souma W and Fujiwara Y 2003 Growth and fluctuations of personal and company's income *Physica A* **324** 352
- [21] Fiaschi D and Marsili M 2009 Economic interactions and the distribution of wealth arXiv:0906.1512
- [22] Personal income data from the Statistics of Income division of IRS <http://www.irs.gov/taxstats/indtaxstats/article/0,,id=134951,00.html>
- [23] Dent H S 1993 *The Great Boom Ahead* (New York: Hyperion)
- [24] Dent H S 2009 *The Great Depression Ahead* (New York: Free Press)
- [25] Anderson P W A 1997 Some thoughts about distribution in economics *The Economy as an Evolving Complex System II* ed W B Arthur, S N Durlauf and D A Lane (Reading, MA: Addison-Wesley) pp 565–6
- [26] Bhattacharyya P, Chatterjee A and Chakrabarti B K 2007 A common mode of origin of power laws in models of market and earthquake *Physica A* **381** 377
- [27] Kakwani N 1980 *Income Inequality and Poverty* (Oxford: Oxford University Press)
- [28] World Resources Institute data, <http://earthtrends.wri.org/>
- [29] Total energy consumption, http://earthtrends.wri.org/searchable_db/index.php?theme=6&variable_ID=267&action=select_countries
- [30] Total population, both sexes, http://earthtrends.wri.org/searchable_db/index.php?theme=4&variable_ID=363&action=select_countries
- [31] Defilla S 2007 A natural value unit—Econophysics as arbiter between finance and economics *Physica A* **382** 42
- [32] MacKay D J C 2009 *Sustainable Energy—Without the Hot Air* (Cambridge: UIT)
- [33] Brown J H *et al* 2010 Energetic limits to economic growth *Proc. Nat. Acad. Sci. USA* at press
- [34] Link to the GDP data, http://earthtrends.wri.org/searchable_db/index.php?theme=5&variable_ID=224&action=select_countries
- [35] Daly H 1996 *Beyond Growth: The Economics of Sustainable Development* (Boston: Beacon Press)