Economic inequality: Is it natural?

Arnab Chatterjee1,*, Sitabhra Sinha2 and Bikas K. Chakrabarti1

1Theoretical Condensed Matter Physics Division and Centre for Applied Mathematics and Computational Science, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700 064, India
2Institute of Mathematical Sciences, C.I.T. Campus, Taramani, Chennai 600 113, India

Mounting evidences are being gathered suggesting that income and wealth distribution in various countries or societies follows a robust pattern, close to the Gibbs distribution of energy in an ideal gas in equilibrium, but also deviating significantly for high-income groups. Application of physical models seems to provide illuminating ideas and understanding, complementing the observations.

Keywords: Asset exchange model, kinetic theory, Pareto law, wealth distribution.

We are all aware of the hard fact: neither wealth nor income is ever uniform for us all. Justified or not, they are unevenly distributed; few are rich, many are poor! Such socio-economic inequalities seem to be a persistent fact of life ever since civilization began. Can it be that it only reflects a simple natural law, understandable from the application of physics?

Income and wealth distribution in society

Investigations over more than a century and the recent availability of electronic databases of income and wealth distribution (ranging from national sample survey of household assets to the income tax return data available from governmental agencies) have revealed some remarkable features. Irrespective of many differences in culture, history, social structure, indicators of relative prosperity (such as gross domestic product or infant mortality) and, to some extent, the economic policies followed in different countries, income distribution seems to follow a particular universal pattern, as does wealth distribution: After an initial rise, the number density of people rapidly decays with their income, the bulk described by a Gibbs or log-normal distribution crossing over at the very high income range (for 5–10% of the richest members of the population) to a power law with an exponent (known as Pareto exponent) value between 1 and 3. This seems to be an universal feature: from ancient Egyptian society1 through nineteenth century Europe2,3 to modern Japan4,5. The same is true across the globe today: from the advanced capitalist economy of USA4,5 to the developing economy of India6.

The power-law tail, indicating a much higher frequency of occurrence of very rich individuals (or households) than would be expected by extrapolating the properties of the bulk of the distribution, was first observed by Vilfredo Pareto2 in the 1890s for income distribution of several societies at very different stages of economic development. Later, wealth distribution was also seen to follow similar behaviour. Subsequently, there have been several attempts starting around the 1950s, mostly by economists, to explain the genesis of the power-law tail (for a review, see Champernowne3). However, most of these models involved a large number of factors that made understanding the essential reason behind the occurrence of inequality difficult. Following this period of activity, a relative lull followed in the 70s and 80s when the field lay dormant, although accurate and extensive data were accumulated that would eventually make possible precise empirical determination of the distribution properties. This availability of large quantity of electronic data and their computational analysis has led to a recent resurgence of interest in the problem, specifically over the last one and half decades.

Although Pareto2 and Gini7 had respectively, identified the power-law tail and the log-normal bulk of income distribution, demonstration of both features in the same distribution was possibly first by Montroll and Shlesinger8, through an analysis of fine-scale income data obtained from the US Internal Revenue Service (IRS) for the year 1935–36. It was observed that while the top 2–3% of the population (in terms of income) followed a power law with Pareto exponent $\nu = 1.63$, the rest followed a log-normal distribution. Later work on Japanese personal income data based on detailed records obtained from the Japanese National Tax Administration9 indicated that the tail of the distribution followed a power law with $\nu$ value that fluctuated from year to year around the mean value of 2. Further work10 showed that the power law region described the top 10% or less of the population (in terms of income), while the remaining income distribution was well described by the log-normal form. While the value of $\nu$ fluctuated significantly from year to year, it was observed that the parameter describing the log-normal bulk, the Gibrat index, remained relatively unchanged. The change of income from year to year, i.e. the growth rate as measured by the log ratio of the income tax paid in successive years, was observed by Fujiwara et al.11 to be

*For correspondence. (e-mail: arnab.chatterjee@saha.ac.in)
also a heavy-tailed distribution, although skewed, and centred about zero. Later work on the US income distribution based on data from IRS for the period 1997–98, while still indicating a power-law tail (with \( \nu \approx 1.7 \)), has suggested that the lower 95% of the population has income whose distribution may be better described by an exponential form\(^ {12,13} \). The same observation has been made for income distribution in the UK for the period 1994–99, where the value of \( \nu \) was found to vary between 2.0 and 2.3, but the bulk seemed to be well described by an exponential decay.

It is interesting to note that when one shifts attention from the income of individuals to the income of companies, one still observes the power-law tail. A study of the income distribution of Japanese firms\(^ {14} \) concluded that it follows a power law with \( \nu = 1 \), which is also often referred to as the Zipf's law. Similar observation has been reported for the income distribution of the US companies\(^ {15} \).

Compared to the empirical work done on income distribution, relatively few studies have looked at the distribution of wealth, which consists of the net value of assets (financial holdings and/or tangible items) owned at a given point in time. Lack of an easily available data source for measuring wealth, analogous to income tax returns for measuring income, means that one has to resort to indirect methods. Levy and Solomon\(^ {16} \) used a published list of wealthiest people to generate a rank-order distribution, from which they inferred the Pareto exponent for wealth distribution in USA. An alternative technique was used based on adjusted data reported for the purpose of inheritance tax to obtain the Pareto exponent for the UK\(^ {13,17} \). Another study used tangible asset (namely house area) as a measure of wealth to obtain the wealth distribution exponent in ancient Egyptian society during the reign of Akhenaten (14th century BC)\(^ {1} \). More recently, wealth distribution in India at present has also been observed\(^ {6} \) to follow a power-law tail with the exponent varying around 0.9. The general feature observed in the limited empirical study of wealth distribution is that of a power-law behaviour for the wealthiest 5–10% of the population, and exponential or log-normal distribution for the rest of the population. The Pareto exponent as measured from the wealth distribution is found to be always lower than the exponent for income distribution, which is consistent with the general observation that, in market economies, wealth is much more unequally distributed than income\(^ {18} \).

The striking regularities (Figure 1) observed in income distribution for different countries, have led to several new attempts at explaining them on theoretical grounds. Much of the current impetus is from physicists' modelling of economic behaviour in analogy with large systems of interacting particles, as treated, for example, in the kinetic theory of gases. According to physicists working on this problem, the regular patterns observed in the income (and wealth) distribution may be indicative of a natural law for the statistical properties of a many-body dynamical system representing the entire set of economic interactions in a society, analogous to those previously derived for gases and liquids. By viewing the economy as a thermodynamic system, one can identify income distribution with the distribution of energy among particles in a gas. In particular, a class of kinetic exchange models have provided a simple mechanism for understanding the unequal accumulation of assets. Many of these models, while simple from the perspective of economics, have the benefit of coming to grips with the key factor in socio-economic interactions that results in different societies converging to similar forms of unequal distribution of resources (see Chatterjee et al.\(^ {4} \) and Chakrabarti et al.\(^ {5} \) for a collection of large number of technical papers in this field; see also Hayes\(^ {19} \), Hogan\(^ {20} \) and Ball\(^ {21} \) for some popular discussions and also criticisms).

**Box 1.** Income inequality: Gini coefficient and Pareto law.

\( \text{a)} \) The Gini coefficient \( G \) gives a measure of inequality in any income distribution and is defined as the proportional area between the Lorenz curve \( (I, \text{giving the cumulative fraction of the people with the fraction of wealth}) \) and the perfect equality curve \( (E, \text{where the fraction of wealth possessed by any fraction of population would be strictly linear}) \): \( G = 1 - (A_I / A_E) \), where \( A_I \) and \( A_E \) are the areas under curves \( I \) and \( E \) respectively.

\( \text{b)} \) When one plots the cumulative wealth (income) distribution against the wealth (income), almost 90–95% of the population fits the Gibbs distribution (indicated by the shaded region in the distribution; often fitted also to lognormal form) and for the rest (very rich) 5–10% of the population in any country, the number density falls off with their wealth (income) much slowly, following a power law, called the Pareto law. The second part of this law, which we do not discuss here, states that about 40–60% of the total wealth of any economy is possessed by 5–10% of the people in the Pareto tail. Although this seems to be qualitatively true, we do not have any recent data to support it.
Figure 1. a, Cumulative probability \(Q(m)\) of annual income \(m\) of the US personal for IRS data of 2001 (from Silva et al. in ref. 4), Pareto exponent \(\nu \approx 1.5\) (given by the slope of the solid line). b, Cumulative income distribution in India during 1929–30, collected from Income Tax and Super Tax data. (Inset) Cumulative distribution of employment income for the top 422 salaried Indians (Business Standard Survey, 2006) showing a power-law tail with \(\nu = 1.75 \pm 0.01\) (in spite of the best of our efforts in collecting the equivalent data from the Income Tax Department of the Government of India or the Reserve Bank of India, we are unable to give or compare with any better data). c, Cumulative probability distribution of Japanese personal income in the year 2000. The power-law region approximately fits to \(\nu = 1.96\) (data from Fujiwara et al. in ref. 4). d, Cumulative probability distribution of firm size (total assets) in France in the year 2001 for 669,620 firms. The power-law region approximately fits to \(\nu = 0.84\) (data from Fujiwara et al. in ref. 4).

A simple ideal gas-like model

Think of an exchange game like the following in an economy where the different commodities are not being explicitly considered, but rather their value in terms of an uniform asset (money) is considered. In such an asset exchange game, there are \(N\) players participating, with each player having an initial capital of one unit of money. \(N\) is very large, and the total money \(M = N\) remains fixed over the game as does the number of players.

(a) In the simplest version, the only allowed move at any time is that two of these players are randomly chosen and they decide to divide their pooled resources randomly among themselves. As no debt is allowed, none of the players can end up with a negative amount of assets. As one can easily guess, the initial delta function distribution of money (with every player having the same amount) gets destabilized with such moves and the state of perfect equality, where every player has the same amount, disappears quickly. Let us ask, what will be the eventual steady-state distribution of assets among the players after many such moves? The answer is well established in physics for more than a century – soon, there will be a stable asset distribution and it will be the Gibbs distribution: \(P(m) \sim \exp[-m/T]\), where the parameter \(T = M/N\) corresponds to the average money owned by an agent^{12,22,23}.

(b) Now think of a modified move in this game: each player ‘saves’ a fraction \(\lambda\) of his/her total assets during every step of the game, from being pooled, and randomly divided the rest with the other (randomly chosen) player. If everybody saves the same fraction \(\lambda\), what is the steady-state distribution of assets after a large number of such moves?
REVIEWS

It becomes Gamma-function like\(^{24,25}\), whose parameters depend on \(\lambda: P(m) \sim m^{\alpha} \exp[-m/T(\lambda)]; \alpha = 3\lambda(1 - \lambda).\) Angle\(^{26,27}\), utilizing a different stochastic model, arrived at somewhat similar (numerical) results, considerably earlier. Although qualitative explanation and limiting results for \(\lambda \to 0\) or \(\lambda \to 1\) are easy to obtain, no exact treatment of this problem is available so far.

(c) What happens to the steady-state asset distribution among these players if \(\lambda\) is not the same for all players, but is different for different players? Let the distribution \(\rho(\lambda)\) of saving propensity \(\lambda\) among the agents be such that \(\rho(\lambda)\) is non-vanishing when \(\lambda \to 1\). The actual asset distribution in such a model will depend on the saving propensity distribution \(\rho(\lambda)\), but for all of them the asymptotic form of distribution will become Pareto-like\(^{28-30}\): \(P(m) \sim m^{-(1+\lambda)}\), \(\nu = 1\) for \(m \to \infty\). This is valid for all such distributions\(^{29}\) (unless \(\rho(\lambda) \propto (1 - \lambda)^{\delta}\) when \(P(m) \sim m^{-(1+2\delta)}\)). However, for variation of \(\rho(\lambda)\) such that \(\rho(\lambda) \to 0\) for \(\lambda < \lambda_0\), one will get an initial Gamma function form for \(P(m)\) for small and intermediate values of \(m\), with parameters determined by \(\lambda_0\) (\(\neq 0\)), and this distribution will eventually become Pareto-like for \(m \to \infty\) with \(\nu = 1\) (Figure 2; cf. refs 28–30). Analytical understanding is now available\(^{31,32}\) and a somewhat rigorous analytical treatment of this problem has been given\(^{49}\).

It may be mentioned that there are a large number of random multiplicative asset exchange models\(^{34,35}\) to explain the Pareto (power-law) tail of wealth or income distribution. The advantage of the kind of model discussed above is that it can accommodate all the essential features of \(P(m)\) for the entire range of \(m\), not only the Pareto tail.

(d) One can of course argue that the random division of pooled assets among players is not a realistic approximation of actual trading carried out in society. As Hayes\(^1,9\) points out, in most exchanges between an individual and a large company, it is unlikely that the individual will end up with a significant fraction of the latter’s assets. Strict enforcement of this condition leads to a new type of game, the minimum exchange model, where the maximum amount that can change hands over a move, is a fraction of the poorer player’s assets. Although the change in the rules does not seem significant from the simple random exchange game, the outcome is astonishingly different: in the steady state, one player ends up with all the assets.

If we now relax the condition that the richer player does not completely dictate the terms of exchange, so that the amount exchanged need not be limited by the total asset owned by the poorer player, we arrive at a game which is asymmetric in the sense of generally favouring the player who is richer than the other, but not so much that the richer player dominates totally. Just like the previously defined savings propensity for a player, one can now define ‘thrift’ \(\tau\), which measures the ability of a player to exploit its advantage over a poorer player\(^2\). For the two extreme cases of minimum (\(\tau = 0\)) and maximum (\(\tau = 1\)) thrift, one gets back the random asset exchange and minimum asset exchange models respectively. However, close to the maximum limit, at the transition between the two very different steady-state distributions given by the two models, we see a power-law distribution! As in the case of \(\lambda\), we can now consider the case when instead of having the same \(\tau\), different players are endowed with different thrift abilities. For such heterogeneous thrift assignment in the population, where \(\tau\) for each player is chosen from a random distribution, the steady-state distribution reproduces the entire range of observed distributions of income (as

Box 2. Kinetic theory of ideal gas: Gibbs and Maxwell–Boltzmann distributions.

In a classical ideal gas in thermodynamic equilibrium, the state variables like pressure \((P)\), volume \((V)\) and the absolute temperature \((T)\) maintains a very simple relationship \(PV = NkT\). Here \(N\) is the number of basic constituents (atoms or molecules; \(N \sim N\)) and \(k\) is a constant called Boltzmann constant. Statistical mechanics of ideal gas, also called the kinetic theory of gas, intends to explain the above gas law in terms of the constituents’ mechanics or kinetics. According to this picture, for a classical ideal gas, each constituent is a Newtonian particle and they undergo random elastic collisions (which conserve kinetic energy \(E\)) among themselves and the walls of the container. These collisions eventually set up a non-uniform (kinetic) energy distribution \(\mathcal{D}(E)\) among the constituents, called the Maxwell–Boltzmann distribution: \(\mathcal{D}(E) = f(E)g(E), g(E) = \exp[-E/kT]\) for a perfect gas in a 3-dimensional container (proportional to the square root of the momentum), and the temperature \(T\) as the average (over the distribution \(\mathcal{D}(E)\) energy, one immediately gets the above mentioned gas law (relating \(P\), \(V\) and \(T\).
The trading markets can be easily modelled to be composed of two-body scatterings. The money $m_i(t)$ of an agent $i$ at time $t$ changes due to trading/scattering with a random agent $j$ in the market; the scattering locally conserves the total money. Each agent saves a fraction $\lambda_i$ of its money $m_i(t)$ at that time $t$ and the same is true for the other, and the rest of the money $(1 - \lambda_i)m_i(t) + (1 - \lambda_j)m_j(t)$ is shared randomly ($\epsilon$ is a random fraction between 0 and 1). We assume $\epsilon$ to be an annealed variable (changes with trading or time), while $\lambda$ are quenched variables (do not change with time).

For uniform $\lambda$, a Gamma distribution $P(m)$ for money occurs. For a white distribution of $\lambda$, a Pareto law $P(m) \sim m^{-2}$ (i.e. Pareto exponent $\nu = 1$) sets in.

Asset distribution in the asymmetric asset exchange game where the players have different thrift values (randomly chosen from a uniform distribution over the unit interval) also exhibits a power-law tail, with Pareto exponent $\nu = 1.5$. In comparing with the cumulative probability $Q(m)$ in Figure 1, one should note that $Q(m)$ is given by $\int_{\lambda_0}^{\lambda} P(m) \, dn$.

well as wealth) in the society: the tail follows a power law, while the bulk is described by an exponential distribution. The tail exponent depends on the distribution of $\tau$, with the value of $\nu = 1.5$ suggested originally by Pareto, obtained for the simplest case of uniform distribution of $\tau$ between $[0, 1]$ (Figure 2 d). However, even extremely different distributions of $\tau$ (e.g. U-shaped) always produce a power-law tailed distribution that is exponentially decaying in the bulk, underlining the robustness of the model in explaining inequality.

An extension

A major limitation of these asset exchange models considered earlier (and summarized above) is that it does not make any explicit reference to the commodities exchanged whose asset values we were considering so far and to the constraints they impose. We have also studied the effect of explicitly introducing a single non-consumable commodity (which is bought and sold in terms of money) on the asset distributions in the steady state. Here again two of the agents are arbitrarily chosen for interaction (or trading) and the commodity exchanged for money, provided of course the two agents have the required amounts of commodity and money (since no credit purchases are allowed). Otherwise, no exchange takes place and a new pair of agents is chosen. The global price of the commodity (ratio of total money to total amount of commodity in the market) is normalized but has temporal fluctuations. Here, we distinguish between money and wealth; wealth of any agent is composed of money and the money equivalent of the commodity with the agent. In spite of many significant effects, the general feature of Gamma-like form of
the asset distributions (for uniform $\lambda$) and the power-law tails (for random $\lambda$) for both money and wealth, with identical exponents, are seen to remain unchanged.

These studies indicate that the precise studies (theories) for asset exchange models are extremely useful and relevant. Also this helps address the question of identifying a money-like asset with wealth in simple asset exchange models and suggests that the absurd simplicity can be relaxed, yet the quantitative features are not affected.

Relevance of gas-like models

All these gas-like models of trading markets are based on the assumption of (a) asset conservation (globally in the market; as well as locally in any trading) and (b) stochasticity. Questions on the validity of these points are natural and have been raised. We now forward some arguments in their favour.

Asset conservation

If we view the trading as scattering processes, one can see the equivalence. Of course, in any such ‘asset exchange’ trading process, one receives some profit or service from the other and this does not appear to be completely random, as assumed in the models. However, if we concentrate only on the ‘cash’ exchanged (even using bank cards!), every trading is an asset conserving one (like the elastic scattering process in physics!) As discussed earlier, conservation of asset can be extended to that of total wealth (including money) and relaxed, as given by the temporally fluctuating price (effectively allows for slight relaxation over this conservation), yet keeping the overall distribution same (with unchanged $\nu$ value). It is also important to note that the frequency of asset exchange in such models defines a timescale in which total asset in the market does not change. In real economies, total asset changes much slowly, so that in the timescale of exchanges, it is quite reasonable to assume the total asset to be conserved in these exchange models.

Stochasticity

Is the trading random? Surely not, when looked from an individual’s point of view. When one maximizes his/her utility by money exchange for the $p$th commodity, he/she may choose to go to the $q$th agent and for the $r$th commodity he/she will go to the $s$th agent. But since $p \neq q \neq r \neq s$ in general, when viewed from a global level, these trading/scattering events will all look random (although for individuals this is a defined choice or utility maximization). It may be noted in this context that in the stochastically formulated ideal gas models in physics (developed in late 1800/early 1900), physicists already knew for more than a hundred years that each of the constituent particles (molecules) follows a precise equation of motion, namely that due to Newton. The assumption of stochasticity in asset-exchange models, even though each agent might follow an utility maximizing strategy (like Newton’s equation of motion for molecules), is therefore not unusual in the context.

Support from economic data

Analysis of high-quality income data from the UK and USA shows peaked Gamma distributions for the low- and middle-income ranges, which suggests a strong case in favour of models discussed earlier. This has already been seen in studies of isolated groups of similar individuals, and has been modelled in a similar fashion.

Concluding remarks

The enormous amount of data available on the income and wealth distribution of various countries clearly establishes a robust feature: Gamma (or log-normal) distribution for the majority (almost 90–95%), followed by a Pareto power law (for the richest 5–10% of the population), as seen in Figure 1. We show that this ‘natural’ behaviour of income inequality comes from a simple ‘scattering picture’ of the market (Figure 2 a), when the agent in the market has got random saving propensity. Models studied in physics (in the kinetic theory of gases) more than a hundred years ago, help in formulating and understanding the ‘natural’ behaviour of the markets.

References


Received 29 January 2007; revised accepted 20 February 2007